Happy Numbers

The Impact of Happy Numbers Personalized Instruction on Student Learning Gains: Evidence from Union School District

- **District:** Union School District
- **Evaluation School Year:** 2024–2025
- **Product:** Happy Numbers
- **Grades:** K–5

Research Overview

This study examined fall-to-spring score gains on the Happy Numbers Diagnostic in a single school district, comparing two usage groups. Students who maintained the recommended level of program usage throughout the year achieved greater gains than those in the lower-usage group. This positive difference was statistically significant in the grades where students met the recommended usage level.

Introduction

Happy Numbers is an online supplemental educational tool designed to enhance mathematics instruction. It provides individualized instruction, teacher-assigned practice, printable materials, and additional educational resources for students in grades PK through 5.

Upon first logging in, students in grades K through 5 complete a diagnostic assessment[1] (referred to as the Placement Test), which determines their initial grade level and starting point within the Happy Numbers curriculum. This starting point is tailored to address each student's most significant areas of need. From there, students progress through the curriculum, receiving immediate feedback and tailored scaffolding to support their performance at every step. A new task is offered only when the previous one is completed with understanding.

Adjustments to the learning path may be made based on the Mid-Year Test results (not mandatory). In the spring, the End-of-Year (EOY) assessment (not mandatory) can be administered to measure students' growth and evaluate their readiness for the next grade level.

---

[1] The Happy Numbers assessments are powered by The Quantile® Framework for Mathematics

Study Purpose

This study examined how two levels of Happy Numbers usage related to students' fall-to-spring growth on the Happy Numbers Diagnostic within a single school district.

Research Questions

- **Usage metric**: How do students in the higher-usage group differ from those in the lower-usage group in terms of tasks completed, curriculum progress, and average time spent per week on Happy Numbers during the 2024-2025 school year?
- **Growth metric**: Do students in the higher-usage group achieve greater fall-to-spring gains on the Happy Numbers Diagnostic than students in the lower-usage group?

The difference between Placement Test and EOY Test scores in HN Score was selected as the metric for growth. For reference, 100 points of growth in HN Score is equivalent to 1 grade level of progress. To learn more about the HN Score scale, please follow this link.

Methodology

Initial Data Sample

The overall analytical sample for grades K through 5 included 1,715 students from 6 schools in the Union School District, located in a suburban area. All students in the sample completed both the Placement Test and the EOY Test and began practicing on Happy Numbers between August 1, 2024, and November 1, 2024. Table 1 displays the number of students in the sample by grade.

Grade	Number of Students
Kindergarten	246
Grade 1	368
Grade 2	370
Grade 3	353
Grade 4	262
Grade 5	116
Total	**1,715**

Table 1. Number of students in the sample by grade.

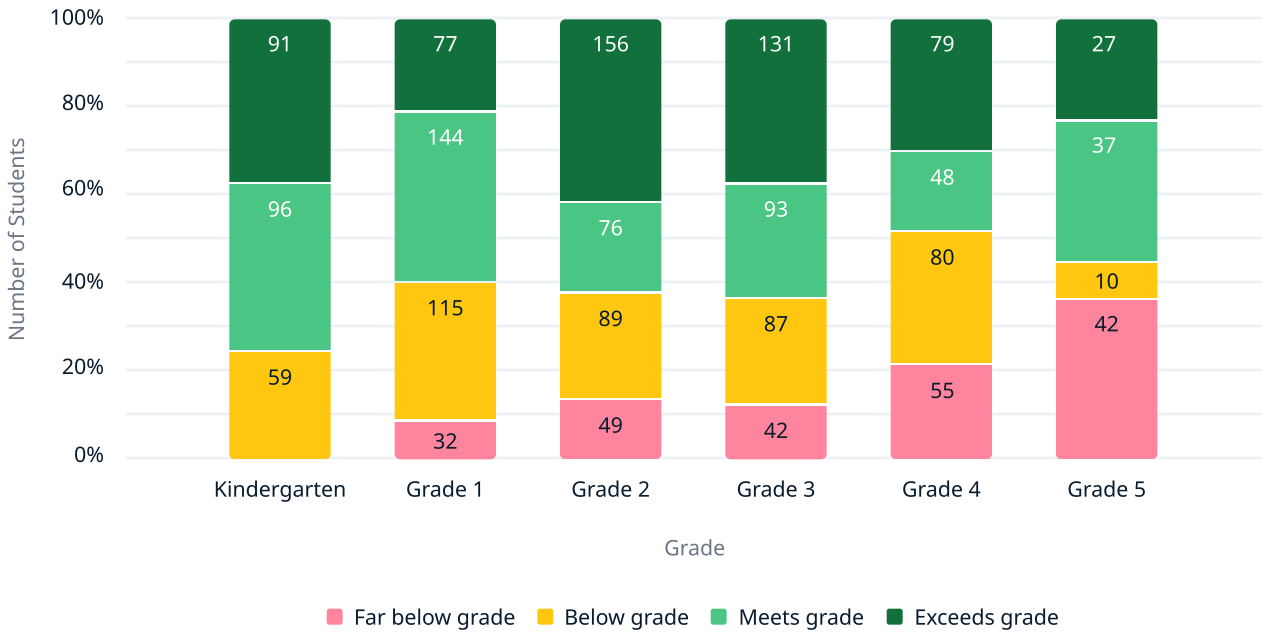
Students are categorized into one of four initial achievement levels: far below grade level, below grade level, met grade level, or exceeded grade level. This grouping was based on the placement map provided here. Specifically, these readiness levels are defined as follows:

- **Far below grade level**: More than one grade level behind
- **Below grade level**: Up to one grade level behind
- **Met grade level**: Meeting grade-level expectations
- **Exceeded grade level**: 0.5 or more grade levels ahead

Methodology

The chart below displays each grade's distribution of initial achievement levels at the start of the year. Across Grades K-5, between 48% and 76% of students began in the met or exceeded grade level categories, while 24%-52% fell into the below or far below grade level categories. The numbers on the bars represent the number of students within each category. A detailed breakdown by grade and initial achievement level can be found in Appendix I.

Number of Students by Grade and Initial Grade Level

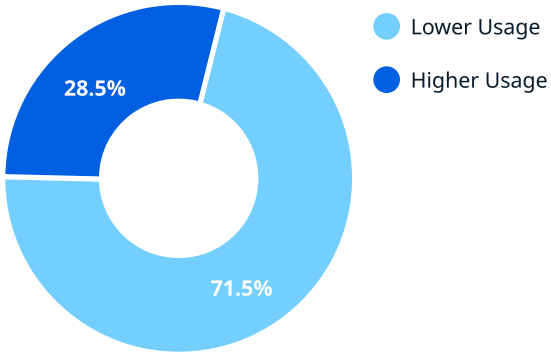


Methodology

Design

Because usage patterns and goals may vary within each grade, each one was analyzed separately. Further analysis included categorizing students into two usage groups - lower and higher. The number of tasks completed on Happy Numbers was selected as the metric for usage, and the k-means algorithm was applied. In addition to task completion, other metrics such as curriculum progress, average time spent on Happy Numbers per week, and their relationship to fall-to-spring math growth were analyzed.

Usage Patterns



Usage Group Size

For each grade, k-means clustering was conducted based on the number of tasks solved. This approach resulted in two clusters: **lower** usage and **higher** usage.

While the distribution varies across individual grades, the overall sample reflects the general usage trends.

*Notably, 28.5% of students across K-5 were classified into the **higher**-usage group.*

Methodology

The mean curriculum progress (measured as the equivalent of grades of the Happy Numbers curriculum completed by students) and the mean time students spent on Happy Numbers per week were calculated for each cluster. These metrics provide an easier-to-interpret framework for understanding student engagement. Table 2 below shows the average time per usage group, as well as the average curriculum progress for each usage group, expressed in grade-level equivalents.

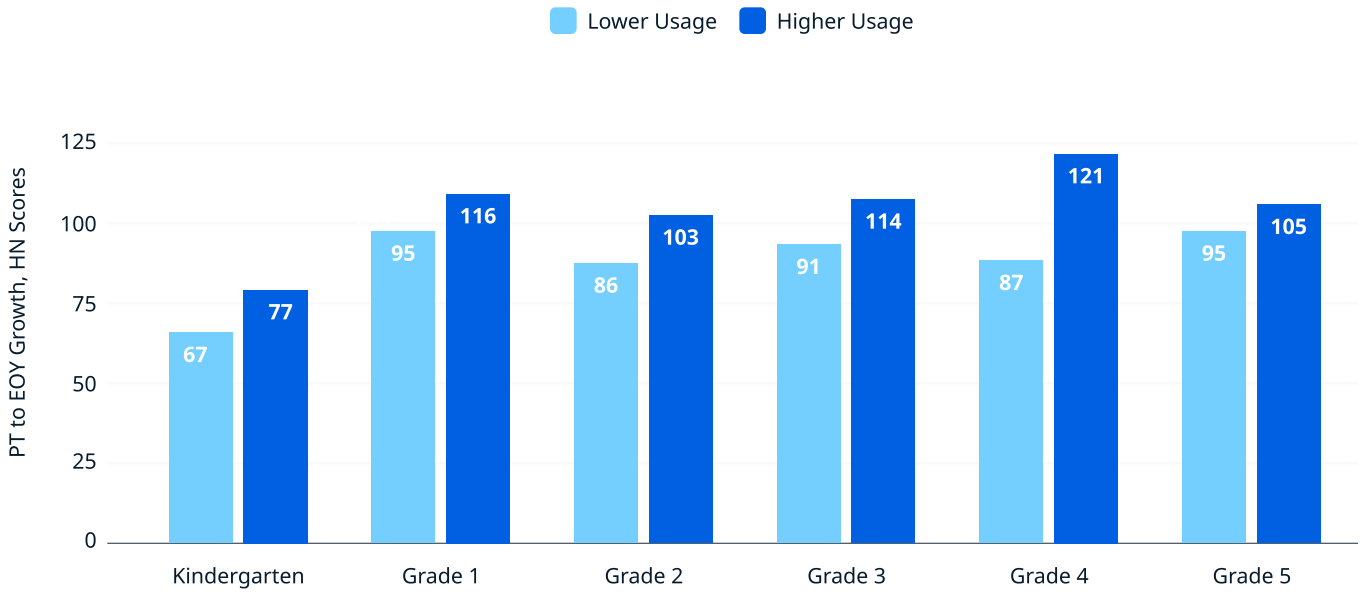
Grade	Usage Cluster	Number of Students	Avg. Curr. Progress	Avg. Weekly Time (min)
Kindergarten	Lower	158	36%	9
	Higher	88	90%	20
Grade 1	Lower	268	46%	15
	Higher	100	147%	**44***
Grade 2	Lower	295	35%	16
	Higher	75	136%	**52***
Grade 3	Lower	215	32%	16
	Higher	138	96%	**44***
Grade 4	Lower	210	30%	15
	Higher	52	97%	36
Grade 5	Lower	81	13%	9
	Higher	35	51%	25

Table 2. Grade-Level Usage Clusters: Size, Curriculum Progress, and Weekly Time

***** Asterisks denote meeting Happy Numbers' weekly usage guideline *(30 minutes for Pre-K, 35 minutes for Kindergarten, and 45 minutes for Grades 1–5).*

Kindergarten students in the high-usage group averaged 20 minutes per week on Happy Numbers, falling short of the recommended 35 minutes. In **Grades 1 through 3**, students classified in the high-usage group consistently met the platform's recommended weekly usage targets. In **Grade 4**, high-usage students came close to meeting the recommended time. In **Grade 5**, students in the high-usage group still fell short of the platform's recommended weekly usage target.

Results



Average Growth from Placement to End-of-Year by Usage Group

The graph illustrates a consistent trend. In **Grades 1-4**, students with higher Happy Numbers usage outperformed their lower-usage peers by 17 to 34 points in BOY-to-EOY growth. These differences are **statistically significant** ($p = .001–.022$) and correspond to effect sizes (Cohen's $d = 0.27–0.39$).

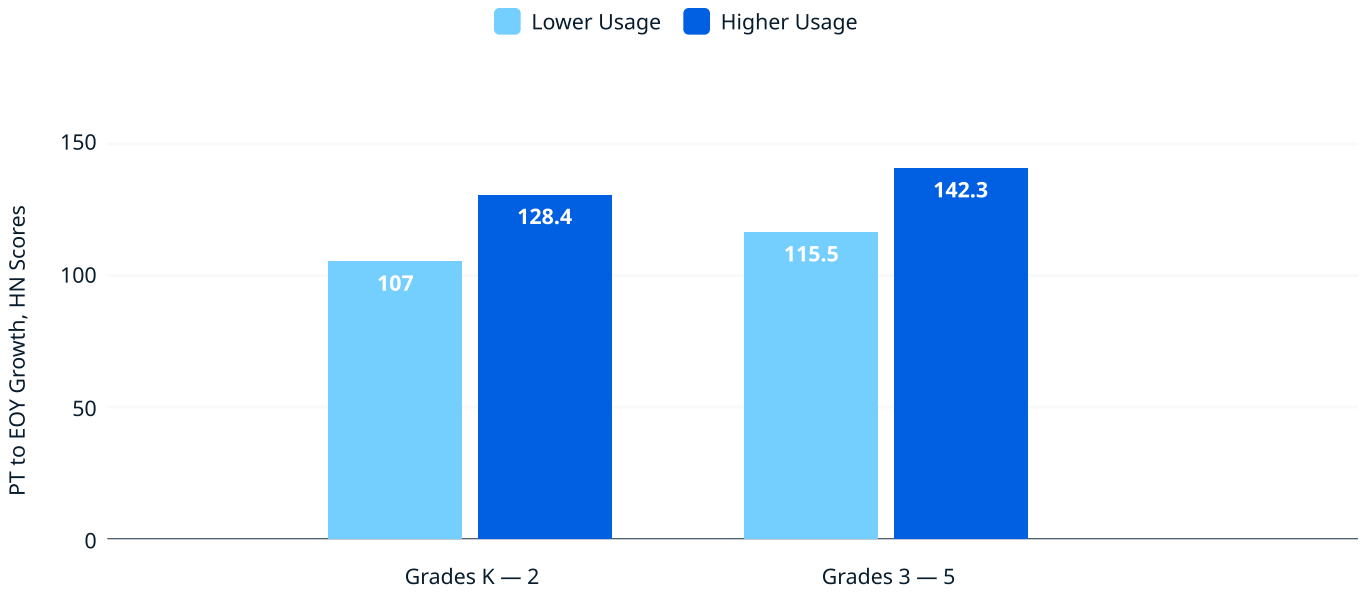
Kindergarten and **Grade 5** both had a ten-point increase, but did not show statistically reliable advantages, which likely reflects insufficient usage in those grades. For more details on statistical significance, refer to Appendix II.

Results

Effect of Usage for Below-Grade-Level Learners

To determine whether heavier engagement benefits the students who start below or far below grade level, we combined grades K-2 and Grades 3-5.

In the **K-2 pool**, higher-usage students gained an average of 21 points more than their lower-usage peers – a **statistically significant difference** ($t \approx 2.42$, $p = .016$, $d = .33$). In **Grades 3-5**, the advantage widened to almost 27 points ($t \approx 2.45$, $p = .015$, $d = .31$). A growth of over 140 HN Score points translates to **1.4 grade-level** equivalents, or almost a year and a half of extra progress.



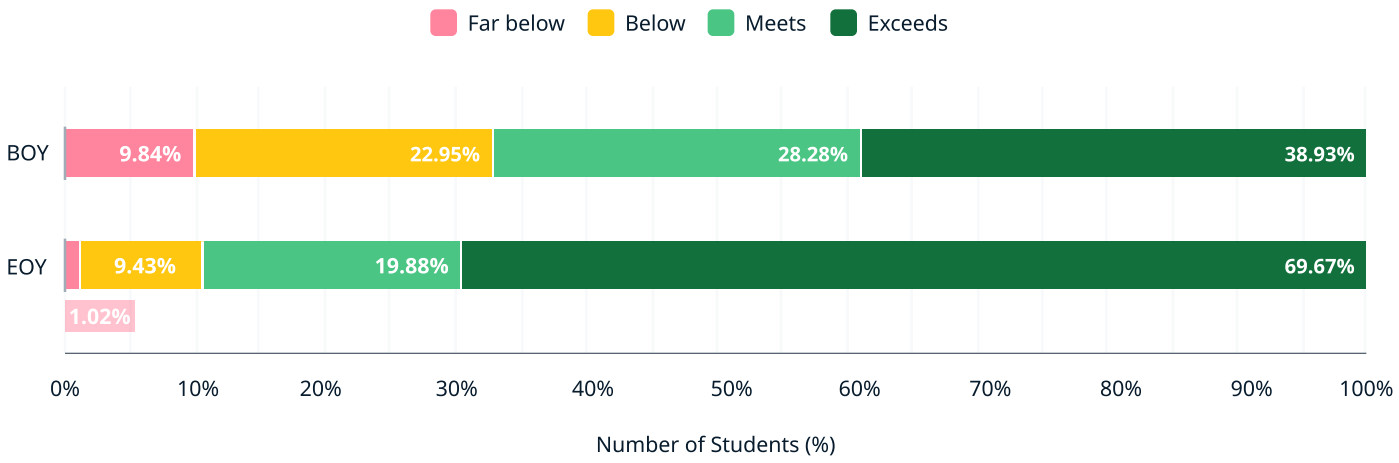
Average Growth from Placement to End-of-Year by Usage Group within 'Below' and 'Far below' Initial Groups

These differences indicate that for students who begin the year behind, sustained use of Happy Numbers is associated with markedly greater math growth over the course of the year. For more details on statistical significance, refer to Appendix III.

Results

Grade level transitions

To illustrate the magnitude of fall-to-spring gains, the figure below reproduces the Happy Numbers assessment report for students in the higher usage group.



K-5 Student Performance Growth: Beginning to End of Year (Higher Usage Group)

Note that the “Exceeds grade” threshold is deliberately stricter at the end of the year (one full grade level ahead) than at the beginning (half a grade level). Consequently, students had to clear a higher bar to remain in that category at EOY. The rest of the categories remain the same.

At BOY, 39% of recommended-usage students already exceeded grade level, and 29% started on grade level—adding up to 67% at or above grade level—while 32% were below or far below grade level. By EOY, the proficient share (meets + exceeds) had climbed from 67% to 90%, while the below/far below group shrank from 33% to 11% (far below ≈ 1%). This shift is highly significant ($p < .0001$): 115 of the 160 students who began below benchmark - about 72% - moved up to meeting or exceeding the grade level, whereas only six students slipped the other way. For more details, refer to Appendix IV.

Conclusion

Across one full school year, students who consistently met Happy Numbers' recommended usage targets achieved measurably greater growth than their lower-usage peers. In Grades 1-4, the advantage ranged from 17 to 34 points - a statistically and educationally meaningful difference ($p < .05$; Cohen's $d \approx 0.3$). For learners who began **below** or **far below** grade level, heavier engagement raised BOY to EOY gains by 21 points in K-2 and 27 points in Grades 3-5, again with significant t-test results. Overall proficiency rates tell the same story: the share of higher-usage students meeting or exceeding grade-level expectations climbed from 67% in the fall to 90% in the spring, a shift confirmed by a highly significant McNemar test ($p < .0001$).

The two grades that did **not** reach significance - Kindergarten and Grade 5 - were also those where even the "high-usage" cluster fell short of the recommended guideline, underscoring the dosage effect. While these findings are observational and cannot prove causation, they provide strong evidence that sustained, guideline-level use of Happy Numbers is linked to accelerated math growth, especially for students who start the year behind. Future work should explore causal designs, longer follow-ups, and additional covariates (e.g., teacher implementation fidelity) to refine these insights.

Appendix I

The table shows the distribution of students across proficiency categories by grade level, based on their initial placement test results. The data includes the number of students (n) and the percentage (%) for each performance category within each grade.

	Far below grade level		Below grade level		Met grade level		Exceeded grade level	
Grade	Students (n)	%	Students (n)	%	Students (n)	%	Students (n)	%
Kindergarten	–	–	59	24.0	96	39.0	91	37.0
Grade 1	32	8.7	115	31.3	144	39.1	77	20.9
Grade 2	49	13.2	89	24.1	76	20.5	156	42.2
Grade 3	42	11.9	87	24.6	93	26.3	131	37.1
Grade 4	55	21.0	70	30.5	48	18.3	79	30.2
Grade 5	42	36.2	10	8.6	37	31.9	27	23.3

Appendix II

Grade	n (High)	n (Low)	Mean Growth (High)	Mean Growth (Low)	Δ (High – Low)	t-stat	p-value	Cohen's d
K	88	158	76.7	67.28	+9.43	1.65	0.1	0.19
1	100	268	116.15	94.65	+21.5	3.1	**0.002**	0.35
2	75	295	103.33	86.08	+17.25	2.32	**0.022**	0.27
3	138	215	113.88	90.65	+23.23	3.34	**0.001**	0.34
4	52	210	121.44	87.07	+34.37	2.56	**0.012**	0.39
5	35	81	105.29	95.06	+10.22	0.74	0.463	0.15

Welch's t-tests Comparing BOY to EOY Growth Between Usage Clusters

Because the two usage groups (lower and higher) within a grade had unequal sample sizes and could differ in score variability, the Welch two-sample t-test was applied. For every grade, we obtained a t-statistic, a p-value, and an effect-size estimate (Cohen's d) that expresses the magnitude of the difference in standard-deviation units.

The findings show a clear pattern. In Grades 1 through 4, the mean growth for higher-usage students exceeded that of lower-usage peers by roughly 17 to 34 points, and the p-values - ranging from about .001 to .022 - indicate these gaps are statistically significant at the 5 percent level. The corresponding Cohen's d values fall between 0.27 and 0.39. In Kindergarten, the higher-usage group gained about nine points more, but the p-value of .10 suggests this result is not statistically significant. In Grade 5, the ten-point advantage for high users was also not significant ($p \approx .46$).

Appendix III

Appendix III reports the statistics for students initially classified as **below** or **far below** grade level. For each grade band, the table lists sample sizes, mean growth by usage group, the mean difference (Δ), the two-tailed p-value from the Welch t-test, and Cohen's d effect size.

Grade	n (High)	n (Low)	Mean Growth (High)	Mean Growth (Low)	Δ (High – Low)	t-stat	p-value	Cohen's d
K — 2	73	271	128.4	107	21.3	2.42	**0.016**	0.33
3 — 5	87	229	142.3	115.5	26.8	2.45	**0.015**	0.31

Appendix IV

The table below cross-classifies the 488 higher-usage students by proficiency status at the beginning (BOY) and end (EOY) of the year. For this analysis, **Proficient** includes students in the **met** and **exceeded grade level** categories, while **Not Proficient** includes those **below** or **far below grade level**. The McNemar test focuses on the 121 "discordant" students whose status changed: 115 moved up to Proficient, while only 6 moved down. The resulting $\chi^2 \approx 98.2$ (1 df) yields $p < 0.0001$, confirming that the shift toward proficiency is far too large to be attributed to chance.

	EOY Proficient	EOY Not Proficient	Row Total
BOY Proficient	322	6	328
BOY Not Proficient	115	45	160
Column Total	437	51	488

Discordant pairs: **b = 6** (Proficient → Not), **c = 115** (Not → Proficient)