THE EDTECH COLLECTIVE
Instructure Partner Ecosystem

DREAMBOX MATH

ESSA Level III Study (2023–24)

Prepared for:
Discovery Education®

Prepared by Instructure:
Meetal Shah, Ph.D., Lead Researcher
Andrew Scanlan, M.A., Senior Researcher

May 2025

INSTRUCTURE
LEVEL 3 ESSA
2025

EXECUTIVE SUMMARY

Discovery Education contracted with Instructure, a third-party edtech research company, to examine the impact of *DreamBox Math* (from hereon, *DreamBox*) usage on student math outcomes. Instructure designed the study to satisfy Level III requirements (Promising Evidence) according to the Every Student Succeeds Act (ESSA, 2015).

Study Sample and Methodology

This study was conducted with data from the 2023–24 school year and included 8,393 K–8 students across 42 elementary and middle schools in one public school district in California. Researchers conducted analyses by grade band to allow for better interpretability of findings: Kindergarten–grade 2 (35%); grades 3–5 (37%); and grades 6–8 (28%). In terms of demographics, the total sample was racially diverse[1] and included White (49%), Asian (23%), Filipino (2%), and Black/African American (2%) students. Seventy-five percent of the sample identified as Hispanic. In terms of socioeconomic status (SES), 77% of students were classified as economically disadvantaged. Twenty-one percent of students were designated as English learners (EL). Finally, 12% percent of students were classified as students with disabilities.

Researchers used three measures to provide insights into *DreamBox* implementation and potential impacts of *DreamBox* on student math 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* use related to student math outcomes controlling for prior math achievement. The analyses included student-level covariates to control for potential selection bias (i.e., grade-level, ethnicity, EL status, disability status, and SES). In addition, researchers calculated standardized effect sizes (Hedges' *g*) to determine the magnitude of changes in student outcomes.

[1] Twenty-three percent of the sample declined to report their racial background. As a result, race was excluded from the analytic models to avoid issues related to missing data and the loss of a substantial portion of the sample. However, subgroup analyses were conducted to examine how *DreamBox* usage was associated with outcomes for students from specific demographic backgrounds.

Main Research Findings

Main Research Findings	
	Grades K–2
+	K–2 students who completed more *DreamBox* lessons per week had higher winter 2024 NWEA MAP® math scores.
+	In subgroup analyses, K–2 Hispanic students, students from low socioeconomic backgrounds, English learners, and low achievers who completed more DreamBox lessons per week had higher NWEA MAP® scores in winter 2024.
	Grades 3–5
+	Grades 3–5 students who completed more *DreamBox* lessons per week had higher winter 2024 NWEA MAP® math scores.
+	In subgroup analyses, grades 3–5 Hispanic students, Hispanic students, students from low SES backgrounds, English learners, and students with disabilities who completed more *DreamBox* lessons per week, had higher winter 2024 NWEA MAP® scores.
+	Grades 3–5 students who completed 2–5 weekly lessons (moderate use) and more than 5 (high use) weekly lessons had significantly higher state EOG mathematics math assessment scores than students who completed fewer than 2 weekly lessons (low use).
+	Grades 3–5 Hispanic students who completed 2–5 weekly lessons (moderate use) and more than 5 (high use) weekly lessons had significantly higher state EOG mathematics math assessment scores than students who completed fewer than 2 weekly lessons (low use).
	Grades 6–8
+	In subgroup analyses, only grade 6–8 students from low SES backgrounds who completed more *DreamBox* lessons per week, had higher winter 2024 NWEA MAP® scores.

Note: All the above findings were statistically significant at $p < .05$.

Conclusions

Given the positive findings, this study provides results to satisfy ESSA evidence requirements for Level III (Promising Evidence).

TABLE OF CONTENTS

INTRODUCTION

In the wake of the COVID-19 pandemic, student performance in math in the United States experienced significant setbacks. While the immediate disruptions have subsided, the impact on learning persists. The 2024 administration of the National Assessment of Educational Progress (NAEP) revealed that fourth-grade math scores increased by 2 percentile points compared to 2022, yet they remain 3 percentile points below pre-pandemic levels (National Center for Education Statistics [NCES], 2024). Eighth-grade math scores, however, showed no significant change from 2022 but were 8 points lower compared to 2019, highlighting ongoing challenges (NCES, 2024).

Internationally, the U.S. ranked 26th on the 2022 Programme for International Student Assessment (PISA), with fewer than 1 in 10 students achieving advanced levels and over a third failing to meet basic achievement standards (Organization for Economic Co-operation and Development [OECD], 2023). In terms of eighth-grade math, the U.S. ranks 24th out of 45 education systems (Trends in International Mathematics and Science Study [TIMSS], 2023). Additionally, 71% of high school seniors who took the ACT college entrance exam in 2024 scored below the college readiness benchmark (ACT, 2024). By the end of the 2023–24 school year, eighth-grade students were still approximately 0.27 standard deviations behind pre-pandemic norms in math (Lewis & Kuhfeld, 2024).

Discovery Education recognizes the urgency of addressing these ongoing challenges. The *DreamBox* K–8 Math learning solution aims to accelerate the acquisition of critical foundational math skills in elementary and middle school, setting students up for success in Algebra I and beyond (Peters, 2024).

As part of their ongoing efforts to demonstrate the effectiveness of their solution, *Discovery Education* contracted with Instructure, a third-party edtech research company, to examine the impact of *DreamBox* on elementary school students' math learning outcomes. Using the Every Student Succeeds Act (ESSA) standards as guidance in developing a study design, findings in this report align with Level III requirements (Promising Evidence). The following research questions guided this study:

Implementation

1. How many *DreamBox* lessons were completed per week by students during the 2023–24 school year?
2. Among *DreamBox* users, what were the usage patterns?

Student Outcomes

3. Did students' NWEA MAP® and end-of-grade (EOG) state assessment scores for math improve significantly, when controlling for prior math achievement and student demographics?
    a. How did the findings differ by subgroup (English learners, low SES, disability status)?
    b. How did the findings differ by grade band and achievement level (i.e., students in the lowest achievement quartile for their grade band)?

This report details the study design and methods, implementation, findings, and conclusions.

STUDY DESIGN AND METHODS

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

Study Design

This study used a correlative design to align with ESSA Level III evidence standards. It included all students in the district who used *DreamBox* during the 2023–24 school year.

Setting and Participants

This study was conducted with data from the 2023–24 school year and included 8,393 K–8 students across 42 elementary and middle schools in one public school district in California. Researchers conducted analyses by grade band to allow for better interpretability of findings: Kindergarten–grade 2 (35%); grades 3–5 (37%); and grades 6–8 (28%). In terms of demographics, the total sample was racially diverse[2] and included White (49%), Asian (23%), Filipino (2%), and Black/African American (2%) students. Seventy-five percent of the sample identified as Hispanic. In terms of socioeconomic status (SES), 77% of students were classified as economically disadvantaged. Twenty-one percent of students were designated as English learners (EL). Finally, 12% percent of students were classified as students with disabilities.

Measures

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

DreamBox Usage Metrics. Researchers utilized 2023–24 student-level usage data to inform the extent to which students used *DreamBox* during the school year and whether students' use of *DreamBox* was related to outcomes. According to the *Discovery Education* 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) and posttest (i.e., and winter 2024) for the kindergarten to grade 8 sample. For the EOG analysis, researchers used spring 2023 scores as the baseline measure and spring 2024 as the outcome measure for

[2] Twenty-three percent of the sample declined to report their racial background. As a result, race was excluded from the analytic models to avoid issues related to missing data and the loss of a substantial portion of the sample. However, subgroup analyses were conducted to examine how *DreamBox* usage was associated with outcomes for students from specific demographic backgrounds.

the grades 3–7 sample. The state uses Smarter Balanced assessments, which have vertically scaled scores. This allows for analysis across a multi-grade sample.

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* use related to student math outcomes controlling for their prior math achievement. The analyses included school-level random effects and student-level covariates to control for potential selection bias (i.e., grade-level, ethnicity, EL status, disability status, and SES). In addition, researchers calculated standardized effect sizes (Hedges' *g*) to determine the magnitude of changes in student outcomes.

IMPLEMENTATION

The charts below highlight *DreamBox* use during the 2023–24 school year based on *DreamBox* internal usage data (Table 1; details in Appendix B).

Table 1. *DreamBox* average weekly lessons completed by grade band for NWEA sample

Grade	*n*	Average (# of Weekly Lessons)	SD	Minimum	Maximum
K–2	2,984	2.8	2.0	0.04	28.2
3–5	3,094	2.7	2.2	0.02	22.5
6–8	2,351	1.3	1.7	0.02	18.4

Researchers used *DreamBox'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*. For average weekly lessons, K–8 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).

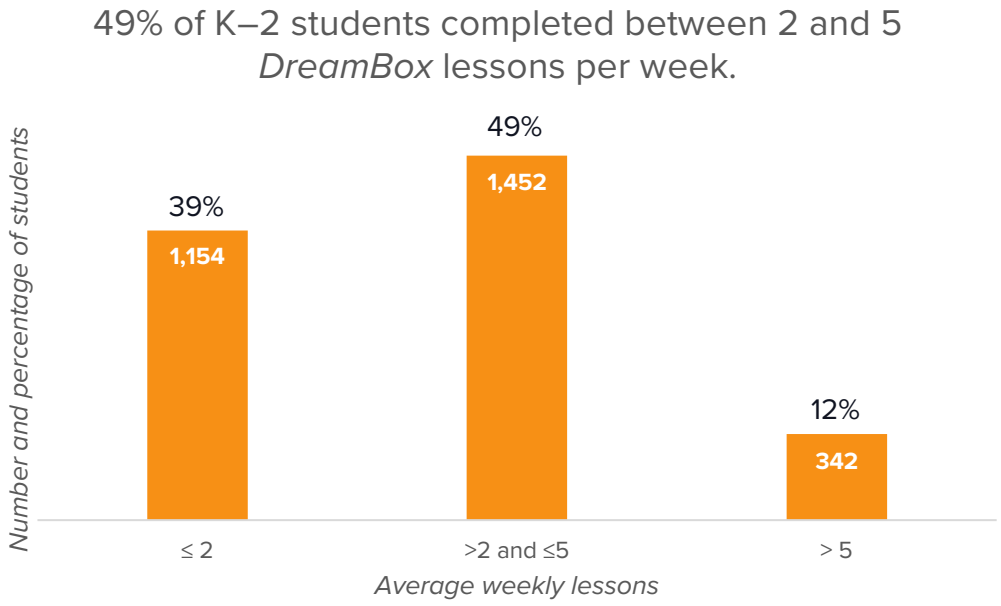


Figure 1. Overall distribution of average weekly lessons completed on DreamBox by grade K–2 students (*n* = 2,948)

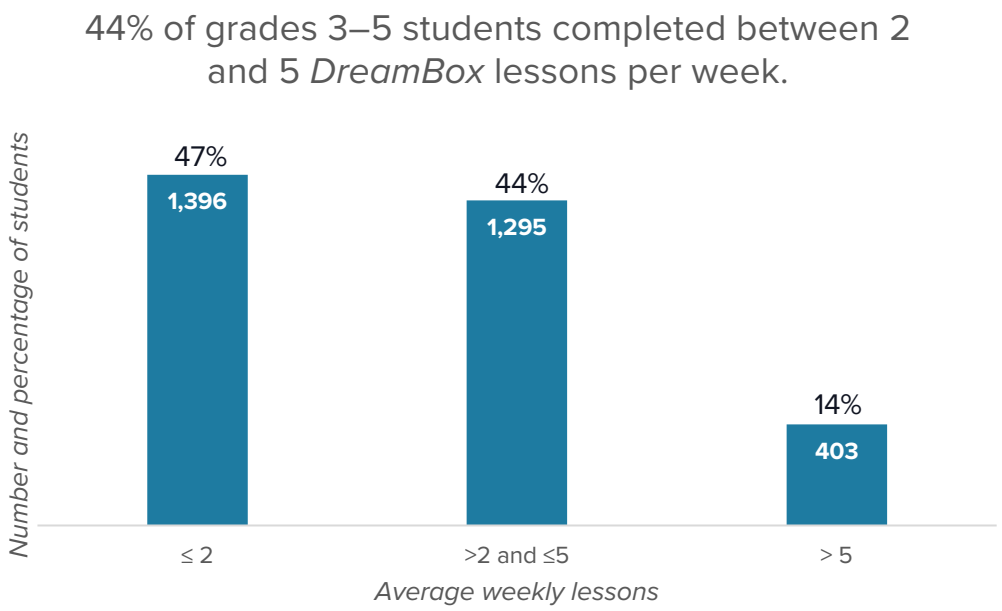


Figure 2. Overall distribution of average weekly lessons completed on *DreamBox* by grade 3–5 students (*n* = 3,094)

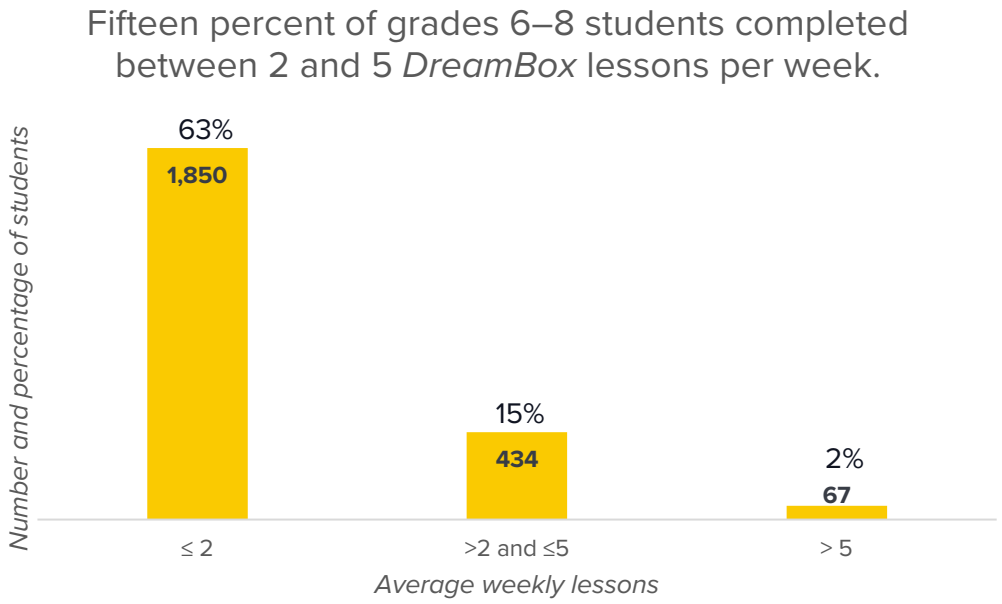


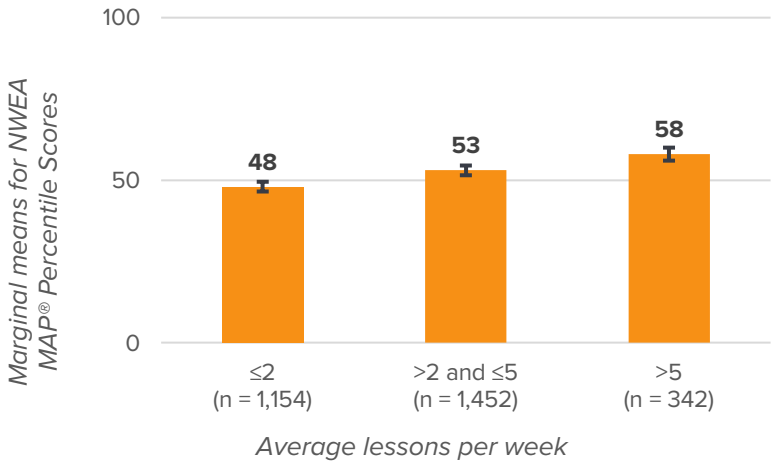
Figure 3. Overall distribution of average weekly lessons completed on DreamBox by grade 6–8 students (*n* = 2,351)

NWEA MAP® OUTCOME FINDINGS FOR K–8 STUDENTS

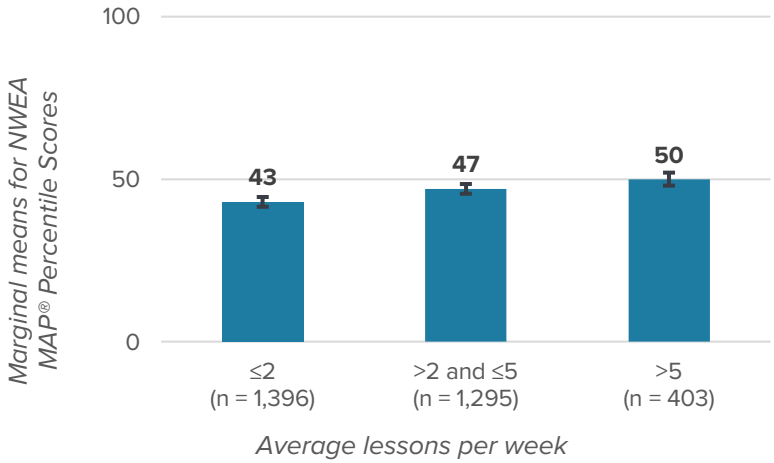
Researchers used two-level multilevel modeling analysis (i.e., students nested in schools) to examine how *DreamBox* use related to student math outcomes controlling for prior math achievement. These analyses also included student-level covariates to control for potential selection bias. To allow for better interpretability of results, marginal means charts are presented below. The vertical lines at the top of each bar represent a 95% confidence interval (see Appendix C for more details about the model and the corresponding Hedges' *g* effect sizes).

Association Between Average Weekly Lessons and K–8 Students' Outcomes on NWEA MAP® by Usage Groups

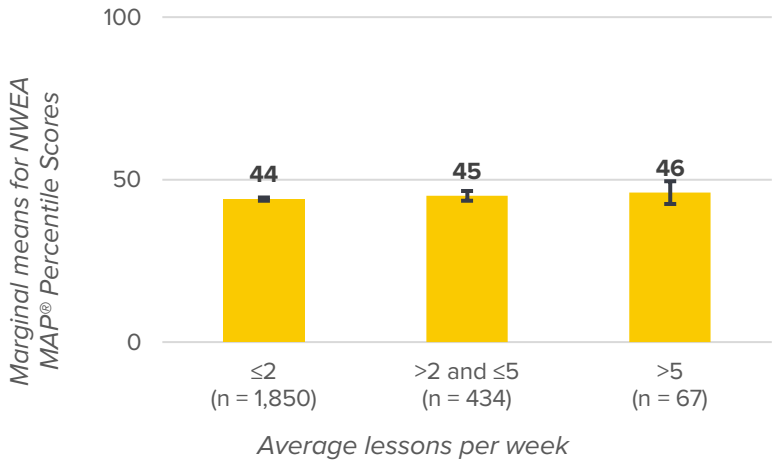
Grade K–2 students who completed 2–5 (moderate use) and more than 5 (high use) weekly lessons had **higher** NWEA MAP® scores than students who completed fewer than 2 lessons (low use). These results were statistically significant at the $p < .05$ level.



Grade 3–5 students who completed 2–5 (moderate use) and more than 5 (high use) weekly lessons had **higher** NWEA MAP® scores than students who completed fewer than 2 lessons (low use). These results were statistically significant at the $p < .05$ level.

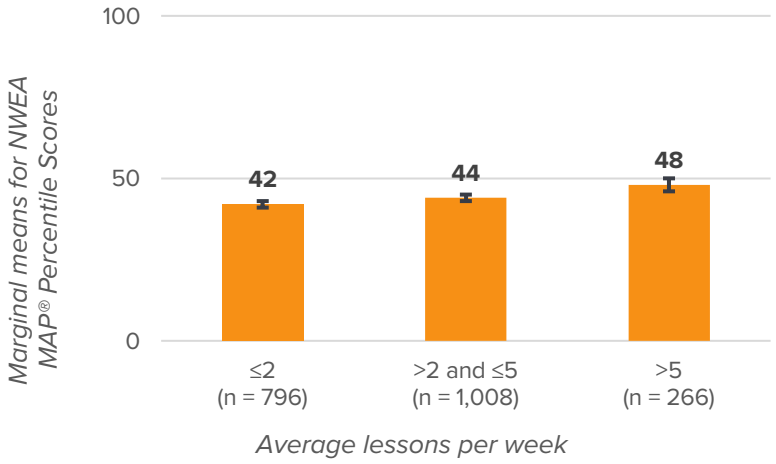


Grade 6–8 students who completed 2–5 (moderate use) and more than 5 (high use) weekly lessons had **similar** NWEA MAP® scores compared to students who completed fewer than 2 lessons (low use). These results were not statistically significant.

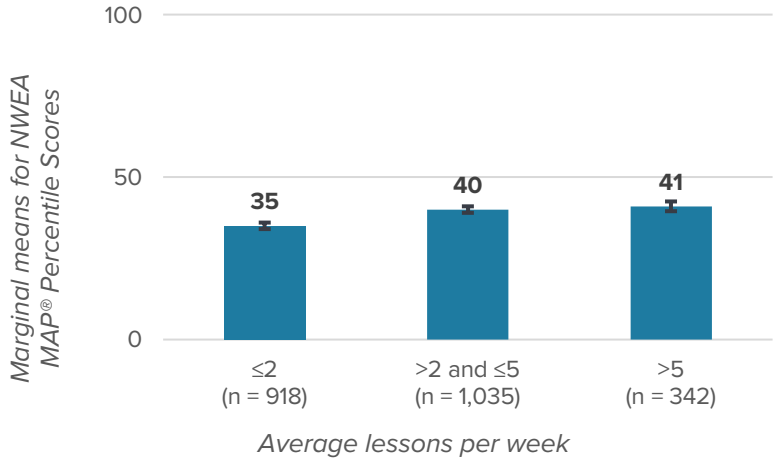


Association Between Average Weekly Lessons and K–8 Students' Outcomes on NWEA MAP® for Hispanic Students

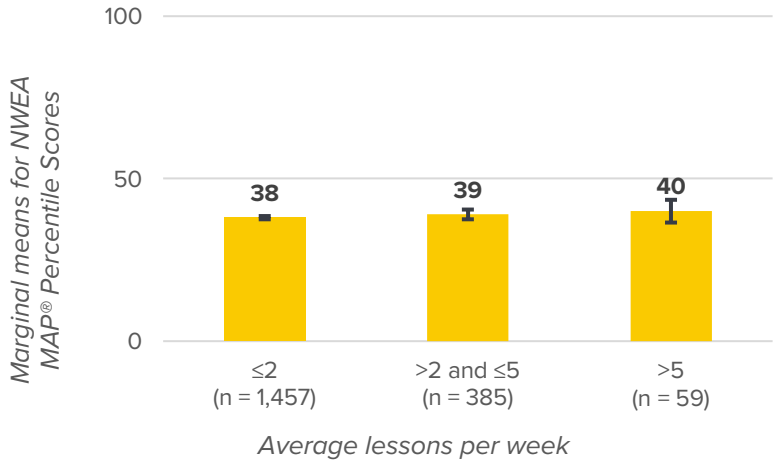
Grade K–2 Hispanic students who completed 2–5 (moderate use) and more than 5 (high use) weekly lessons had **higher** NWEA MAP® scores than students who completed fewer than 2 lessons (low use). These results were statistically significant at the p <.05 level.



Grade 3–5 Hispanic students who completed 2–5 (moderate use) and more than 5 (high use) weekly lessons had **higher** NWEA MAP® scores than students who completed fewer than 2 lessons (low use). These results were statistically significant at the p <.05 level.



Grade 6–8 Hispanic students who completed 2–5 (moderate use) and more than 5 (high use) weekly lessons had **similar** NWEA MAP® scores compared to students who completed fewer than 2 lessons (low use). These results were not statistically significant.



Association Between Average Weekly Lessons and K–8 Students' Outcomes on NWEA MAP® by Subgroups

To determine whether the positive findings for the full sample at each grade band hold for demographic subgroups (i.e., low SES students, EL students, students with disabilities) and achievement subgroups (i.e., lowest achievement quartile), researchers conducted linear regressions to examine how use of *DreamBox* related to students' winter 2024 math achievement after controlling for prior math achievement and any statistically significant demographic characteristics.

As illustrated by point estimates in the figure 4, among low SES 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.

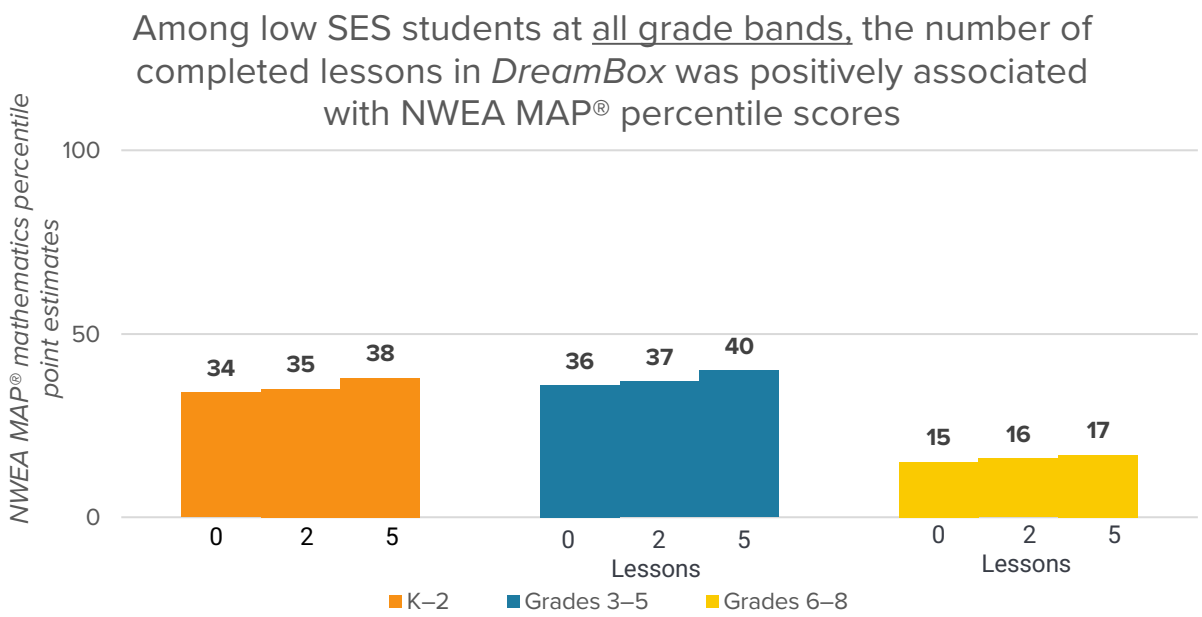


Figure 4. NWEA MAP® mathematics percentile point estimates for low SES students completing 2 and 5 *DreamBox* lessons per week.

While percentile point estimates for EL students, students with disabilities, and students in the lowest achievement quartile in grades K–2 and 3–5 are presented below, we have excluded visualizations for grades 6–8. This decision was made because the linear regression model for this grade band yielded a negative constant term, which limits the interpretability of the results. For transparency, the full model output for grades 6–8 is included in Appendix C.

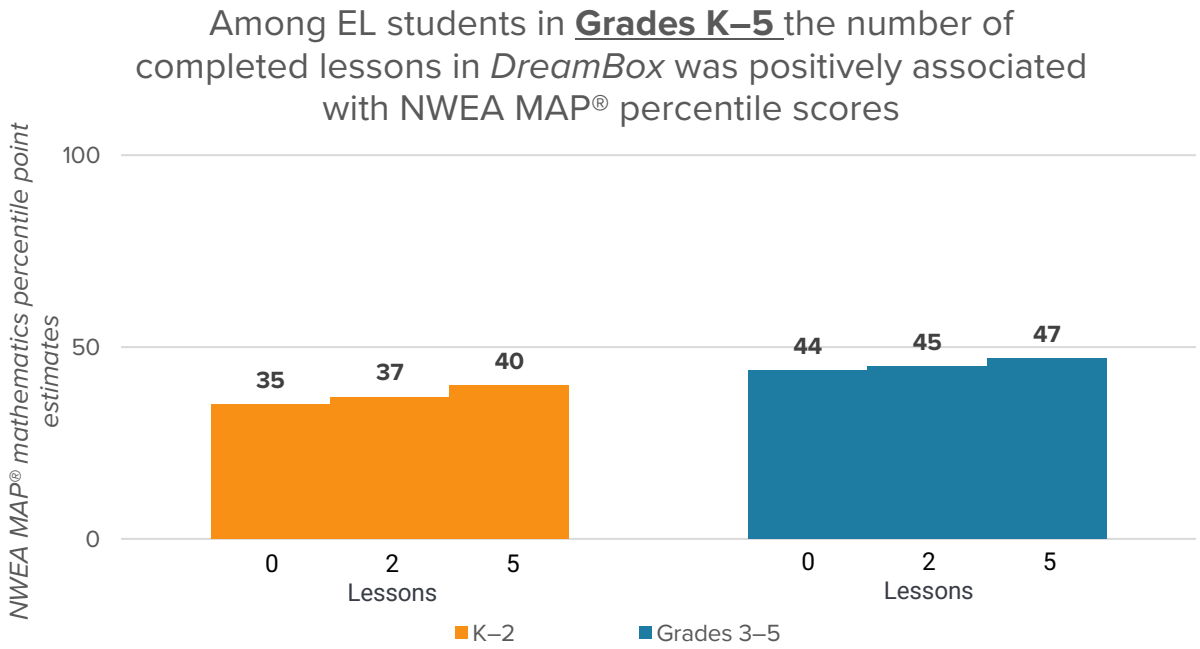


Figure 5. NWEA MAP® mathematics percentile point estimates for EL students completing 2 and 5 *DreamBox* lessons per week.

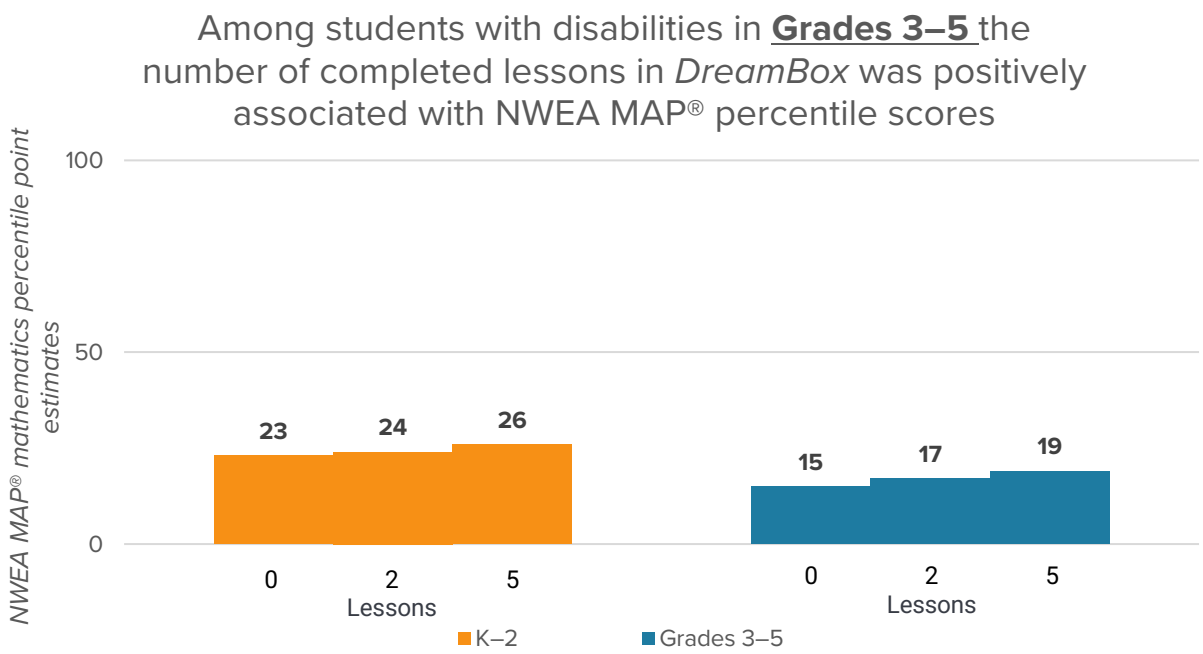


Figure 6. NWEA MAP® mathematics percentile point estimates for students with disabilities completing 2 and 5 *DreamBox* lessons per week.

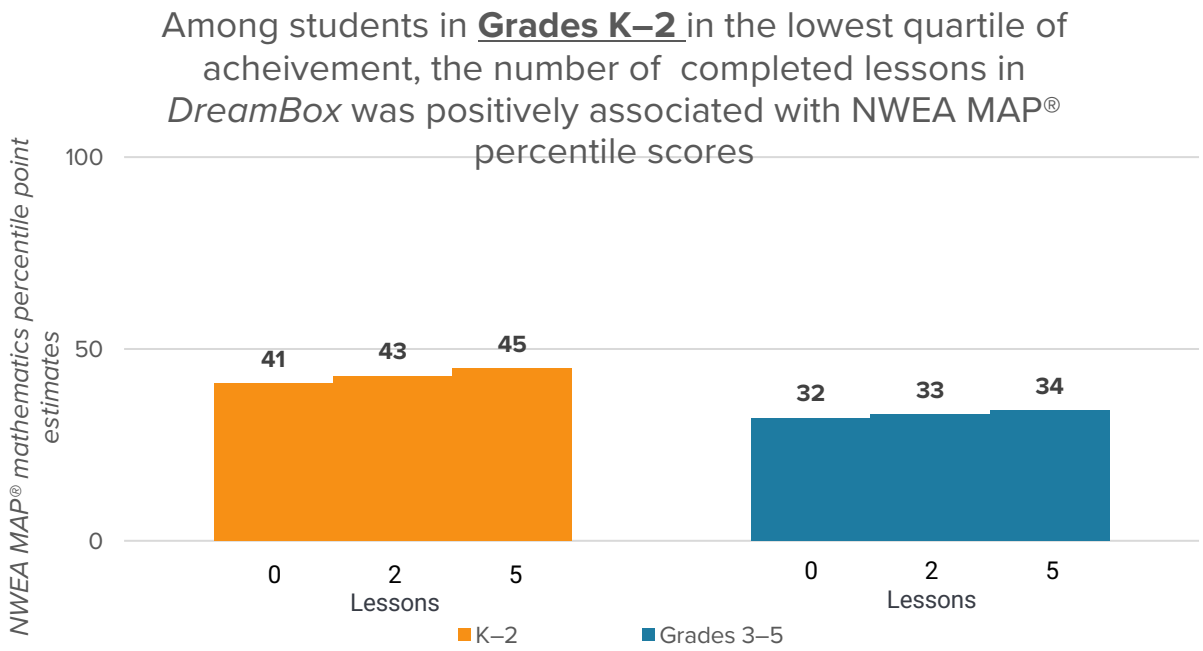
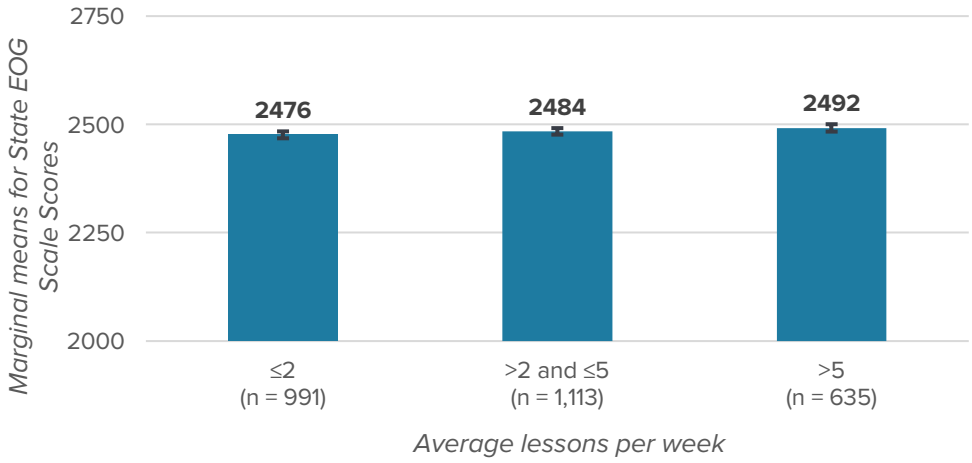


Figure 7. NWEA MAP® mathematics percentile point estimates for student in the lowest quartile at baseline completing 2 and 5 *DreamBox* lessons per week.

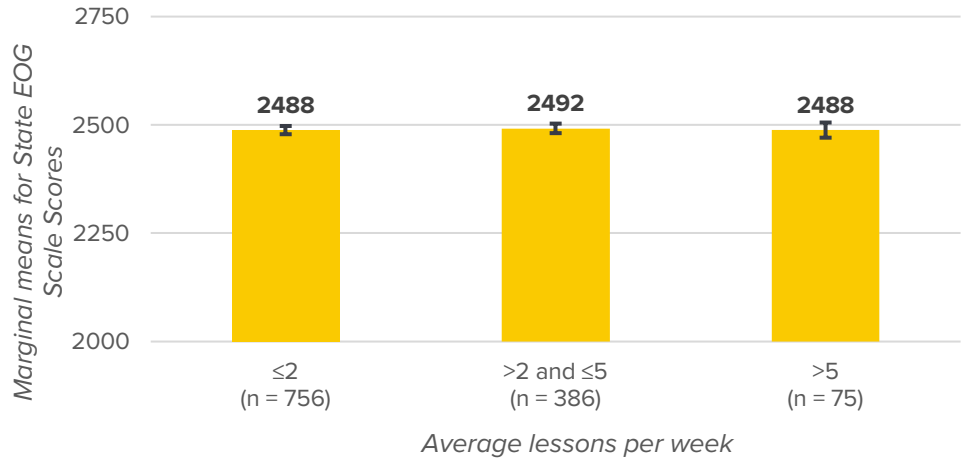
STATE EOG OUTCOME FINDINGS FOR GRADES 3–7 STUDENTS

Researchers examined whether greater usage of *DreamBox* related to higher spring 2024 state EOG assessment achievement using spring 2023 (prior year) EOG scores and student demographic characteristics as covariates. Researchers used two-level multilevel modeling analysis (i.e., students nested in schools) to examine how *DreamBox* use related to student math outcomes. To allow for better interpretability of results, marginal means charts are presented below. The vertical lines at the top of each bar represent a 95% confidence interval (see Appendices D 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 vertically scaled. Researchers conducted the analysis separately by grade-band. As such, EOG findings are only listed for grades 3–5 and grades 6–7 students in the sample.

Grade 3–5 students who completed 2–5 (moderate use) and more than 5 (high use) weekly lessons had **higher** state EOG mathematics assessment scores than students who completed fewer than 2 lessons (low use). These results were statistically significant at the $p < .05$ level.



Grade 6–7 students who completed 2–5 (moderate use) and more than 5 (high use) weekly lessons had **similar** state EOG mathematics assessment scores compared to students who completed fewer than 2 lessons (low use). These results were not statistically significant.



Summary of Effect Sizes for NWEA MAP® and EOG State Math Assessment Findings

Researchers calculated Hedges' *g* effect sizes (Hedges, 1981) to compare results across different assessments and grade bands. Hedges' *g* effect sizes are categorized as small (<0.20), medium (0.20–0.49), and large (≥0.50). Alternative methods for quantifying effect sizes exist. For instance, Kraft (2020) argues that due to the rarity of effect sizes exceeding 0.20 in educational research, a more appropriate range would be small (<0.05), medium (0.05–0.19), and large (≥0.20). Kraft also emphasizes that effect sizes from correlational studies should not be interpreted as causal effects, and such studies typically yield larger effect sizes compared to causal studies.

Using these thresholds, the study found statistically significant large effects on NWEA MAP® math outcomes from using DreamBox for more than five weekly lessons compared to two or fewer weekly lessons for K–2 and grade 3–5 students. For Grades 6–8 students, the effects were mostly positive but none of the findings were statistically significant (Table 2).

Table 2. Summary of effect sizes for NWEA MAP® Findings

Grade	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
K–2	**0.14***	**0.31***	**0.18***
3–5	**0.14***	**0.23***	**0.10***
6–8	0.04	0.07	0.04

The study also found statistically significant effects on EOG state math outcomes from using *DreamBox* for more than five weekly lessons compared to two or fewer weekly lessons for grade 3–5 students. For grades 6–7 students, findings were **not** statistically significant (Table 3).

Table 3. Summary of effect sizes for EOG State Math Assessment Findings

Grade	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
3–5	**0.08***	**0.16***	**0.08***
6–8	0.04	0.00	-0.03

CONCLUSIONS AND RECOMMENDATIONS

The findings support an association between *DreamBox* usage and improved NWEA MAP® outcomes for grades K–5 students. This trend was also observed within various subgroups in these grade levels, including Hispanic students, students from low SES backgrounds, and English learners.

Findings also indicate that grades K–5 students with moderate (2–5 weekly lessons) and high (more than 5 weekly lessons) *DreamBox* usage achieved significantly higher state EOG mathematics math assessment scores compared to students with low usage (fewer than 2 lessons). This trend was also observed for Hispanic students in this grade range.

For grades 6–8 students, subgroup analyses also revealed that students from low SES backgrounds who completed more *DreamBox* lessons per week also demonstrated higher winter 2024 NWEA MAP® math scores.

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* compare to elementary school students using other math programs.
- For the middle grades sample, it may be valuable to assess whether differences in implementation fidelity, usage patterns, or contextual factors (e.g., instructional time, curriculum alignment) are affecting outcomes for grades 6–8. Gathering more detailed qualitative or quantitative data on these factors could help clarify the findings.

REFERENCES

ACT. (2024). *The condition of college and career readiness: 2024 graduating class*. URL

Every Student Succeeds Act, Pub. L. No. 114-95 (2015). URL

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

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

Lewis, K., & Kuhfeld, M. (2024, July). *Recovery still elusive: 2023–24 student achievement highlights persistent achievement gaps and a long road ahead*. NWEA. URL

National Center for Education Statistics. (2024). *The Nation's Report Card: 2024 mathematics and reading assessments*. U.S. Department of Education, Institute of Education Sciences. URL

Organisation for Economic Co-operation and Development. (2023). *PISA 2022 results (Volume I): The state of learning and equity in education*. OECD Publishing. https://doi.org/10.1787/53f23881-en

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

TIMSS & PIRLS International Study Center. (2023). *TIMSS 2023 international results in mathematics and science*. Boston College, Lynch School of Education. URL

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 URL

APPENDIX A. DREAMBOX MATH LOGIC MODEL

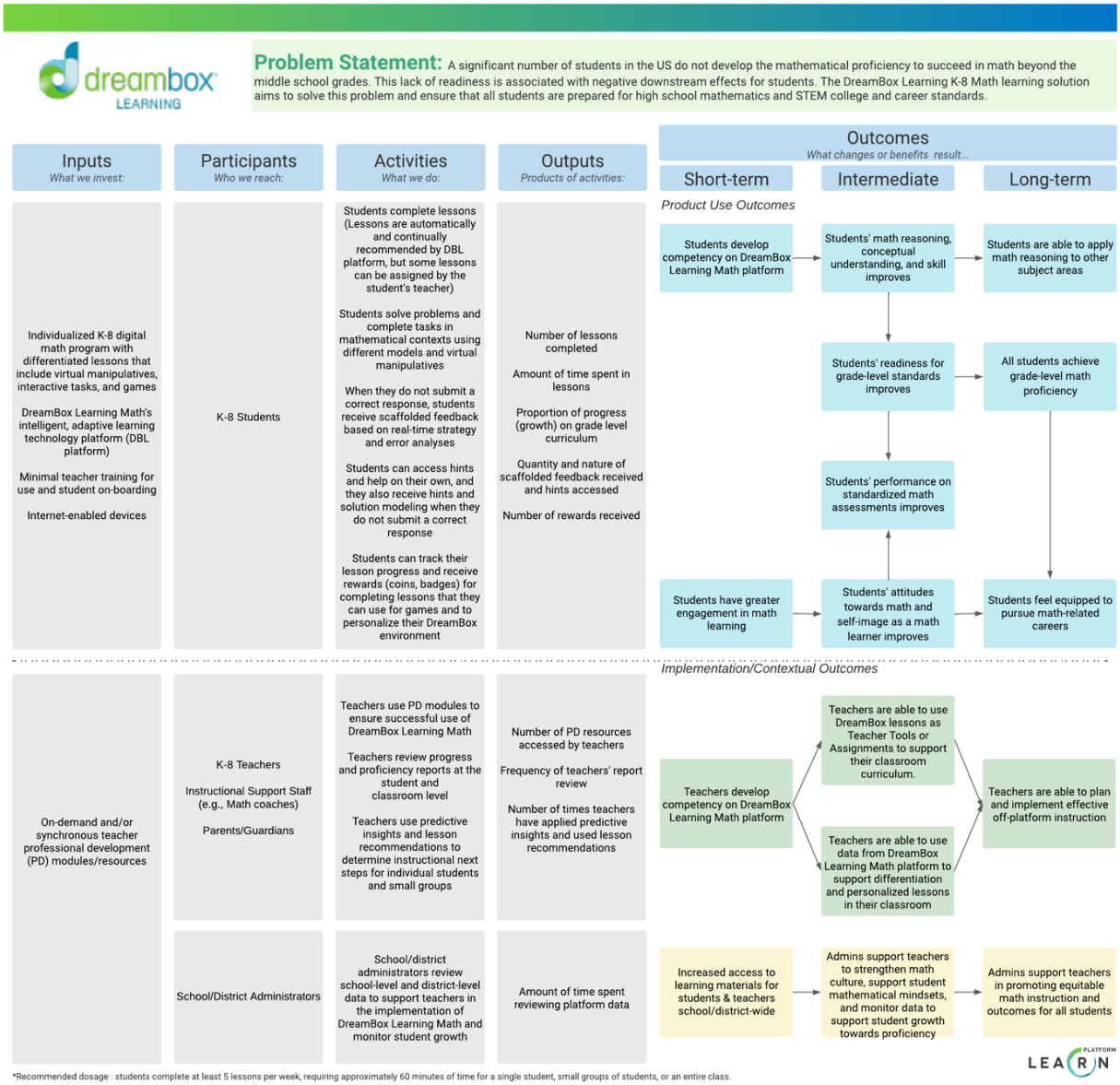


Figure 1. DreamBox Learning Math logic model

Prepared for DreamBox Learning, September 2022

APPENDIX B. ADDITIONAL INFORMATION ON PROGRAM IMPLEMENTATION

Table B1. Descriptive statistics for the weekly lessons' usage categories for **grades K–2 NWEA** sample

Usage categories: weekly lessons		*n*	Mean	SD	Min	Max
Low	≤ 2 weekly lessons	1,154	1.19	0.50	0.04	2.00
Moderate	> 2 and ≤ 5 weekly lessons	1,452	3.22	0.81	2.02	4.98
High	> 5 weekly lessons	342	6.85	2.31	5.02	28.19

Table B2. Descriptive statistics for the weekly lessons' usage categories for **grades 3–5 NWEA** sample

Usage categories: weekly lessons		*n*	Mean	SD	Min	Max
Low	≤ 2 weekly lessons	1,396	1.02	0.58	0.02	2.00
Moderate	> 2 and ≤ 5 weekly lessons	1,295	3.25	0.85	2.02	5.00
High	> 5 weekly lessons	403	6.84	2.12	5.02	22.50

Table B3. Descriptive statistics for the weekly lessons' usage categories for **grades 6–8 NWEA** sample

Usage categories: weekly lessons		*n*	Mean	SD	Min	Max
Low	≤ 2 weekly lessons	1,850	0.58	0.56	0.02	2.00
Moderate	> 2 and ≤ 5 weekly lessons	434	3.11	0.76	2.02	4.98
High	> 5 weekly lessons	67	7.54	2.96	5.04	18.44

Table B4. Descriptive statistics for the weekly lessons' usage categories for t**3rd to 5th grade grades 3–5 State EOG** sample

Usage categories: weekly lessons		*n*	Mean	SD	Min	Max
Low	≤ 2 weekly lessons	991	0.98	0.56	0.03	2.00
Moderate	> 2 and ≤ 5 weekly lessons	1,113	3.43	0.84	2.02	4.98
High	> 5 weekly lessons	635	7.54	2.35	5.00	23.95

Table B5. Descriptive statistics for the weekly lessons' usage categories for **grades 6–8 State EOG** sample

Usage categories: weekly lessons		*n*	Mean	SD	Min	Max
Low	≤ 2 weekly lessons	756	0.86	0.60	0.03	2.00
Moderate	> 2 and ≤ 5 weekly lessons	386	3.46	0.84	2.02	4.95
High	> 5 weekly lessons	75	6.58	3.26	5.00	27.86

APPENDIX C. ADDITIONAL INFORMATION ON NWEA MAP® OUTCOME FINDINGS FOR K–8 STUDENTS

Overall Association Between Average Weekly Lessons and K–8 Students' Outcomes on NWEA MAP® by Usage Groups

Table C1. Association between **grades K–2** *DreamBox* usage groups and winter 2024 NWEA MAP® mathematics percentile scores

Predictor	Unstd. Beta Coefficient	Standard Error	Test statistic[3]	*p*-value
Moderate Use vs. Low Use (Hedges' *g* = **0.14***)	4.05	0.69	5.83	<.001
High Use vs. Low Use (Hedges' *g* = **0.31***)	9.15	1.14	8.00	<.001
High Use vs. Moderate Use (Hedges' *g* = **0.18***)	5.10	1.07	4.79	<.001
Winter 2023 NWEA MAP® percentile scores	0.69	0.01	57.54	<.001
Ethnicity (Hispanic or not)	-6.39	0.86	-7.42	<.001
Disability status	-4.50	1.04	-4.31	<.001
SES	-1.76	0.76	-2.33	.020
School-level random effects	6.67	2.98	24.86	<.001

Table C2. Association between **grades 3–5** *DreamBox* usage groups and winter 2024 NWEA MAP® mathematics percentile scores

Predictor	Unstd. Beta Coefficient	Standard Error	Test statistic	*p*-value
Moderate Use vs. Low Use (Hedges' *g* = **0.14***)	3.81	0.62	6.11	<.001
High Use vs. Low Use (Hedges' *g* = **0.23***)	6.60	0.99	6.68	<.001
High Use vs. Moderate Use (Hedges' *g* = **0.10***)	2.79	0.89	3.13	.002

[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
Winter 2023 NWEA MAP® percentile scores	0.75	0.01	68.83	<.001
Grade	-1.73	0.31	-5.49	<.001
Ethnicity (Hispanic or not)	-4.49	0.76	-5.92	<.001
Disability status	-4.63	0.81	-5.70	<.001
SES	-1.93	0.65	-2.95	<.001
School-level random effects	11.169	3.90	96.16	<.001

Table C3. Association between **grades 6–8** *DreamBox* usage groups and winter 2024 NWEA MAP® mathematics percentile scores

Predictor	Unstd. Beta Coefficient	Standard Error	Test statistic	*p*-value
Moderate Use vs. Low Use (Hedges' g = 0.04)	0.10	0.74	1.35	.179
High Use vs. Low Use (Hedges' g = 0.07)	1.98	1.71	1.16	.246
High Use vs. Moderate Use (Hedges' g = 0.04)	0.99	1.78	0.55	.580
Winter 2023 NWEA MAP® percentile scores	0.85	0.01	64.90	<.001
Grade	1.71	0.36	4.75	<.001
Ethnicity (Hispanic or not)	-6.01	0.79	-7.65	<.001
ELL status	-1.80	0.81	-2.23	.026
Disability status	-5.99	0.84	-7.09	<.001

Association Between Average Weekly Lessons and K–8 Hispanic Students' Outcomes on NWEA MAP® by Usage Groups

Table C4. Association between **grades K–2** *DreamBox* usage groups and winter 2024 NWEA MAP® mathematics percentile scores for **Hispanic Students**

Predictor	Unstd. Beta Coefficient	Standard Error	Test statistic	*p*-value
Moderate Use vs. Low Use (Hedges' *g* = **0.10***)	2.50	0.81	3.07	.002
High Use vs. Low Use (Hedges' *g* = **0.23***)	6.26	1.22	5.14	<.001
High Use vs. Moderate Use (Hedges' *g* = **0.14***)	3.76	1.18	3.20	.001
Winter 2023 NWEA MAP® percentile scores	0.71	0.01	47.84	<.001
Grade	1.07	0.47	2.30	.022
ELL status	-1.91	0.93	-2.05	.040
Disability status	-4.01	1.16	-3.45	.001
SES	-3.24	1.01	-3.21	.001

Table C5. Association between **grades 3–5** *DreamBox* usage groups and winter 2024 NWEA MAP® mathematics percentile scores for **Hispanic Students**

Predictor	Unstd. Beta Coefficient	Standard Error	Test statistic	*p*-value
Moderate Use vs. Low Use (Hedges' *g* = **0.17**)	4.15	0.67	6.23	<.001
High Use vs. Low Use (Hedges' *g* = **0.23**)	5.62	0.93	6.05	<.001
High Use vs. Moderate Use (Hedges' *g* = 0.06)	1.48	0.92	1.61	.106
Winter 2023 NWEA MAP® percentile scores	0.74	0.01	58.71	<.001
Grade	-2.04	0.38	-5.41	<.001
Disability status	-4.31	0.89	-4.85	<.001
SES	-2.10	0.85	-2.47	.013

Table C6. Association between **grades 6–8** *DreamBox* usage groups and winter 2024 NWEA MAP® mathematics percentile scores for **Hispanic Students**

Predictor	Unstd. Beta Coefficient	Standard Error	Test statistic	*p*-value
Moderate Use vs. Low Use (Hedges' *g* = 0.05)	1.16	0.80	1.44	.149
High Use vs. Low Use (Hedges' *g* = 0.07)	1.73	1.84	0.94	.347
High Use vs. Moderate Use (Hedges' *g* = 0.02)	0.57	1.92	0.30	.765
Winter 2023 NWEA MAP® percentile scores	0.85	0.02	55.87	.001
Grade	1.72	0.41	4.23	.001
ELL status	-2.59	0.91	-2.85	.004
Disability status	-5.78	0.90	-6.39	.001

Association Between Average Weekly Lessons and K–8 Students' Outcomes on NWEA MAP® by Race and Demographic Subgroups

Table C7. Association between **grades K–2** *DreamBox* 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)	Mean Weekly Lessons	SD	Min	Max
Low SES Students	2,213	0.94	0.17	5.38	<.001	+5	2.87	2.08	0.04	28.19
ELLs	774	0.89	0.32	2.80	.005	+4	2.91	2.07	0.04	19.56
Student with Disabilities	305	0.55	0.42	1.32	.189	-	2.97	2.53	0.04	28.19
Lowest Achievement Quartile	737	0.79	0.30	2.66	.008	+4	2.74	2.23	0.04	28.19
Quad 1	518	0.10	0.30	0.33	.739	-	3.45	2.42	0.04	19.56
Quad 2	632	1.57	0.52	3.00	.003	+8	2.22	1.37	0.04	9.04
Quad 3	914	1.17	0.26	4.61	<.001	+6	2.62	2.11	0.04	17.81
Quad 4	884	1.79	0.31	5.76	<.001	+9	3.17	1.87	0.06	28.19

Table C8. Association between **grades 3–5** *DreamBox* 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)	Mean Weekly Lessons	SD	Min	Max
Low SES Students	2,363	0.87	0.14	6.39	<.001	+4	2.87	2.22	0.02	22.50
ELLs	621	0.64	0.25	2.54	.011	+3	2.95	2.37	0.02	22.50
Student with Disabilities	378	0.86	0.30	2.86	.004	+4	3.11	2.75	0.02	16.50
Lowest Achievement Quartile	786	0.26	0.22	1.18	.239	-	2.79	2.11	0.02	22.50

Subgroup	*N*	Unstd. Beta Coeff.	SE	*t*-value	*p*-value	Impact of 5 lessons (percentile point change)	Mean Weekly Lessons	SD	Min	Max
Quad 1	540	0.34	0.30	1.13	.260	-	3.56	2.11	0.08	22.50
Quad 2	670	2.03	0.40	5.10	<.001	+10	1.86	1.42	0.02	11.92
Quad 3	946	0.38	0.20	1.89	.059	-	2.09	2.47	0.02	16.00
Quad 4	938	1.18	0.25	4.67	<.001	+6	3.45	1.84	0.04	10.96

Table C9. Association between grades 6–8 *DreamBox* 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)	Mean Weekly Lessons	SD	Min	Max
Low SES Students	1,858	0.46	0.19	2.42	.016	+2	1.34	1.73	0.02	18.44
ELLs	365	0.26	0.35	0.74	.459	-	1.43	1.97	0.02	18.44
Student with Disabilities	360	0.34	0.27	1.25	.214	-	1.69	2.27	0.02	17.38
Lowest Achievement Quartile	603	0.13	0.24	0.57	.571	-	1.46	1.84	0.02	17.38
Quad 1	568	0.24	0.40	0.60	.551	-	1.71	1.50	0.02	7.08
Quad 2	467	0.64	0.34	1.88	.061	-	1.00	1.86	0.02	18.44
Quad 3	503	-0.90	0.86	-1.04	.297	-	0.37	0.67	0.02	8.50
Quad 4	813	0.77	0.31	2.45	.014	+4	1.61	1.80	0.02	17.38

APPENDIX D. ADDITIONAL INFORMATION ON STATE EOG OUTCOME FINDINGS FOR GRADES 3–7 STUDENTS

Overall Association Between Average Weekly Lessons and K–8 Students' Outcomes on State EOG Assessment by Usage Groups

Table D1. Association between **grades 3–5** *DreamBox* usage groups and spring 2024 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**)	7.93	2.66	2.99	.003
High Use vs. Low Use (Hedges' *g* = **0.16**)	16.13	3.47	4.65	<.001
High Use vs. Moderate Use (Hedges' *g* = **0.08**)	8.20	3.14	2.61	.009
Spring 2023 (prior year) EOG scores	0.82	0.01	58.12	<.001
Ethnicity (Hispanic or not)	-19.19	3.30	-5.81	<.001
ELL status	-13.56	2.80	-4.84	<.001
Disability status	-15.31	3.38	-4.53	<.001
SES	-6.49	2.82	-2.30	.022
School-level random effects	274.69	94.13	129.34	<.001

Table D2. Association between **grades 6–8** *DreamBox* usage groups and spring 2024 state EOG mathematics assessment scaled scores

Predictor	Unstd. Beta Coefficient	Standard Error	Test statistic	*p*-value
Moderate Use vs. Low Use (Hedges' *g* = 0.04)	3.86	4.60	0.84	.401
High Use vs. Low Use (Hedges' *g* = 0.00)	-0.09	8.23	-0.01	.992
High Use vs. Moderate Use (Hedges' *g* = -0.03)	-3.94	7.99	-0.49	.622
Spring 2023 (prior year) EOG scores	0.76	0.02	35.88	<.001

Predictor	Unstd. Beta Coefficient	Standard Error	Test statistic	*p*-value
Ethnicity (Hispanic or not)	-19.19	5.75	-3.34	.001
ELL status	-12.86	5.06	-2.54	.011
Disability status	-28.09	5.42	-5.18	<.001