Happy Numbers ESSA Level III Study (2021 - 2022)

Prepared for:
Happy Numbers

Prepared by LearnPlatform by Instructure:
Alexandra Lee, Ph.D., Researcher, LearnPlatform
Chelsae Long, M.A., Jr. Researcher, LearnPlatform
Meetal Shah, Ph.D., Senior Researcher, LearnPlatform

May 15, 2023

EXECUTIVE SUMMARY

Happy Numbers contracted with LearnPlatform by Instructure, a third-party edtech research company, to examine the relationship between student usage of the Happy Numbers math program and math achievement. LearnPlatform designed the study to satisfy Level III requirements (*Promising Evidence*) according to Every Student Succeeds Act (ESSA)[1].

Study Sample, Measures, and Methods

This study utilized data from the 2021-22 school year and analyses included 540 elementary school students (Kindergarten and Grade 1) across 30 schools in one school district.

Researchers used measures of usage and achievement to provide insights into how schools implement Happy Numbers and the potential impacts on student achievement. Happy Numbers provided LearnPlatform with usage metrics and scores from their placement test and end-of-year math assessments[2]. The school district provided student demographic data, which researchers used to evaluate whether there were differences in usage and achievement depending on students' subgroup membership.

Researchers used a variety of quantitative analytic approaches. Specifically, descriptive statistics were used to examine participant characteristics and support analyses of implementation, as well as multilevel regression analyses to investigate how use of Happy Numbers is related to students' achievement on end-of-year math assessments. The analyses included student-level covariates to control for potential selection bias. Classroom-level effects were accounted for by the multilevel models. In addition, researchers calculated standardized effect size estimates to determine the magnitude of changes in math achievement at the beginning- and end-of-year.

[1] To show promising evidence, an ESSA Tier 3 study must be a well-designed and implemented correlation study with statistical controls for selection bias, have at least one statistically significant positive result on a relevant outcome, and no strong negative findings from other quasi-experimental or experimental studies.
[2] The assessments used are aligned with the Quantile® ranges developed by MetaMetrics®.

Key Takeaways

- Kindergarten and Grade 1 students that completed more tasks in Happy Numbers had higher end-of-year math scores (statistically significant relationships $p < .05$).
- On average, Kindergarten and Grade 1 students completed a total of 168 tasks over the course of the school year. There were no significant differences in usage depending on demographic subgroup membership (i.e., gender, special education status, or race/ethnicity).
- Kindergarten and Grade 1 students who completed at least 132 Happy Numbers tasks had significantly higher end-of-year math assessment scores compared to students who completed less than 132 tasks per year[3].

Conclusions

Given positive outcome findings, this study provides results to satisfy ESSA evidence requirements for Level III (*Promising Evidence*). Specifically, this study met the following criteria for Level III:

- ✓ Correlative study with non-independent outcome measure
- ✓ Proper design and implementation
- ✓ Statistical controls through covariates
- ✓ At least one statistically significant, positive correlation with statistical controls for selection bias

[3] An ANOVA test to check baseline equivalence indicated that there were no significant differences depending on cluster membership (F (512, 2) = 0.14, p = .872).

TABLE OF CONTENTS

Introduction

Happy Numbers contracted LearnPlatform by Instructure, a third-party edtech research company, to examine the association between student usage of its math program and achievement. LearnPlatform designed the study to satisfy Level III requirements (*Promising Evidence*) according to the Every Student Succeeds Act (ESSA).

Happy Numbers recognizes that half of U.S. high school students are not ready for college-level mathematics at time of enrollment, and the trajectory of college students who lag their peers in mathematics begins as early as elementary school. Happy Numbers was designed to assist teachers with math instruction through differentiated, individualized math instruction (see logic model in Appendix A; Henschel & Shah, 2022).

The present study had the following research questions:

Implementation Question

1. During the 2021-22 school year, what is the average number of tasks completed by students on Happy Numbers?

Outcome Question

2. Do students who complete more tasks on Happy Numbers have higher math achievement?

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

Study Design and Methods

This section of the report briefly describes the study's design, setting, participants, measures, and analysis methods.

Study Design

This study used a correlative design to examine the relation between math achievement with Happy Numbers usage during the 2021–2022 school year. This analysis controlled for pre-test math scores, demographic variables that were significantly associated with post-test math scores (i.e., special education status), and classroom-level effects.

Setting

This study included data from the 2021–22 school year and included 540 Kindergarten and Grade 1 students enrolled in math courses across 30 elementary schools in a mid-sized midwestern school district.

Participants

After removing extreme outliers (i.e., students that completed more than 400 tasks in a school year and were greater than 2.5 standard deviation units about the mean), there were 512 students in the final analytic sample. According to demographic data provided by the district, 51% of students were enrolled in Kindergarten, and 49% were enrolled in Grade 1. The racial breakdown of students in the sample was as follows: White (84.96%), Hispanic or Latino (10.16%), African American (2.93%), followed by Asian (1.95%). Females made up 49.2% of the group, while males accounted for 50.8%. In addition, 3.13% of the Kindergarten and Grade 1 students were identified as participating in special education (i.e., they have an individualized education plan (IEP)).

Measures

This study included multiple measures to provide insights into Happy Numbers implementation and evidence about the potential impacts of the learning solution on math achievement.

Happy Numbers Assessments. The Happy Numbers assessment[4] is aligned to the Quantile® Framework for mathematics and is designed to assess students' mathematics readiness in Kindergarten through Grade 5. The beginning-of-year (i.e., fall 2021), mid-year (i.e., winter 2022), and end-of-year assessments (i.e., spring 2022) are adaptive, consist of 20 multiple-choice items

[4] Happy Numbers refers to the beginning-of-year administration of this assessment as the "placement test".

for each grade level, and cover the skills and concepts associated with the Happy Numbers curriculum and increase in complexity according to students' grade level (i.e., Grade 1 students answer more complex questions than Kindergarteners).

Happy Numbers Usage Metrics. Researchers utilized 2021–22 student-level usage (i.e., total tasks completed in Happy Numbers). Usage data informed the extent to which students used Happy Numbers during the school year and whether students' use of the product related to students' math achievement. Happy Numbers provides recommendations about the number of tasks a student should complete (i.e., a task goal) based on their initial placement level determined by the Happy Numbers Assessments. The range for task goals varies by grade and proficiency on the placement test. For Kindergarten students, the task goals are 225 (above grade), 250 (on grade), and 330 (below grade). For Grade 1 students the task goals are 220 (above grade), 260 (on grade), and 330 (below grade).

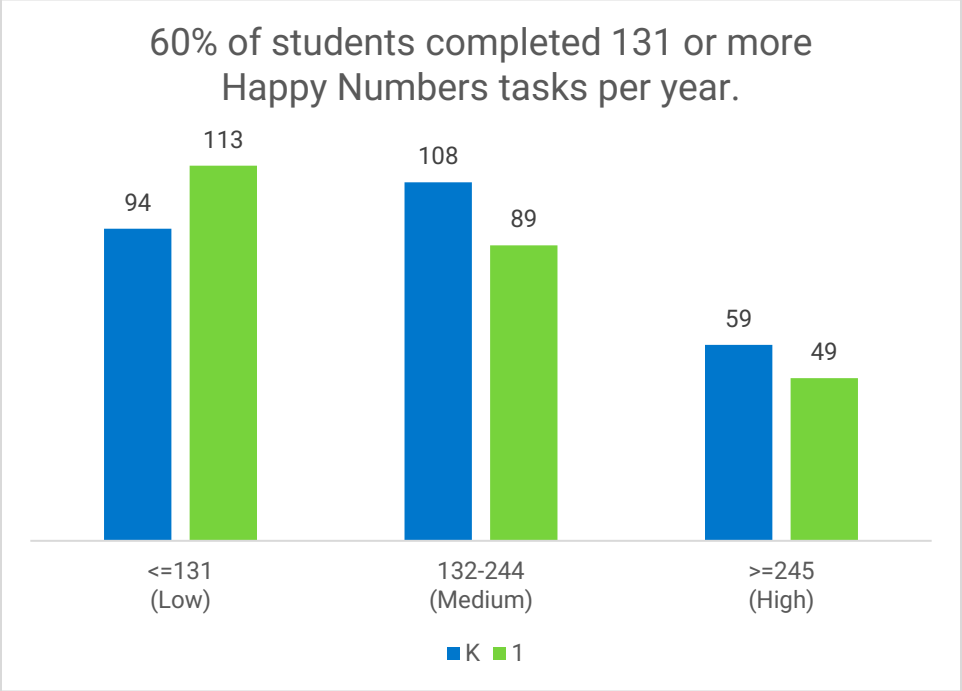
Data Analysis

Researchers used a variety of quantitative analytic approaches. First, researchers conducted descriptive statistics to describe participant characteristics and support implementation analyses. Researchers conducted a *k*-means cluster analysis to group students by similar levels of Happy Numbers usage based on the total number of tasks completed. Clusters were identified at the grade level (i.e., Kindergarten and Grade 1) to evaluate whether there were differences in usage and achievement depending on grade level.

Researchers then conducted multilevel regression analyses to examine how Happy Numbers use related to student math achievement on the Happy Numbers Assessment in one school year. The multilevel regression analyses included student-level demographic variables and beginning-of-year math scores to control for potential selection bias. In addition, researchers calculated standardized effect sizes to determine the magnitude of changes in student achievement or the standardized difference between students depending on usage group membership (i.e., difference in performance between Happy Numbers low, moderate, and high users). In addition to using multilevel regression models, ANCOVAs were also used to assess whether usage cluster membership was significantly associated with differences in end-of-year math scores, controlling for beginning-of-year math scores.

Program Implementation

The charts below highlight Happy Numbers use during the 2021-22 school year based on internal usage data. Overall, Kindergarten and Grade 1 students completed an average of 168 tasks on Happy Numbers across a school year (SD = 91.15, range 9 – 398).



Study Findings

To answer the study research questions, researchers conducted regressions analysis. All findings are statistically significant at the $p < .05$ and include standardized effect sizes (Hedge's *g*) to assist with interpretation.

? Do students who complete more tasks on Happy Numbers have higher math achievement?

Researchers examined whether there was a significant relationship between completing Happy Numbers tasks and math achievement during the 2021–22 school year among Kindergarten and Grade 1 students. Of note, Appendix B provides additional information on these analyses and findings.

To determine whether there was a significant relationship between Happy Numbers use and math achievement, researchers conducted multilevel regression analysis that included end-of-year math assessment scores as the outcome of interest. These models included special

education status as a student-level covariate since it was identified in preliminary analysis to be significantly associated with end-of-year test scores.

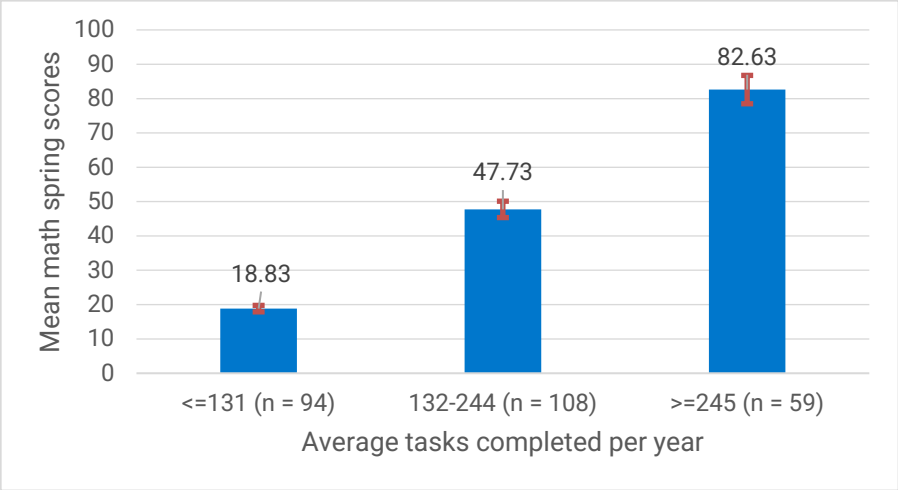
Overall Relationship Between Tasks Completed in Happy Numbers and Math Achievement

The multilevel regression model included end-of-year math achievement scores as the outcome of interest, controlling for classroom-level random effects, beginning-of-year math scores, and special education status. The results show that the number of tasks completed in Happy Numbers had a statistically significant positive relationship with end-of-year math scores across the full analytic sample (b = 330, p < .001). The partial eta-squared effect size[5] indicated there was a large positive effect (η^2 = .29).

These findings were further confirmed by a second regression analysis that specified usage cluster membership as a predictor variable of end-of-year math scores (controlling for beginning-of-year scores and special education status). We found that students in the clusters with greater usage had higher math scores, even though there were no significant differences in beginning-of-year test scores across the clusters (F (512, 2) = 0.14, p = .872). Specifically, students in the high total tasks cluster had statistically significant higher end-of-year math scores compared to the low total tasks cluster (b = 67.940, p <.001). Students in the high usage tasks clusters were using Happy Numbers nearly aligned to the recommended number of tasks for students who placed on grade level on the baseline test (i.e., 250 total tasks for Kindergarteners and 260 total tasks for Grade 1 students). See the graphs below for detailed results.

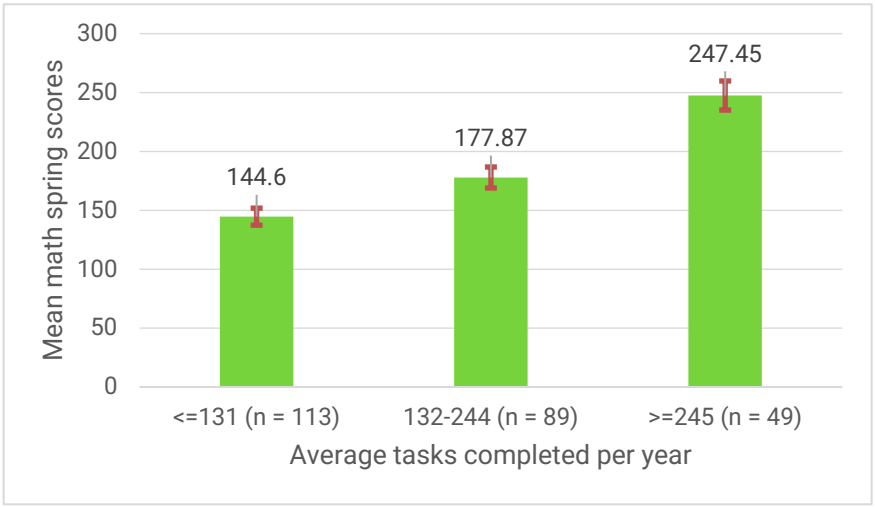
[5] The effect sizes reported are η^2 for the results of the multiple regression tests (i.e., variance explained). The magnitude of η^2 can be interpreted using the following thresholds: small effect = .01, medium effect = .06, and large effect = .14.

Key Finding 1. Kindergarten students that completed more Happy Numbers tasks had higher scores on the end-of-year math assessment than students who completed fewer tasks (low vs. medium Hedge's *g* effect size[6] = 0.21; medium vs. high effect size =0.25, low vs. high effect size = 0.45).



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

Key Finding 2. Grade 1 students that completed more Happy Numbers tasks had higher scores on the end-of-year math assessment than students who completed fewer tasks (low vs. medium Hedge's g effect size = 0.18; medium vs. high effect size =0.42, low vs. high effect size = 0.56).



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

[6] The effect sizes reported are Hedge's *g* for the results of the adjusted mean comparison tests (i.e., mean difference). The magnitude of Hedge's *g* can be interpreted using the following thresholds: small effect = .20, medium effect = .50, and large effect = .80.

Conclusions & Recommendations

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

Researchers recommend the following next steps:

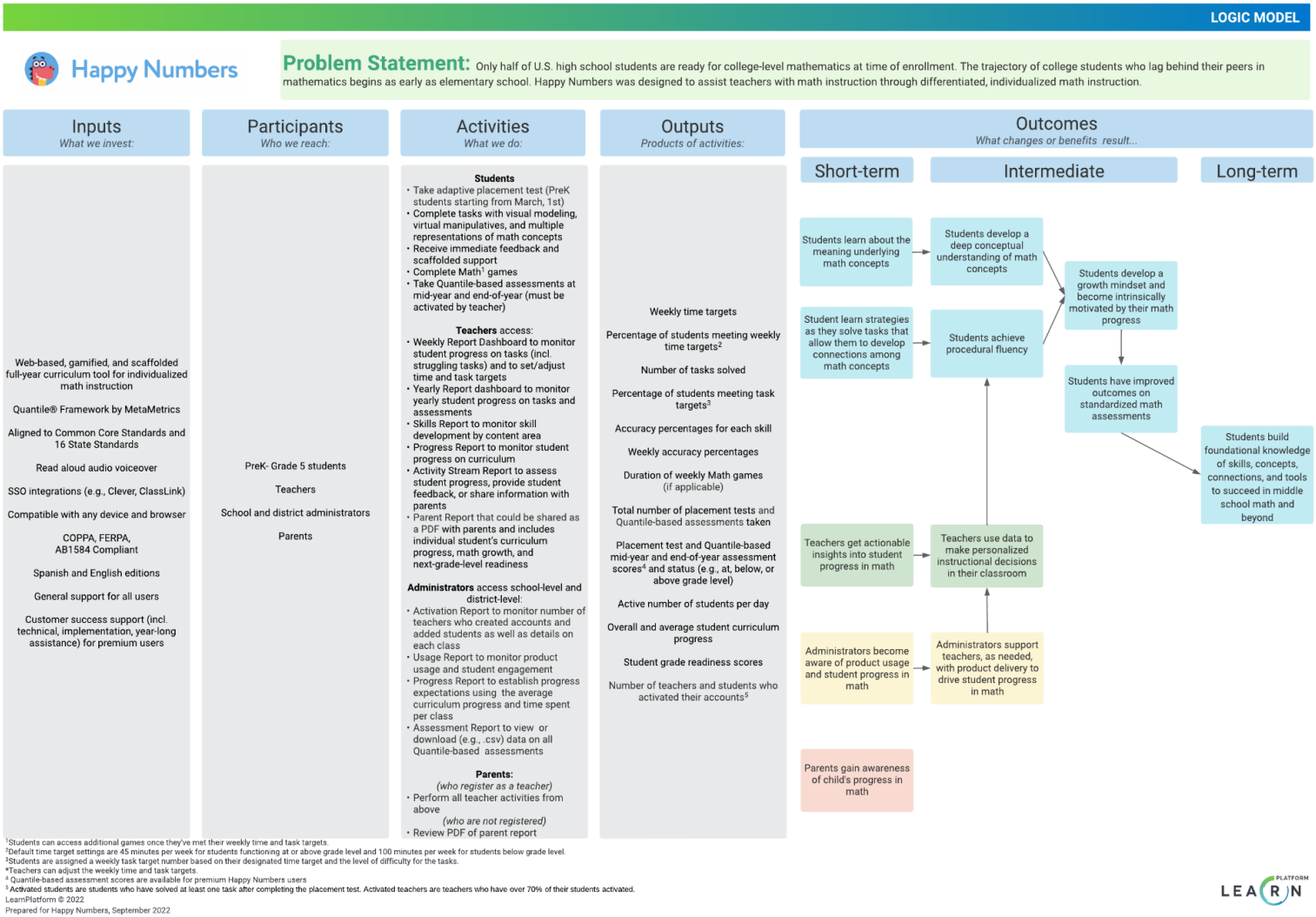
- Evaluation of Happy Numbers in another school district will allow stakeholders to better understand Happy Numbers' impact on math achievement in a different learning context.
- There was a statistically significant, positive relationship between using Happy Numbers (completing > 131 tasks) and math achievement. Furthermore, students who completed the number of tasks aligned with Happy Numbers recommendations for students on grade level (i.e., > 244 tasks) had the largest effect size compared to those who used Happy Numbers the least. This finding suggests that students should aim to reach the task benchmarks detailed by Happy Numbers.
- Happy Numbers should consider recruiting districts that use standardized assessments that are not specifically aligned with the product and meet WWC's standards for assessment outcomes (ESSA Level II).

References

Henschel, M. & Shah, M. 2022. *Happy Numbers Logic Model. Study Type: ESSA Evidence Level IV*. LearnPlatform.

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. Happy Numbers Logic Model



Appendix B. Additional Information on Student Findings

The following section shows additional information regarding the study's findings. Researchers report statistically significant findings at the $p = .05$ level and calculated standardized effect sizes.

Relationship between Average Tasks Per Year and Math Scores for Kindergarten and Grade 1 Students

To examine the relationship of Kindergarten and Grade 1 students' average tasks completed per year on Happy Numbers with math scores, researchers ran two types of regression analyses to examine whether Happy Numbers usage was significantly associated with end-of-year math scores. First, we used multilevel regression models to assess whether total tasks completed on Happy Numbers significantly predicted end-of-year math scores, controlling for classroom-level random effects, beginning-of-year math scores, and special education (SPED) status. The overall model fit was statistically significant (Wald $\chi^2 = 209.69$, $p < .001$) and the detailed results by parameter are included in Table B1.

Table B1. *Math scores by total tasks per year on the Happy Numbers platform across both Kindergarten and Grade 1 students*

Predictor	***b* Coefficient**	**Standard Error**	***z*-value**	***p*-value**
Total tasks	.326	.077	4.22	<.001
Fall math score	.595	.047	12.60	<.001
SPED status	-89.49	34.10	-2.62	<.001
Intercept	71.32	18.95	3.76	<.001
Classroom-level random effects	**Estimate**	**Standard Error**	**95% Confidence Interval**	
Intercept	4732.62	1501.21	2541.54	8812.66
Residual	15874.17	1022.84	13990.86	18010.99

The second type of regression model assessed whether average tasks usage groups were significantly associated with end-of-year math scores, after controlling for beginning-of-year scores, and SPED status. To account for all possible combinations among usage groups, researchers ran two separate models. In the first model, the low-usage group (i.e., <=131 average tasks per year) served as the reference group and was compared with the middle (i.e., 132–244 average tasks per year) and high-usage groups (i.e., >=245 tasks per year). In the second model, the medium-usage group served as the reference group and was compared with the two other categories. To determine the magnitude of the relationship, researchers calculated standardized Hedge's *g* effect sizes (see table B2).

Table B2. *Math scores by average tasks per year on the Happy Numbers platform across both Kindergarten and Grade 1 students*

Group Comparisons	*b* Coefficient	Standard Error	*t*-value	*p*-value	Effect Size
<=131 average tasks per year compared to 132–244 tasks per year	-15.25	14.33	-1.06	0.288	.10
<=131 average tasks per year compared to >=245 tasks per year	67.94	17.11	3.97	<0.001*	.38*
132-244 tasks per year compared to >=245 tasks per year	52.69	17.25	3.05	0.002*	.31*

*Statistically significant at the 0.05 level. Effect sizes are Hedge's g.

To evaluate the practical implications of these regression analyses, we examined the estimated end-of-year mean scores by grade and cluster. On average, a Kindergarten student in the lowest usage cluster had an average spring math score of 19, whereas a student in the highest usage cluster had an average spring math score of 83. Both sets of students fell within the 50th to 90th percentiles according to the Quantile® student assessment equivalent scores[7] (for Kindergarten, the 50th percentile = EM115Q, and the 90th percentile = 190Q). On average, a Grade 1 student in the lowest usage cluster had an average spring math score of 145 and a student in the higher usage cluster had an average spring math score of 247. Similarly, to Kindergarten students, both sets of students fell within the 50th to 90th percentiles according to the Quantile student measures (for Grade 1, the 50th percentile = 120Q, and the 90th percentile = 390Q).

[7] These determinations were made using the Quantile Measure recommendations (https://hub.lexile.com/quantile-grade-level-charts).