DreamBox Math (2022–23)

Study Type: ESSA Evidence Level III

Prepared for:
Discovery Education

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

July 1, 2024

EXECUTIVE SUMMARY

Discovery Education 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 III requirements (Promising Evidence) according to the Every Student Succeeds Act (ESSA).

Study Sample, Measures, and Methods

This study occurred during the 2022–23 school year and included 57,146 K–8 students from across 161 schools in one district. Researchers conducted analyses by grade band to allow for better interpretability of findings: Kindergarten (16%); grades 1–2 (33%); grades 3–5 (45%); and grades 6–8 (6%). In terms of demographics, the total sample was racially diverse and included Black/African American (34%), Hispanic (31%), White (24%), Asian (7%), and mixed race (4%) students. The district did not provide data on students' socioeconomic status (e.g., low-income indicator, or free and reduced lunch status). Forty-nine percent of students in the sample identified as female.

Researchers used three measures to provide insights into DreamBox implementation and potential impacts of DreamBox on student learning outcomes: DreamBox usage data, NWEA MAP® mathematics percentile scores, and end-of-grade (EOG) state mathematics assessment scores. Researchers used two-level multilevel modeling analysis (i.e., students nested in schools) to examine how DreamBox Math use related to student math end-of-year outcomes controlling for their beginning-of-year scores. The analyses included student-level covariates to control for potential selection bias (e.g., grade-level, race, and gender). In addition, researchers calculated standardized effect sizes (Hedges' *g*) to determine the magnitude of changes in student outcomes.

Student Outcomes

- For all grades from K–5, students who completed more DreamBox Math lessons per week had higher spring NWEA MAP® math percentile scores. These were statistically significant relationships (*p values* < .05).
- For all grades from K–5, students who spent more time on DreamBox Math lessons per week had higher spring NWEA MAP® math percentile scores. These were statistically significant relationships (*p values* < .05).
- Positive associations between use of DreamBox Math and NWEA MAP® math percentile scores hold across all race- and gender-based subgroups of interest including Black/African American and Hispanic males (*p values* < .05).
- Grades 4 and 5 students who completed 2–5 (moderate use) and more than 5 (high use) weekly lessons had significantly higher state EOG mathematics assessment scores than students who completed fewer than 2 lessons (low use). These were statistically significant relationships (*p values* < .05).
- Grade 4 students who spent more than 60 weekly minutes on DreamBox Math (high use) had significantly higher State EOG scores than students who spent 30–60 (moderate use) and less than 60 (low use) weekly minutes.
- Grades 5 students who spent 30–60 (moderate use) more than 30 weekly minutes on DreamBox Math (high and moderate use) had significantly higher State EOG scores than students who spent less than 30 weekly minutes (low use).

Conclusions

This study provides results to satisfy ESSA evidence requirements for Level III (Promising Evidence) given the study design and positive statistically significant findings.

TABLE OF CONTENTS

Introduction

As a result of disruptions caused to schooling by the COVID-19 pandemic, student performance in math plummeted to alarmingly low levels in the United States and K–12 students are yet to overcome this learning loss (Sparks, 2022). Recent math assessment results have shed light on the severity of learning loss; the 2022 administration of National Assessment of Educational Progress (NAEP) showed largest drops in math performance among 4th and 8th grade students since the program's inception in 1990 (NCES, 2022). In the same year, the U.S. ranked 26th on the Programme for International Student Assessment (PISA), with fewer than 1 in 10 students achieving advanced levels and more than a third failing to meet basic achievement standards (Sparks, 2023). Seventy percent of high school seniors who took the ACT college entrance exam in 2023 scored below the college readiness benchmark (ACT, 2023). By the end of the 2022–23 school year, eighth-grade students were approximately 0.27 standard deviations behind pre-pandemic norms in math, according to the assessment company NWEA (Peters, 2024). Discovery Education recognizes the urgency of accelerating learning in math to mitigate pandemic-related learning loss. The DreamBox K–8 Math learning solution aims to accelerate learning of critical foundational math skills in elementary and middle school, setting students up for success in Algebra I and beyond.

As part of their ongoing efforts to demonstrate the efficacy of DreamBox Math, Discovery Education 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 Math (Shah & Styers, 2022), LearnPlatform designed a study to satisfy ESSA Level III requirements (Promising Evidence) with the following research questions.

Program Implementation Research Questions

1. How many DreamBox Math lessons were completed per week by students during the 2022–23 school year?
2. How much time in DreamBox Math lessons was spent per week by students during the 2022–23 school year?
3. Among DreamBox Math users, what were the usage patterns?

Effectiveness Research Questions

4. After controlling for students' prior math achievement, grade, gender, and race, to what extent was students' use of DreamBox usage associated with NWEA MAP® percentile scores and end-of-grade (EOG) state assessment scores for mathematics?
    a. Do positive findings hold across subgroups of interest (i.e., Black/African American and Hispanic males)?
    b. Do positive findings hold across grade levels and achievement levels (e.g., students in the low proficiency band for their grade)?

Methods

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

Participants

This study occurred during the 2022–23 school year and included 57,146 K–8 students from across 161 schools in one district. Researchers conducted analyses by grade band to allow for better interpretability of findings: Kindergarten (16%); grades 1–2 (33%); grades 3–5 (45%); and grades 6–8 (6%).

This study occurred during the 2022–23 school year and included 57,146 K–8 students from across 161 schools in one district. Researchers conducted analyses by grade band to allow for better interpretability of findings: Kindergarten (16%); grades 1–2 (33%); grades 3–5 (45%); and grades 6–8 (6%). In terms of demographics, the total sample was racially diverse and included Black/African American (34%), Hispanic (31%), White (24%), Asian (7%), and mixed race (4%) students. The district did not provide data on students' socioeconomic status (e.g., low-income indicator, or free and reduced lunch status). Forty-nine percent of students in the sample identified as female.

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 2022–23 student-level usage data 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. According to the DreamBox team, measuring intended usage of the product aligns most closely with the number of lessons students complete. Discovery Education 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 might 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 NWEA MAP® mathematics percentile scores and state end-of-grade (EOG) mathematics assessment scores. NWEA MAP® an adaptive, research-based assessment that reliably measures math knowledge and progress from kindergarten through to Grade 12. Researchers used the NWEA MAP® percentile score as an overall measure of math achievement at two time points: pretest (i.e., winter 2023 for kindergarten sample[1] and fall 2022 for

[1] The district does not administer the fall NWEA MAP® test to kindergarten students. Therefore, to align with the assessment time points, researchers utilized student-level usage data from winter 2023 to spring 2023. This is also another reason for analyzing the kindergarten students separately from the Grades 1–2 sample.

grades 1–8 sample) and posttest (i.e., spring 2023). For the EOG analysis, researchers used spring 2022 scores as the baseline measure.

Data Analysis

Researchers used a variety of quantitative analytic approaches. First, researchers conducted descriptive statistics to examine participant characteristics and support implementation analyses.

Researchers then used two-level multilevel modeling analysis (i.e., students nested in schools) to examine how DreamBox Math use related to student math end-of-year outcomes controlling for their beginning-of-year scores. The analyses included student-level covariates to control for potential selection bias (e.g., grade-level, race, and gender). In addition, researchers calculated standardized effect sizes (Hedges' g) to determine the magnitude of changes in student outcomes. Given that the grades 6–8 sample was about 6% of the total sample and the main implementation of DreamBox was among the elementary grades at this district, the main report contains findings for the grades K–5 sample while the grades 6–8 implementation and outcome findings are outlined in Appendix E.

Program Implementation Findings for K–5 Students

The charts below highlight DreamBox Math use during the 2022–23 school year based on DreamBox's internal usage data (Table 1; details in Appendix B).

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	**9,276**	**3.5**	**30**
Grades 1–2	**18,545**	**3.6**	**34**
Grade 3–5	**26,178**	**2.1**	**29**

Researchers used DreamBox Math's dosage recommendations to group students by similar levels of usage based on the number of average weekly lessons completed and the amount of time (average weekly minutes) spent on DreamBox Math. For average weekly lessons, K–5 students were sorted into three usage categories ranging from low usage (2 lessons or fewer) to moderate usage (between 2 and 5 lessons), and high usage (more than 5 lessons) (Figures 1–3).

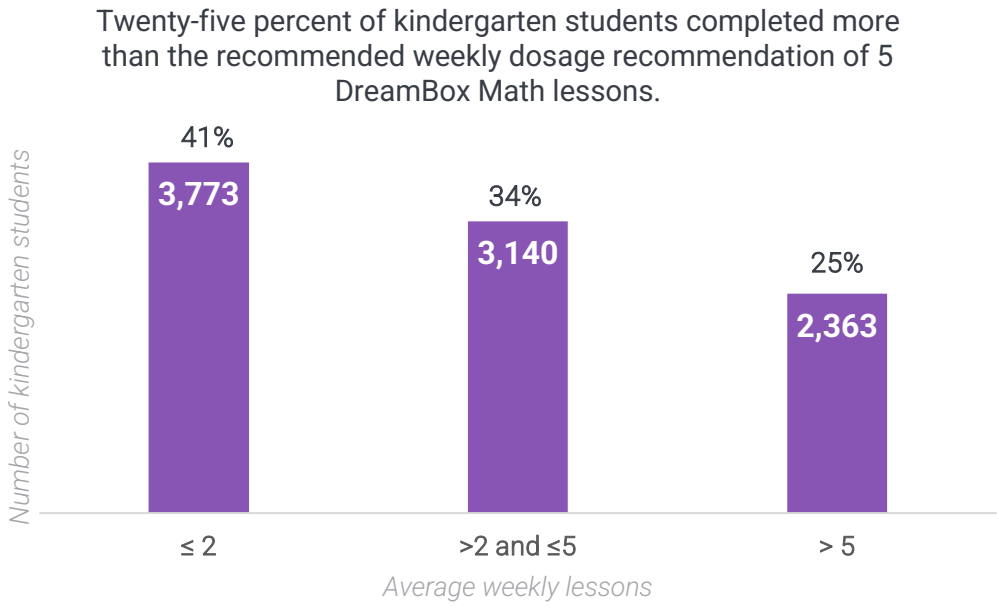


Figure 1. Overall distribution of average weekly lessons completed on DreamBox Math by kindergarten students.

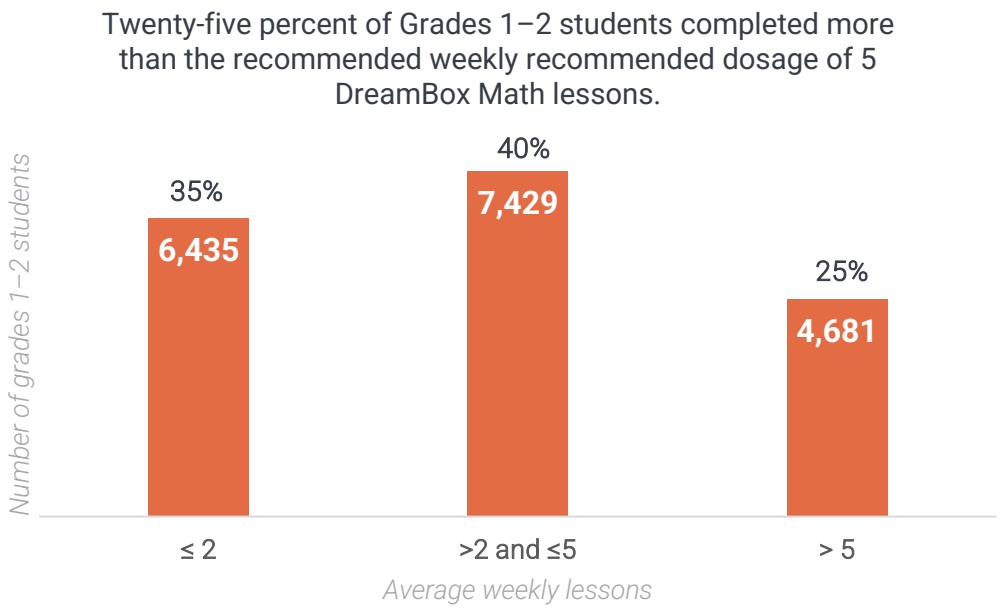


Figure 2. Overall distribution of average weekly lessons completed on DreamBox Math by Grades 1–2 students.

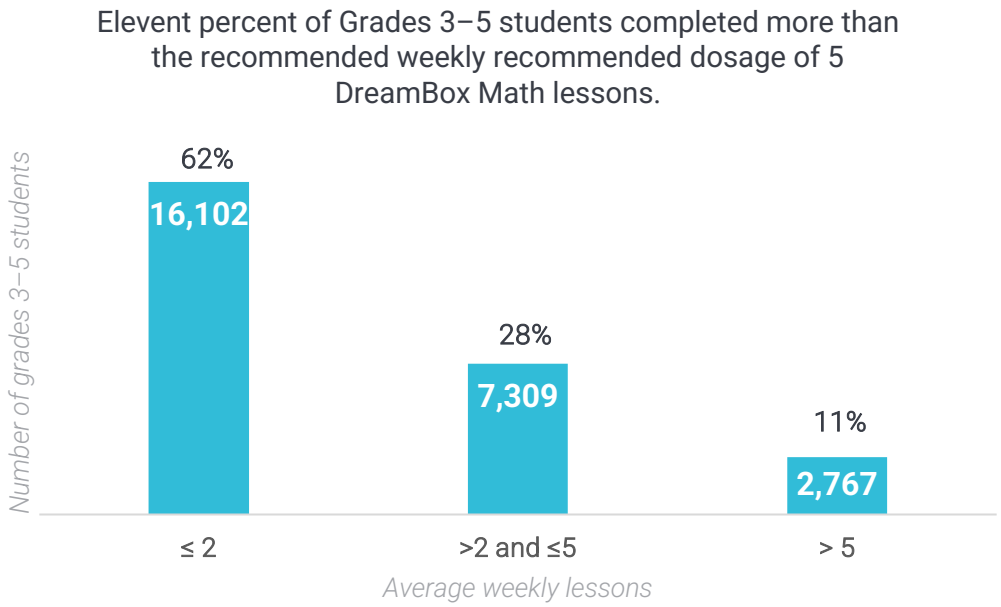


Figure 3. Overall distribution of average weekly lessons completed on DreamBox Math by Grades 3–5 students.

For average weekly minutes, K–5 students were sorted into three usage categories ranging from low usage (30 minutes or fewer) to moderate usage (between 30 and 60 minutes), and high usage (more than 60 minutes) (Figures 4–6).

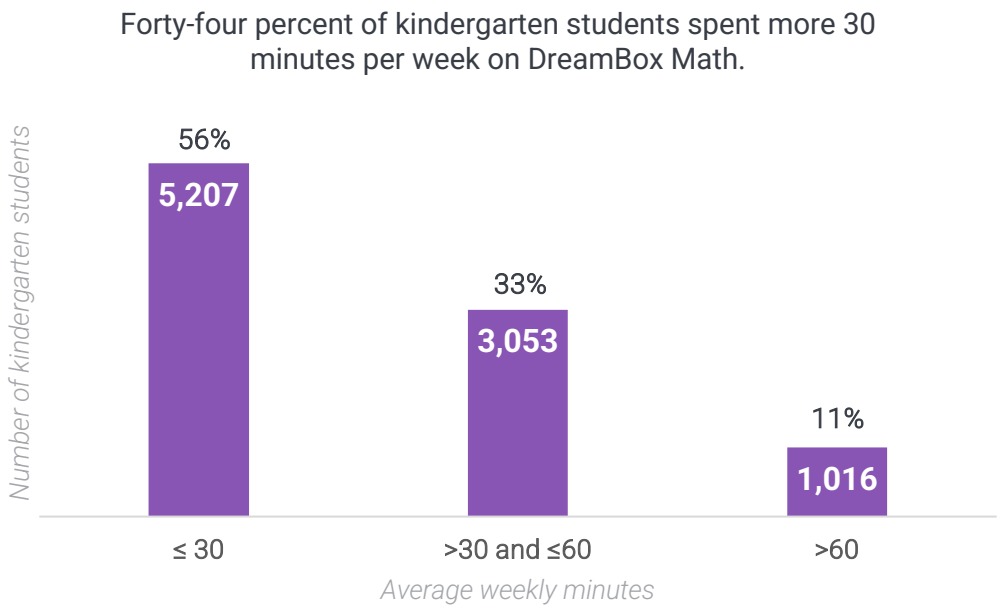


Figure 4. Overall distribution of average weekly minutes spent on DreamBox Math by kindergarten students.

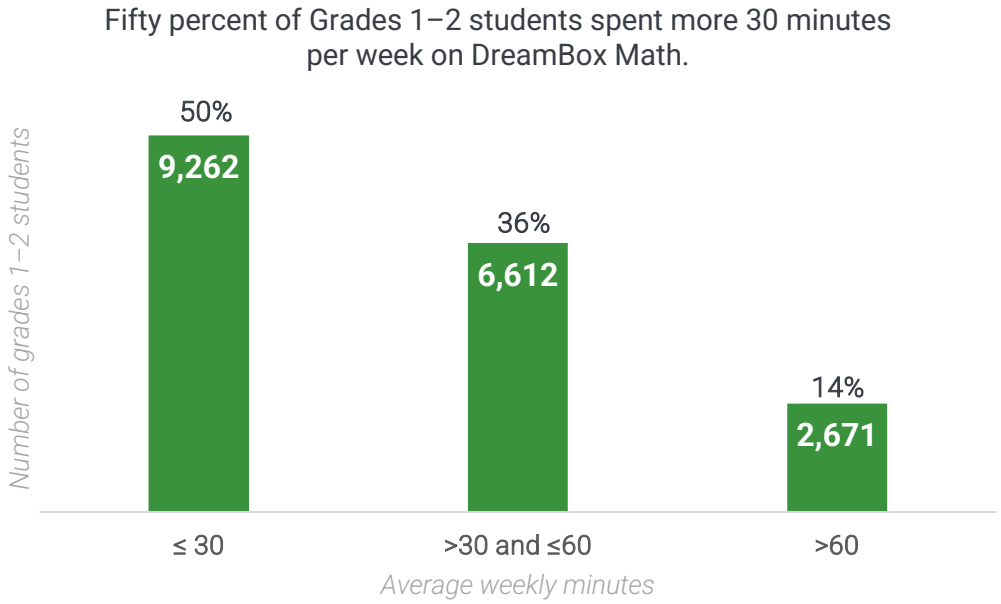


Figure 5. Overall distribution of average weekly minutes spent on DreamBox Math by Grades 1–2 students.

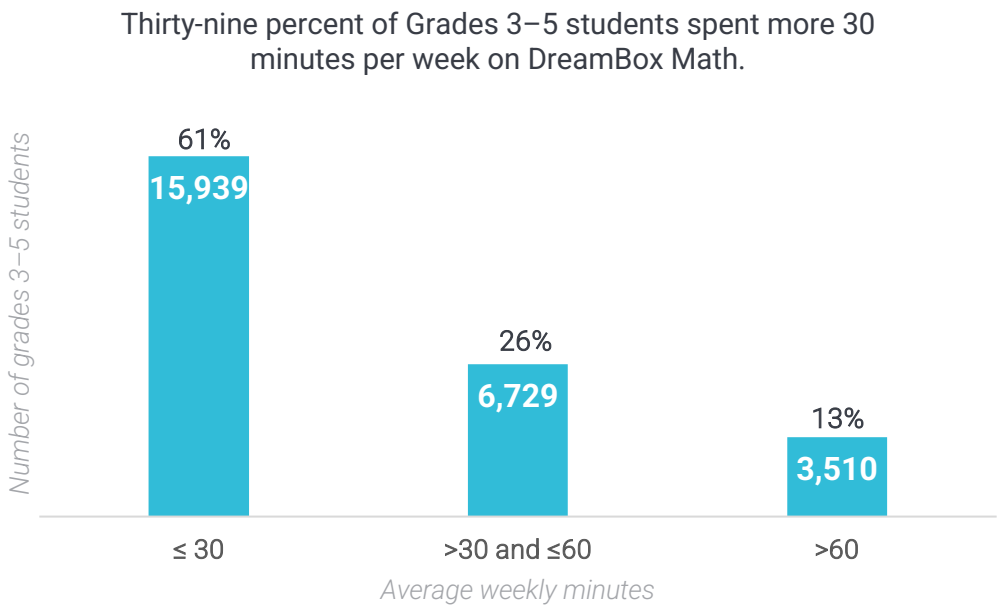


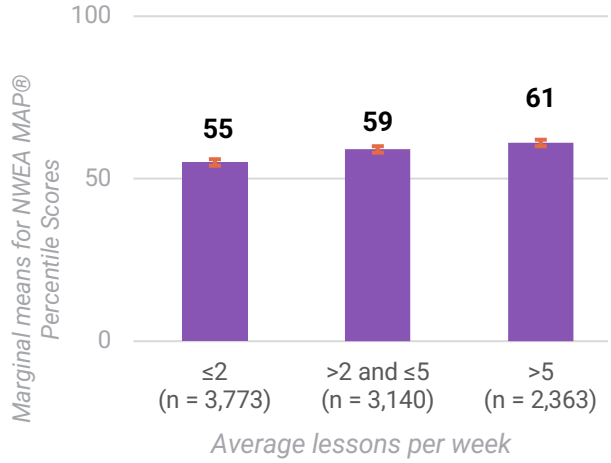
Figure 6. Overall distribution of average weekly minutes spent on DreamBox Math by Grades 3–5 students.

NWEA MAP® Outcome Findings for K–5 Students

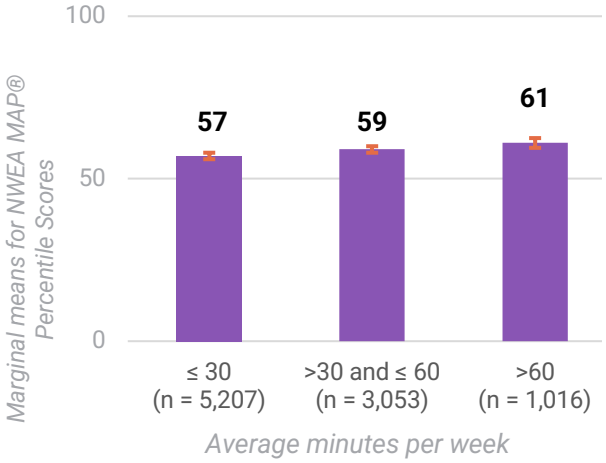
Researchers used two-level multilevel modeling analysis (i.e., students nested in schools) to examine how DreamBox Math use related to student math end-of-year outcomes controlling for their beginning-of-year scores. These analyses also included student-level covariates to control for potential selection bias. 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. **All the usage group findings by grade-band listed in the charts below are statistically significant at the $p < .05$ level** (see Appendix C for more details about the model and the corresponding Hedges' *g* effect sizes).

Association Between Average Weekly Lessons and Minutes and K–5 Students' Outcomes on NWEA MAP® by Usage Groups

Kindergarten students who completed 2–5 (moderate use) and more than 5 (high use) *weekly lessons* had significantly higher NWEA MAP® scores than students who completed fewer than 2 lessons (low use).

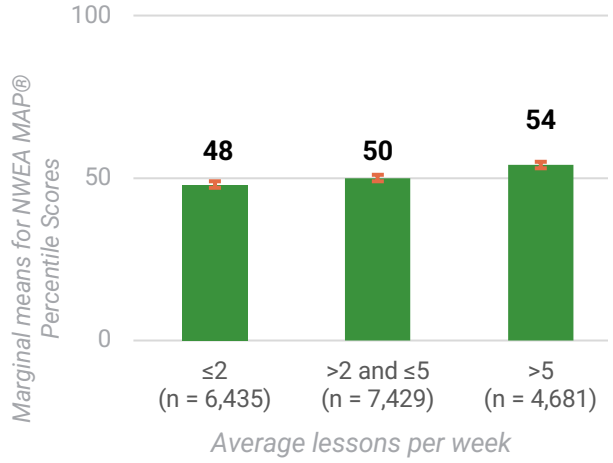


Kindergarten students who spent 30–60 (moderate use) and more than 60 (high use) *weekly minutes* had significantly higher NWEA MAP® scores than students who spent less than 30 weekly minutes (low use).

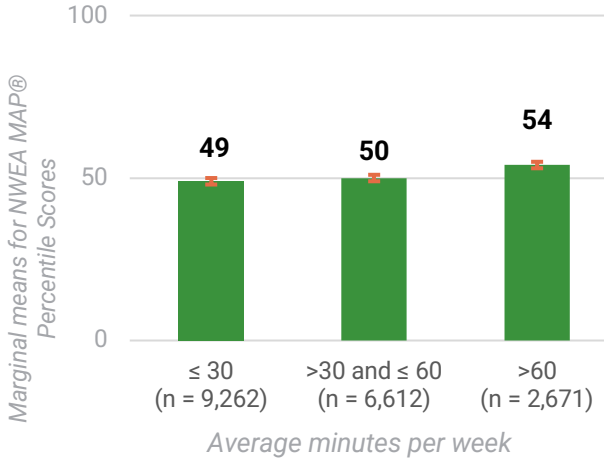


Note: Findings for Kindergarten students spanned a shorter time frame of use from winter 2023 to spring 2023.

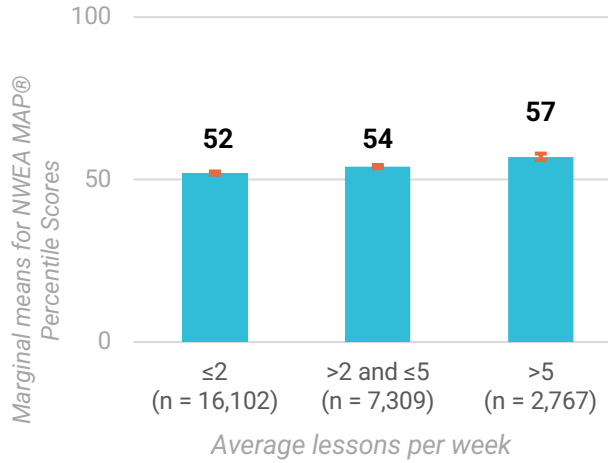
Grade 1–2 students who completed 2–5 (moderate use) and more than 5 (high use) *weekly lessons* had significantly higher NWEA MAP® scores than students who completed fewer than 2 lessons (low use).



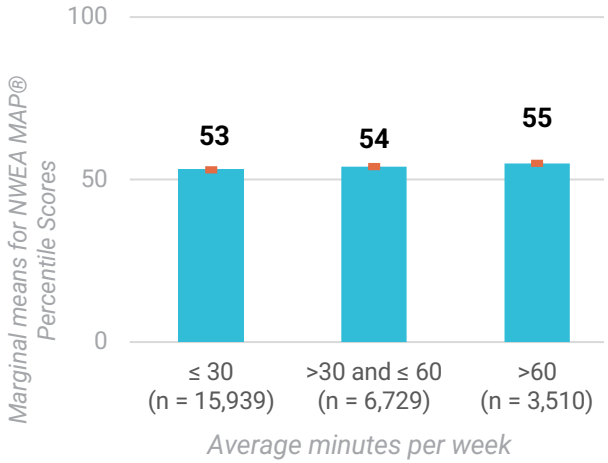
Grade 1–2 students who spent 30–60 (moderate use) and more than 60 (high use) *weekly minutes* had significantly higher NWEA MAP® scores than students who spent less than 30 weekly minutes (low use).



Grade 3–5 students who completed 2–5 (moderate use) and more than 5 (high use) *weekly lessons* had significantly higher NWEA MAP® scores than students who completed fewer than 2 lessons (low use).



Grade 3–5 students who spent 30–60 (moderate use) and more than 60 (high use) *weekly minutes* had significantly higher NWEA MAP® scores than students who spent less than 30 weekly minutes (low use).



Association Between DreamBox Usage and K–5 Students' Outcomes on NWEA MAP® by Race and Gender

To determine whether the positive findings for the full sample at each grade band hold for race- and gender-based subgroups, researchers draw on newer research that helps outline what role participants' social identities play in quantitative analyses (Keane et al., 2023; Suzuki et al., 2021). In doing so, researchers categorized participants into one of eight groups formed by their gender and racial identities directly (e.g., African American/Black males, African American/Black females, Hispanic Males, Hispanic females, etc.). Researchers then conducted linear regressions to examine how use of DreamBox Math related to students' spring 2023 math achievement after controlling for prior math achievement and grade.

Among all eight categories of students, there was a positive, statistically significant association such that students who completed more lessons in DreamBox (lessons weekly) had higher spring NWEA MAP® mathematics percentile scores (Appendix C for details). As illustrated by point estimates in the figures 7 and 8, Hispanic and African American male students at a typical achievement level who complete 2 and 5 lessons per week we see a statistically significant increase in the NWEA MAP® mathematics percentile scores.

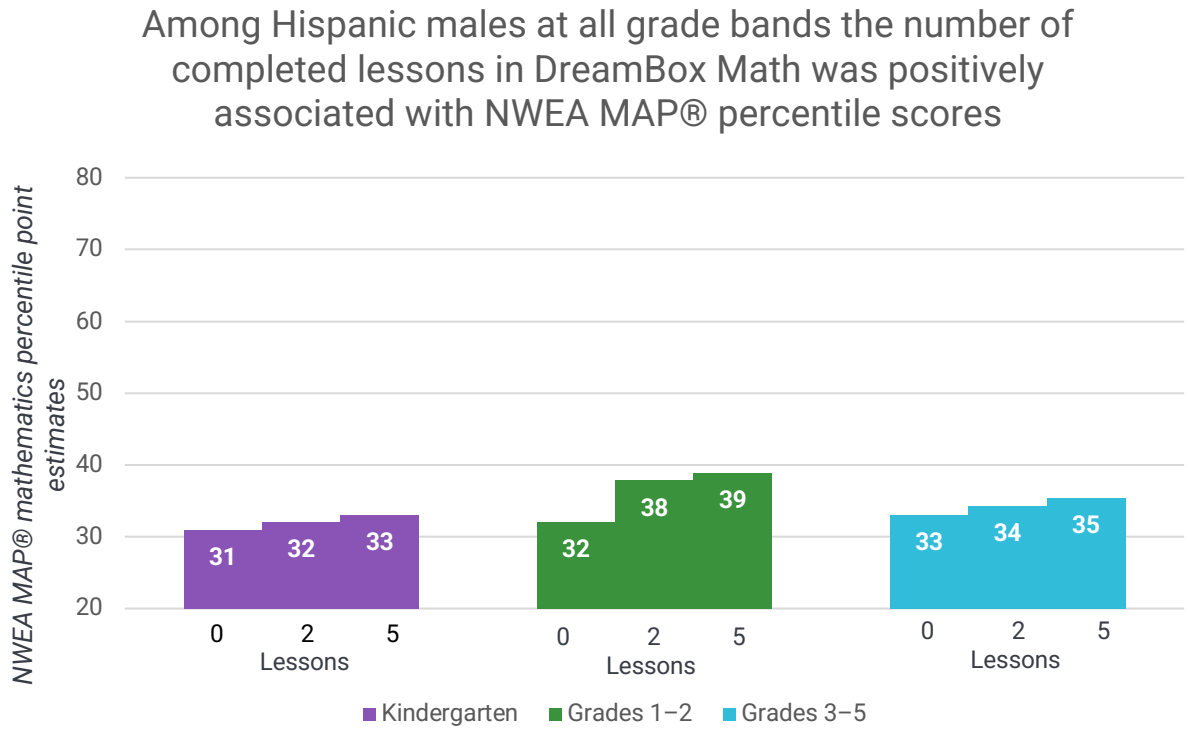


Figure 7. NWEA MAP® mathematics percentile point estimates for Hispanic Males completing 2 and 5 DreamBox Math lessons per week.

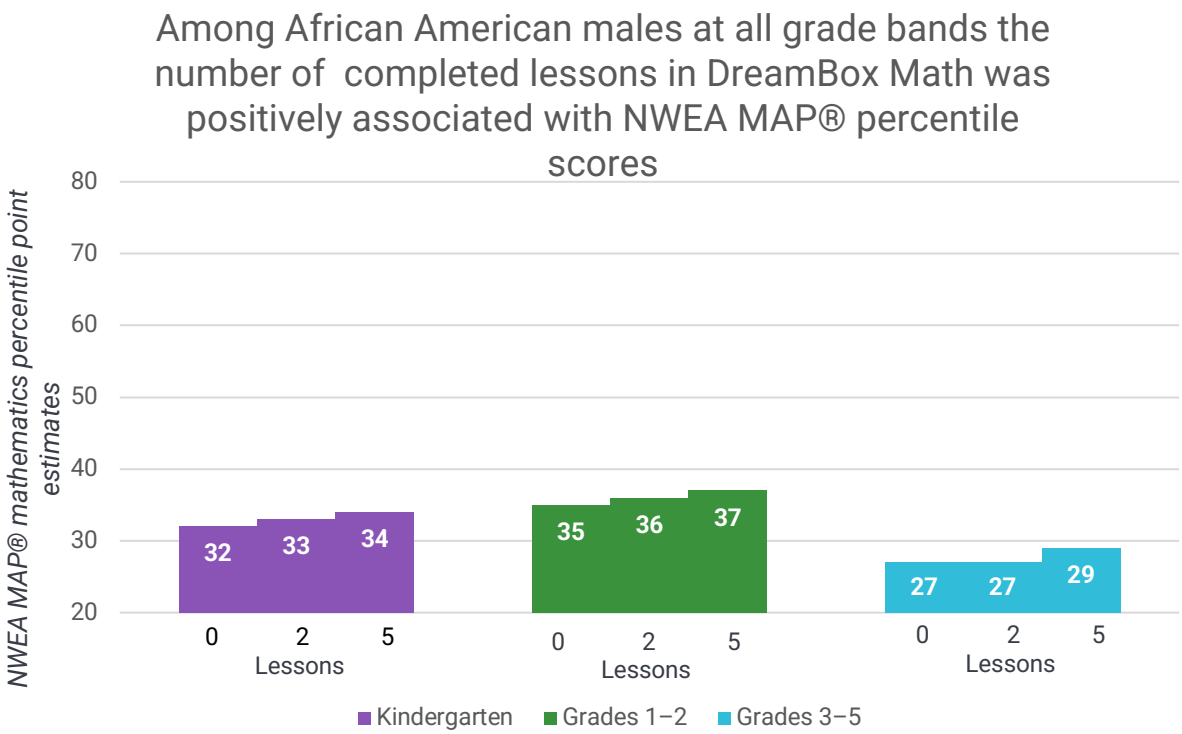


Figure 8. NWEA MAP® mathematics percentile point estimates for African American Males completing 2 and 5 DreamBox Math lessons per week.

State EOG Outcome Findings for Grades 4 and 5 Students

Researchers examined whether greater usage of DreamBox Math related to higher spring 2023 state EOG assessment achievement using spring 2022 (prior year) EOG scores, gender, and/or race, as covariates. Researchers used two-level multilevel modeling analysis (i.e., students nested in schools) to examine how DreamBox Math use related to student math outcomes from beginning- to end-of-year. 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. **All the usage group findings listed in the charts below are statistically significant at the $p < .05$ level** (see Appendices D and E for more details about the model and the corresponding Hedge's *g* effect sizes).

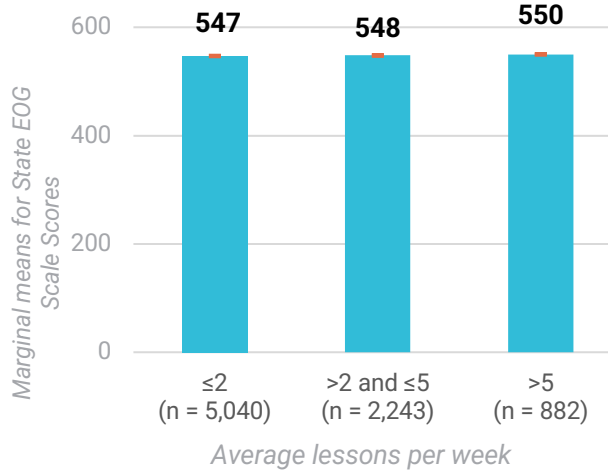
The state EOG assessments start at grade 3 and these state EOG scores are not vertically scaled. Therefore, researchers did not analyze EOG data for grade 3 students in the sample because they did not have a baseline measure of achievement. Further, researchers conducted the analysis separately by grade-level. As such, EOG findings are only listed for grades 4 and 5 students in the sample.

For supporting readers' interpretations of the EOG outcome findings, researchers reference here North Carolina's EOG tests of mathematics grades 3– 5 achievement level ranges (cut scores; NCDPI, 2019):

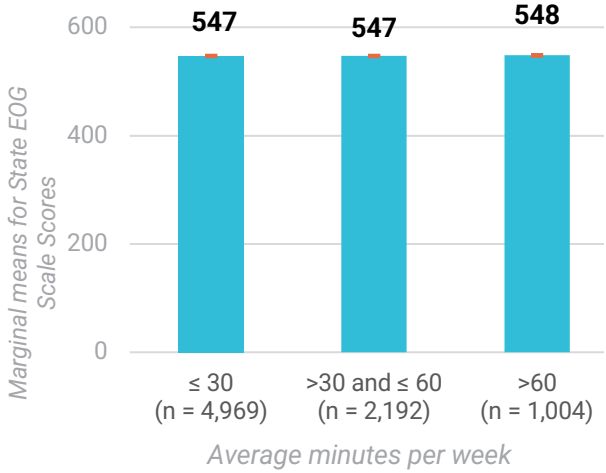
Grade	Not Proficient	Level 3	Level 4	Level 5
3	≤ 544	545–550	551–559	≥ 560
4	≤ 546	547–551	552–559	≥ 560
5	≤ 545	546–550	551–560	≥ 561

Level 4 and 5 meets on-grade-level proficiency and career-and-college readiness standards while Level 3 only meets on-grade-level proficiency standard.

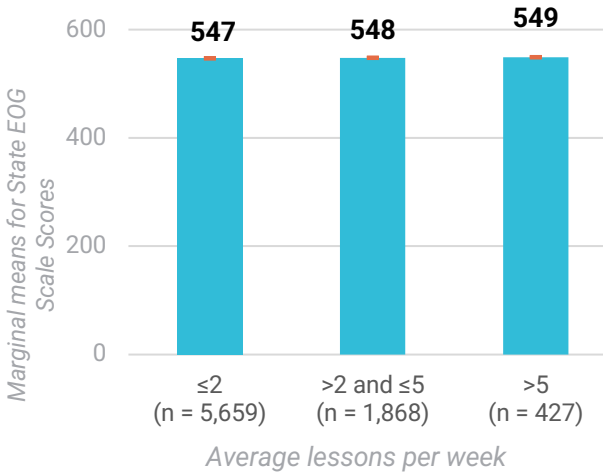
Grades 4 students who completed 2–5 (moderate use) and more than 5 (high use) weekly lessons had significantly higher state EOG mathematics assessment scores than students who completed fewer than 2 lessons (low use).



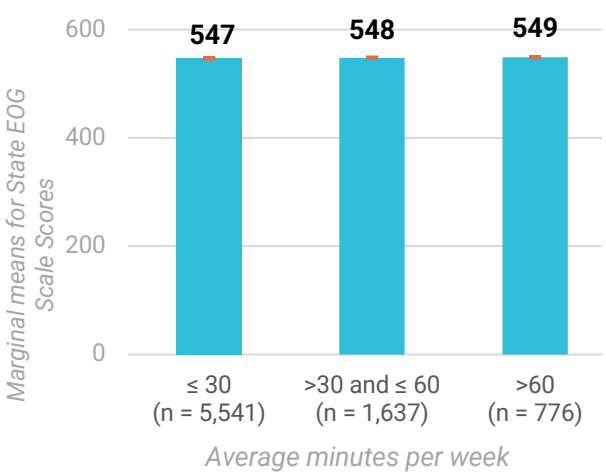
Grades 4 students who spent more than 60 *weekly minutes* on DreamBox Math (high use) had significantly higher State EOG scores than students who spent less than 60 weekly minutes (low and moderate use).



Grades 5 students who completed 2–5 (moderate use) and more than 5 (high use) weekly lessons had significantly higher state EOG mathematics assessment scores than students who completed fewer than 2 lessons (low use).



Grades 5 students who spent more than 30 *weekly minutes* on DreamBox Math (high and moderate use) had significantly higher State EOG scores than students who spent less than 30 weekly minutes (low use).



Summary of Effect Sizes for NWEA MAP® Findings

Researchers calculated Hedges' *g* effect sizes (Hedges, 1981) so that results could be compared between the different assessments and grade bands. It is recommended that these effect sizes be interpreted small if <0.05, medium if between 0.05 and 0.19, and large if >0.20 (Kraft, 2020). Furthermore, it is also important to consider scalability, cost, context, and prior research when determining the practical effectiveness of any intervention beyond and alongside the calculated effect size. Using these recommended thresholds, there was a large effect of using DreamBox for > 5 compared to ≤ 2 weekly lessons for Kindergarten and Grades 1–2 students. There was a medium effect for Grades 3—5 students. Overall, the effects were larger for weekly lessons completed compared to weekly minutes.

Grade Band	Moderate vs. Low Use 2–5 vs. ≤ 2 weekly lessons	High vs. Low Use > 5 vs. ≤ 2 weekly lessons	High vs. Moderate Use > 5 vs. ≤ 2–5 weekly lessons
Kindergarten	0.10	0.20	0.10
Grades 1–2	0.08	0.20	0.13
Grades 3–5	0.07	0.14	0.07

Grade Band	Moderate vs. Low Use 30–60 vs. ≤ 30 weekly minutes	High vs. Low Use > 60 vs. ≤ 30 weekly minutes	High vs. Moderate Use > 60 vs. ≤ 30–60 weekly minutes
Kindergarten	0.08	0.15	0.07
Grades 1–2	0.04	0.14	0.10
Grades 3–5	0.03	0.07	0.04

Summary of Effect Sizes for EOG State Math Assessment Findings

The findings for end-of-grade state math assessment scores were larger than those for NWEA MAP math scores for Grades 4 and 5 students. Using the recommended criteria for interpreting effect sizes (i.e., <0.05 is small, 0.05 - 0.19 is medium, >0.20 is large; Kraft, 2020), there was a large effect of using DreamBox for > 5 compared to ≤ 2 weekly lessons for Grades 4–5 students. Furthermore, weekly lessons was a stronger predictor achievement compared to weekly minutes, which was aligned with the NWEA MAP findings.

Grade Band	Moderate vs. Low Use 2–5 vs. ≤ 2 weekly lessons	High vs. Low Use > 5 vs. ≤ 2 weekly lessons	High vs. Moderate Use > 5 vs. ≤ 2–5 weekly lessons
Grades 4	0.11	0.26	0.15
Grades 5	0.14	0.23	0.09

Grade Band	Moderate vs. Low Use 30–60 vs. ≤ 30 weekly minutes	High vs. Low Use > 60 vs. ≤ 30 weekly minutes	High vs. Moderate Use > 60 vs. ≤ 30–60 weekly minutes
Grades 4	0.02	0.09	0.07
Grades 5	0.08	0.13	0.05

Conclusions and Recommendations

The findings support an association between DreamBox Math usage and improved math skills for K–5 students. Further, the positive outcome findings hold for race- and gender-based subgroups at the district. This study provides results to satisfy ESSA evidence requirements for Level III (Promising Evidence). Specifically, this study met the following criteria:

- ✔ Correlative design
- ✔ Proper design and implementation
- ✔ Statistical controls through covariates
- ✔ At least one statistically significant, positive finding

Researchers recommend the following next steps:

- Discovery Education 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

ACT (2023, October 10) *Fewer high school seniors ready for college as ACT scores continue to decline* URL

Hedges, L. (1981). Distribution theory for Glass's estimator of effect size and related estimators. *Journal of Educational Statistics, 6*, 107–128.

Keane, J., Lee, A. A., Lee, G. A., Perez, T., & Linnenbrink-Garcia, L. (April, 2023). Different approaches to modeling intersectionality: A QuantCrit exploration of stereotype threat patterns. Symposium paper submitted for T. Perez & A. Lira (Chairs), Intersectional and critical approaches to examining the experiences and success of minoritized undergraduates in STEM. American Educational Research Association. Chicago, IL.

Kraft, M. A. (2020). Interpreting effect sizes of education interventions. *Educational Researcher, 49*, 241–253.

NC DPI. (2019). North Carolina End-of-Grade Tests of Mathematics Grades 3-8 https://www.dpi.nc.gov/documents/accountability/testing/eog/eog-mathematics-achievement-level-ranges-and-descriptors/open

National Center for Educational Statistics. (2022). The Nation's Report Card: Results from the 2022 Mathematics and Reading Assessments.

Peters, S. (2024, May 16). *When are students "ready" for algebra?* NWEA URL

Sparks, S. (2022, October 24). *Two decades of progress, nearly gone: national math, reading scores hit historic lows.* EdWeek. URL.

Sparks, Sarah (2023, December 5). *U.S. teenagers decline in global test of math, but hold steady in reading, science.* EdWeek. URL

Suzuki, S., Morris, S. L., & Johnson, S. K. (2021). Using QuantCrit to advance an anti-racist developmental science: Applications to mixture modeling. *Journal of Adolescent Research, 36*(5), 535-560.

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

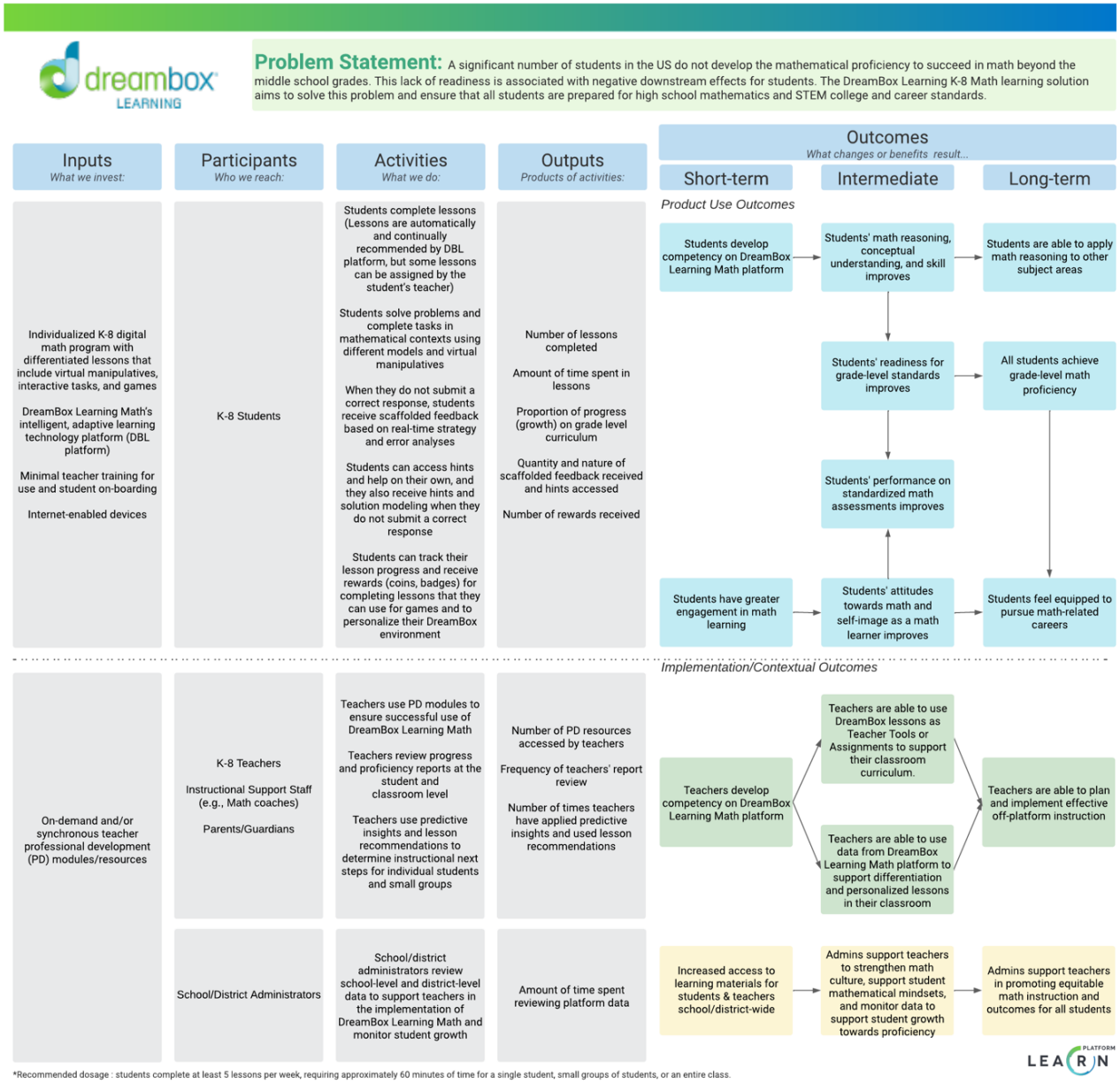


Figure 1. DreamBox Learning Math logic model

Appendix B. Additional Information on K–5 Program Implementation

Table B1. Descriptive statistics for the weekly lessons' usage categories for kindergarten sample

	Usage categories: weekly lessons	*n*	Mean	SD
Low	≤ 2 weekly lessons	3,773	0.9	0.6
Moderate	> 2 and ≤ 5 weekly lessons	3,140	3.4	0.8
High	> 5 weekly lessons	2,363	7.8	2.3

Table B2. Descriptive statistics for the weekly minutes' usage categories for kindergarten sample

	Usage categories: weekly minutes	*n*	Mean	SD
Low	≤ 30 weekly minutes	5,207	14	9
Moderate	> 30 and ≤ 60 weekly minutes	3,053	42	8
High	> 60 weekly minutes	1,016	77	14

Table B3. Descriptive statistics for the weekly lessons' usage categories for grades 1–2 sample

	Usage categories: weekly lessons	*n*	Mean	SD
Low	≤ 2 weekly lessons	6,435	1.1	0.6
Moderate	> 2 and ≤ 5 weekly lessons	7,429	3.4	0.8
High	> 5 weekly lessons	4,681	7.4	2.0

Table B4. Descriptive statistics for the weekly minutes' usage categories for grades 1–2 sample

	Usage categories: weekly minutes	*n*	Mean	SD
Low	≤ 30 weekly minutes	9,262	16	8
Moderate	> 30 and ≤ 60 weekly minutes	6,612	43	8
High	> 60 weekly minutes	2,671	77	14

Table B5. Descriptive statistics for the weekly lessons' usage categories for grades 3–5 sample

	Usage categories: weekly lessons	*n*	Mean	SD
Low	≤ 2 weekly lessons	16,102	0.7	0.6
Moderate	> 2 and ≤ 5 weekly lessons	7,309	3.3	0.9
High	> 5 weekly lessons	2,767	7.0	1.7

Table B6. Descriptive statistics for the weekly minutes' usage categories for grades 3–5 sample

	Usage categories: weekly minutes	*n*	Mean	SD
Low	≤ 30 weekly minutes	15,939	12	8.8
Moderate	> 30 and ≤ 60 weekly minutes	6,729	43	8.6
High	> 60 weekly minutes	3,510	80	15.3

Appendix C. Additional Information on K–5 Outcome Findings

Researchers examined whether greater usage of DreamBox Math related to higher spring NWEA MAP® achievement using partial correlation models that included prior NWEA MAP® percentile scores, grade, gender, and/or race, as covariates. Researchers report statistically significant findings at the $p < .05$ level. Statistically significant findings are marked green (positive effect size).

Table C1. Overall association between average weekly lessons and K–5 Students' NWEA MAP® mathematics percentile scores

Grade Band	Covariates	*n*	Partial correlation coefficient	*p*-value
Kindergarten	Winter 2023 percentile scores and race	9,276	.14	<.001
Grades 1–2	Winter 2023 percentile scores, gender, and race	18,545	.14	<.001
Grades 3–5	Winter 2023 percentile scores, grade, gender, and race	26,178	.09	<.001

Table C2. Overall association between average weekly minutes and K–5 Students' NWEA MAP® mathematics percentile scores

Grade Band	Covariates	*n*	Partial correlation coefficient	*p*-value
Kindergarten	Winter 2023 percentile scores, race	9,276	.11	<.001
Grades 1–2	Winter 2023 percentile scores, gender, and race	18,545	.05	<.001
Grades 3–5	Winter 2023 percentile scores, grade, gender, and race	26,178	.03	<.001

Association Between Average Weekly Lessons and K–5 Students' Outcomes on NWEA MAP® by Usage Groups

Table C3. Association between kindergarten DreamBox Math usage groups and spring 2023 NWEA MAP® mathematics percentile scores

Predictor	Unstd. Beta Coefficient	Standard Error	Test statistic[2]	*p*-value
Moderate Use vs. Low Use (Hedges' *g* = 0.11)	3.37	0.42	8.03	<.001
High Use vs. Low Use (Hedges' *g* = 0.21)	6.13	0.49	12.57	<.001
High Use vs. Moderate Use (Hedges' *g* = 0.10)	2.76	0.47	5.89	<.001
Winter 2023 NWEA MAP® percentile scores	0.77	0.01	120.94	<.001
Race	0.62	0.09	6.37	<.001
School-level random effects	12.29	2.19	213.64	<.001

Table C4. Association between grades 1–2 DreamBox Math usage groups and spring 2023 NWEA MAP® mathematics percentile scores

Predictor	Unstd. Beta Coefficient	Standard Error	Test statistic	*p*-value
Moderate Use vs. Low Use (Hedges' *g* = 0.07)	2.28	0.28	8.28	<.001
High Use vs. Low Use (Hedges' *g* = 0.21)	6.49	0.33	19.74	<.001
High Use vs. Moderate Use (Hedges' *g* = 0.14)	4.21	0.30	14.24	<.001
Fall 2022 NWEA MAP® percentile scores	0.78	0.00	191.34	<.001
Gender	2.37	0.23	10.51	<.001
Race	0.38	0.07	5.81	<.001

[2] Test statistics are a *z*-score for the fixed effects (i.e., usage group, prior performance on NWEA MAP®, and student demographic characteristics) and a chi-square for the random effects (i.e., school-level).

Predictor	Unstd. Beta Coefficient	Standard Error	Test statistic	*p*-value
School-level random effects	19.39	2.84	1021.19	<.001

Table C5. Association between grades 3–5 DreamBox Math usage groups and spring 2023 NWEA MAP® mathematics percentile scores

Predictor	Unstd. Beta Coefficient	Standard Error	Test statistic	*p*-value
Moderate Use vs. Low Use (Hedges' *g* = 0.08)	2.45	0.21	11.93	<.001
High Use vs. Low Use (Hedges' *g* = 0.17)	5.12	0.31	16.56	<.001
High Use vs. Moderate Use (Hedges' *g* = 0.09)	2.67	0.31	8.64	<.001
Fall 2022 NWEA MAP® percentile scores	0.86	0.00	279.46	<.001
Grade	-1.19	0.10	-11.35	<.001
Gender	0.83	0.17	4.94	<.001
School-level random effects	8.93	1.33	862.25	<.001

Association Between Average Weekly Minutes and K–5 Students' Outcomes on NWEA MAP® by Usage Groups

Table C6. Association between kindergarten DreamBox Math usage groups and spring 2023 NWEA MAP® mathematics percentile scores

Predictor	Unstd. Beta Coefficient	Standard Error	Test statistic[3]	*p*-value
Moderate Use vs. Low Use (Hedges' *g* = 0.08)	2.50	0.41	6.05	<.001
High Use vs. Low Use (Hedges' *g* = 0.16)	4.92	0.63	7.76	<.001
High Use vs. Moderate Use (Hedges' *g* = 0.08)	2.42	0.63	3.83	<.001
Winter 2023 NWEA MAP® percentile scores	0.79	0.00	124.18	<.001
Race	0.63	0.10	6.41	<.001
School-level random effects	12.73	2.26	217.47	<.001

Table C7. Association between grades 1–2 DreamBox Math usage groups and spring 2023 NWEA MAP® mathematics percentile scores

Predictor	Unstd. Beta Coefficient	Standard Error	Test statistic	*p*-value
Moderate Use vs. Low Use (Hedges' *g* = 0.04)	1.27	0.27	4.7	<.001
High Use vs. Low Use (Hedges' *g* = 0.15)	4.75	0.38	12.43	<.001
High Use vs. Moderate Use (Hedges' *g* = 0.11)	3.48	0.37	9.49	<.001
Fall 2022 NWEA MAP® percentile scores	0.79	0.00	194.51	<.001
Gender	2.65	0.23	11.75	<.001
Race	0.38	0.07	5.73	<.001

[3] Test statistics are a *z*-score for the fixed effects (i.e., usage group, prior performance on NWEA MAP®, and student demographic characteristics) and a chi-square for the random effects (i.e., school-level).

Predictor	Unstd. Beta Coefficient	Standard Error	Test statistic	*p*-value
School-level random effects	19.50	2.86	982.39	<.001

Table C8. Association between grades 3–5 DreamBox Math usage groups and spring 2023 NWEA MAP® mathematics percentile scores

Predictor	Unstd. Beta Coefficient	Standard Error	Test statistic	*p*-value
Moderate Use vs. Low Use (Hedges' *g* = 0.04)	1.17	0.21	5.5	<.001
High Use vs. Low Use (Hedges' *g* = 0.08)	2.44	0.29	8.39	<.001
High Use vs. Moderate Use (Hedges' *g* = 0.04)	1.27	0.30	4.29	<.001
Fall 2022 NWEA MAP® percentile scores	0.86	0.00	284.38	<.001
Grade	-1.39	0.10	-13.37	<.001
Gender	0.98	0.17	5.84	<.001
School-level random effects	8.75	1.31	826.04	<.001

Association Between Average Weekly Lessons and K–5 Students' Outcomes on NWEA MAP® by Race and Gender Subgroups

Table C9. Association between Kindergarten DreamBox Math weekly lessons and spring 2023 NWEA MAP® mathematics percentile scores by race and gender subgroups

Subgroup	*N*	Unstd. Beta Coeff.	SE	*t*-value	*p*-value	Impact of 5 lessons (percentile point change)
Hispanic Females	1,423	0.36	0.08	4.71	<.001	+2
Asian Females	287	0.34	0.15	2.29	.023	+2
African American Females	1,377	0.33	0.07	4.47	<.001	+2
White Females	1,149	0.33	0.06	5.44	<.001	+2

Subgroup	*N*	Unstd. Beta Coeff.	SE	*t*-value	*p*-value	Impact of 5 lessons (percentile point change)
Hispanic Males	1,498	0.38	0.07	5.45	<.001	+2
Asian Males	339	0.44	0.14	3.21	.001	+2
African American Males	1,543	0.48	0.08	6.27	<.001	+2
White Males	1,196	0.35	0.07	5.09	<.001	+2

Table C10. Association between grades 1–2 DreamBox Math weekly lessons and spring 2023 NWEA MAP® mathematics percentile scores by race and gender subgroups

Subgroup	*N*	Unstd. Beta Coeff.	SE	*t*-value	*p*-value	Impact of 5 lessons (percentile point change)
Hispanic Females	2,847	0.46	0.06	7.74	<.001	+2
Asian Females	629	0.24	0.09	2.62	.009	+1
African American Females	3,025	0.41	0.06	6.65	<.001	+2
White Females	2,308	0.51	0.06	8.98	<.001	+3
Hispanic Males	2,818	0.35	0.06	5.88	<.001	+2
Asian Males	674	0.34	0.10	3.52	<.001	+2
African American Males	3,091	0.48	0.06	7.95	<.001	+2
White Males	2,430	0.38	0.05	7.43	<.001	+2

Table C11. Association between grades 3–5 DreamBox Math weekly lessons and spring 2023 NWEA MAP® mathematics percentile scores by race and gender subgroups

Subgroup	*N*	Unstd. Beta Coeff.	SE	*t*-value	*p*-value	Impact of 5 lessons (percentile point change)
Hispanic Females	3,883	0.30	0.07	4.61	<.001	+2
Asian Females	974	0.24	0.10	2.47	.014	+1
African American Females	4,484	0.37	0.06	5.96	<.001	+2

Subgroup	*N*	Unstd. Beta Coeff.	SE	*t*-value	*p*-value	**Impact of 5 lessons** (percentile point change)
White Females	3,056	0.24	0.05	4.47	<.001	+2
Hispanic Males	4,187	0.33	0.06	5.92	<.001	+2
Asian Males	943	0.19	0.09	2.06	.039	+2
African American Males	4,506	0.40	0.06	6.59	<.001	+2
White Males	3,193	0.15	0.05	2.93	.003	+2

Appendix D. Grades 4 and 5 State End-of-Grade Assessment Outcome Findings

Table D1. Descriptive statistics for the weekly lessons' usage categories for grade 4 sample

Usage categories: weekly lessons	*n*	Mean	SD
≤ 2 weekly lessons	5,040	0.8	0.6
> 2 and ≤ 5 weekly lessons	2,243	3.2	0.8
> 5 weekly lessons	882	6.9	1.6

Table D2. Descriptive statistics for the weekly minutes' usage categories for grade 4 sample

Usage categories: weekly minutes	*n*	Mean	SD
≤ 30 weekly minutes	4,969	12	9
> 30 and ≤ 60 weekly minutes	2,192	43	8
> 60 weekly minutes	1,004	78	14

Table D3. Descriptive statistics for the weekly lessons' usage categories for grade 5 sample

Usage categories: weekly lessons	*n*	Mean	SD
≤ 2 weekly lessons	5,659	0.7	0.6
> 2 and ≤ 5 weekly lessons	1,868	3.2	0.8
> 5 weekly lessons	427	6.4	1.4

Table D4. Descriptive statistics for the weekly minutes' usage categories for grade 5 sample

Usage categories: weekly minutes	*n*	Mean	SD
≤ 30 weekly minutes	5,541	12	8
> 30 and ≤ 60 weekly minutes	1,637	42	8
> 60 weekly minutes	776	79	14

Association Between Average Weekly Lessons and Grades 4 & 5 Students' Outcomes on State EOG Assessment by Usage Groups

Table D5. Association between grade 4 DreamBox Math usage groups and spring 2023 state EOG mathematics assessment scaled scores

Predictor	Unstd. Beta Coefficient	Standard Error	Test statistic	*p*-value
Moderate Use vs. Low Use (Hedges' *g* = 0.11)	1.19	0.16	7.63	<.001
High Use vs. Low Use (Hedges' *g* = 0.26)	2.67	0.23	11.46	<.001
High Use vs. Moderate Use (Hedges' *g* = 0.15)	1.48	0.22	6.68	<.001
Spring 2022 (prior year) EOG scores	0.80	0.01	123.08	<.001
Gender	0.69	0.12	5.88	<.001
Race	0.09	0.03	2.60	.009
School-level random effects	3.42	0.52	660.20	<.001

Table D6. Association between grade 5 DreamBox Math usage groups and spring 2023 state EOG mathematics assessment scaled scores

Predictor	Unstd. Beta Coefficient	Standard Error	Test statistic	*p*-value
Moderate Use vs. Low Use (Hedges' *g* = 0.14)	1.53	0.16	9.34	<.001
High Use vs. Low Use (Hedges' *g* = 0.23)	2.52	0.31	8.18	<.001
High Use vs. Moderate Use (Hedges' *g* = 0.09)	0.99	0.30	3.30	.001
Spring 2022 (prior year) EOG scores	0.87	0.01	134.13	<.001
Gender	-0.47	0.12	-3.98	<.001
School-level random effects	4.31	0.65	825.42	<.001

Association Between Average Weekly Minutes and Grades 4 & 5 Students' Outcomes on State EOG Assessment by Usage Groups

Table D7. Association between grade 4 DreamBox Math usage groups and spring 2023 state EOG mathematics assessment scaled scores

Predictor	Unstd. Beta Coefficient	Standard Error	Test statistic	*p*-value
Moderate Use vs. Low Use (Hedges' *g* = 0.02)	0.18	0.16	1.12	0.262
High Use vs. Low Use (Hedges' *g* = 0.09)	0.93	0.23	4.00	<.001
High Use vs. Moderate Use (Hedges' *g* = 0.07)	0.75	0.22	3.39	0.001
Spring 2022 (prior year) EOG scores	0.83	0.01	128.84	<.001
Gender	0.79	0.12	6.66	<.001
Race	0.08	0.03	2.36	.018
School-level random effects	3.43	0.52	644.83	<.001

Table D8. Association between grade 5 DreamBox Math usage groups and spring 2023 state EOG mathematics assessment scaled scores

Predictor	Unstd. Beta Coefficient	Standard Error	Test statistic	*p*-value
Moderate Use vs. Low Use (Hedges' *g* = 0.08)	0.91	0.17	5.25	<.001
High Use vs. Low Use (Hedges' *g* = 0.13)	1.43	0.27	5.37	<.001
High Use vs. Moderate Use (Hedges' *g* = 0.15)	0.52	0.26	1.99	0.047
Spring 2022 (prior year) EOG scores	0.88	0.01	136.39	<.001
Gender	-0.43	0.12	-3.60	<.001
School-level random effects	4.25	0.65	805.33	<.001

Appendix E. Grades 4 and 5 Outcome Findings for Students Classified as Not Proficient on Spring 2022 (Prior Year) EOG Math Assessment

Researchers also examined whether greater usage of DreamBox Math related to higher spring 2023 state EOG assessment achievement among grade 4 and 5 students classified as not proficient on the spring 2022 (prior year) EOG math assessment. They controlled for spring 2022 (prior year) EOG scores, gender, and/or race. Researchers used two-level multilevel modeling analysis (i.e., students nested in schools) to examine how DreamBox Math use related to student math outcomes. To allow for better interpretability of results, marginal means charts are presented below. The black vertical lines at the top of each bar represent a 95% confidence interval.

Table E1. Descriptive statistics for the weekly lessons' usage categories for grade 4 sample that was classified as not proficient on the spring 2022 EOG math assessment

Usage categories: weekly lessons	*n*	Mean	SD
≤ 2 weekly lessons	2,481	0.8	0.6
> 2 and ≤ 5 weekly lessons	1,419	3.3	0.8
> 5 weekly lessons	686	6.9	1.6

Table E2. Descriptive statistics for the weekly lessons' usage categories for grade 5 sample that was classified as not proficient on the spring 2022 EOG math assessment

Usage categories: weekly lessons	*n*	Mean	SD
≤ 2 weekly lessons	2,494	0.7	0.6
> 2 and ≤ 5 weekly lessons	1,061	3.2	0.8
> 5 weekly lessons	267	6.5	1.5

Association Between Average Weekly Lessons and Grades 4 & 5 Students' Outcomes on State EOG Assessment by Usage Groups

Table E3. Association between grade 4 DreamBox Math usage groups and spring 2023 state EOG mathematics assessment scaled scores

Predictor	Unstd. Beta Coefficient	Standard Error	Test statistic	*p*-value
Moderate Use vs. Low Use (Hedges' *g* = 0.14)	1.05	0.19	5.42	<.001
High Use vs. Low Use (Hedges' *g* = 0.39)	2.82	0.27	10.48	<.001
High Use vs. Moderate Use (Hedges' *g* = 0.24)	1.77	0.25	7.12	<.001
Spring 2022 EOG scores	0.75	0.01	59.35	<.001
Gender	0.80	0.15	5.37	<.001
Race	0.11	0.04	2.68	.007
School-level random effects	4.50	0.74	432.45	<.001

Among Grades 4 students classified as not proficient at the end of the prior year, those who completed 2–5 (moderate use) and more than 5 (high use) weekly lessons had significantly higher state EOG mathematics assessment scores than students who completed fewer than 2 lessons (low use).

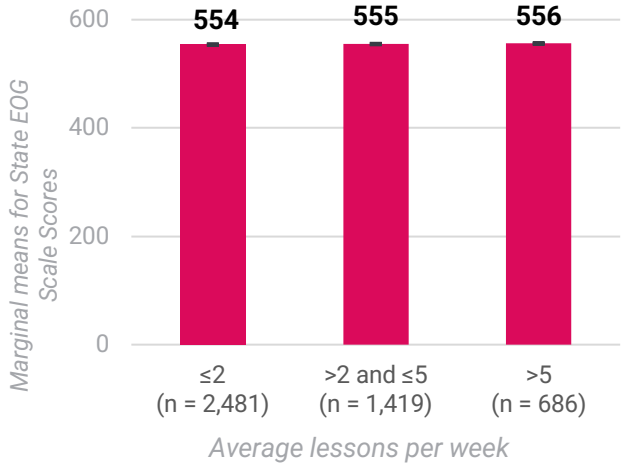
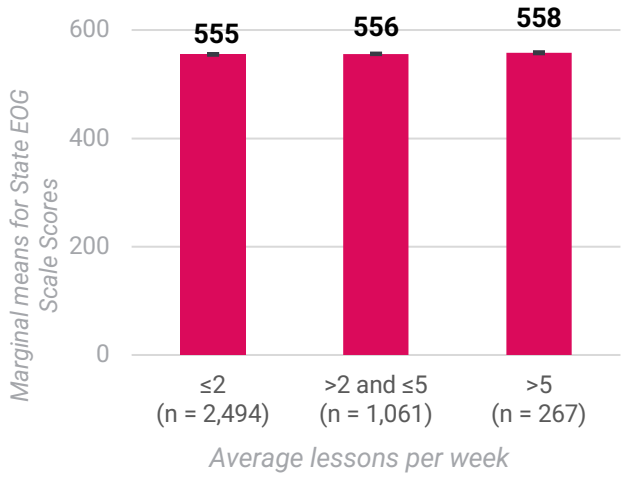


Table E4. Association between grade 5 DreamBox Math usage groups and spring 2023 state EOG mathematics assessment scaled scores

Predictor	Unstd. Beta Coefficient	Standard Error	Test statistic	*p*-value
Moderate Use vs. Low Use (Hedges' *g* = 0.16)	1.17	0.20	5.81	<.001
High Use vs. Low Use (Hedges' *g* = 0.32)	2.31	0.36	6.33	<.001
High Use vs. Moderate Use (Hedges' *g* = 0.24)	1.13	0.35	3.24	.001
Spring 2022 EOG scores	0.82	0.01	59.38	<.001
Gender	-0.40	0.15	-2.56	.01
School-level random effects	4.13	0.72	409.66	<.001

Among Grades 5 students classified as not proficient at the end of the prior year, those who completed 2–5 (moderate use) and more than 5 (high use) weekly lessons had significantly higher state EOG mathematics assessment scores than students who completed fewer than 2 lessons (low use).



Appendix F. Grades 6–8 Implementation and Outcome Findings

Table E1. Descriptive statistics for the weekly lessons' usage categories for grades 6–8 sample

Usage categories: weekly lessons	*n*	Mean	SD
≤ 2 weekly lessons	2,892	0.4	0.5
> 2 and ≤ 5 weekly lessons	228	2.9	0.8
> 5 weekly lessons	27	6.3	1.5

Table E2. Descriptive statistics for the weekly minutes' usage categories for grades 6–8 sample

Usage categories: weekly minutes	*n*	Mean	SD
≤ 30 weekly minutes	2,751	8	8
> 30 and ≤ 60 weekly minutes	337	40	8
> 60 weekly minutes	59	73	12

Given the relatively small sample of DreamBox Math users at grades 6–8 (*n* = 3,147) and right-skewed usage distribution (Table E1 and 2), researchers conducted linear regression analyses to investigate how the use of DreamBox Math was related student math achievement from fall 2022 to spring 2022. Analyses included student-level covariates to control for potential selection bias. None of the findings for grades 6–8 students were statistically significant.

Association Between Average Weekly Lessons and 6–8 Students' Outcomes on NWEA MAP®

Table E3. Association between grades 6–8 DreamBox Math usage and spring 2023 NWEA MAP® mathematics percentile scores

Predictor	Unstd. Beta Coefficient	Standard Error	Test statistic	*p*-value
Weekly Lessons	0.42	0.24	1.73	.083
Fall 2023 NWEA MAP® percentile scores	0.85	0.01	60.35	<.001
Grade	2.55	0.29	8.78	<.001
Race	0.32	0.13	2.45	.014

Association Between Average Weekly Minutes and 6–8 Students' Outcomes on NWEA MAP®

Table E3. Association between grade 6–8 DreamBox Math usage and spring 2023 NWEA MAP® mathematics percentile scores

Predictor	Unstd. Beta Coefficient	Standard Error	Test statistic	*p*-value
Weekly Minutes	0.00	0.02	-0.09	.930
Fall 2023 NWEA MAP® percentile scores	0.85	0.01	60.30	<.001
Grade	2.50	0.29	8.62	<.001
Race	0.33	0.13	2.50	.012