![WestEd logo](WestEd.org)

Impact and Implementation Evaluation of Lalilo: Phase II Executive Summary

Aaron Soo Ping Chow, Katy Nilsen, Jazmin Cruz, Dennis Ciancio

September 2024

The Lalilo English literacy program was developed to enable students to learn and practice foundational pre-reading and reading skills in English. The Lalilo program is constructed with the potential to improve students' emergent and basic reading skills in content areas that include phonics, phonology, sight/high frequency words, word families, independent reading, listening comprehension, and grammar and conventions.

In 2021, WestEd partnered with Renaissance Learning to examine the impact of Lalilo on literacy achievement and reading attitudes among first grade students, and to develop an understanding of the link between Lalilo implementation, contextual factors, and program impact. Phase I of the study, which occurred during the 2022-23 school year, focused on understanding how Lalilo is typically implemented in early elementary grades and teachers' impressions of the program.[1] Phase II included two studies, an impact study focused on estimating the impact of Lalilo on literacy achievement and student motivation and attitudes towards reading and an implementation study focused on teachers' impressions of the Lalilo student program and teacher dashboard, their reported implementation of Lalilo, and student performance within the Lalilo program.

This executive summary provides findings for the Phase II impact and implementation studies.

Impact Study

Fourteen first grade teachers and their students were recruited to be in the treatment group and use Lalilo during the 2023–24 school year, while another fifteen first grade teachers and their students were recruited to be in the comparison group and continue business-as-usual instruction.[2] The final

---

[1] Soo Ping Chow, A., Nilsen, K., Boxerman, J., Cruz, J., & Ciancio, D. (2023). *Impact and implementation evaluation of Lalilo: Phase I brief*. WestEd. https://docs.renaissance.com/R67209

[2] Participating teachers were from schools representing different geographic areas across the country; however, 23 of the 29 teachers came from midwestern states.

analytic sample for the impact study included data from 222 treatment group students and 198 comparison group students.[3]

Impact of Lalilo on Literacy Skills

Treatment group students scored significantly higher on the FastBridge earlyReading composite and subtest measures at pretest (i.e., prior to Lalilo implementation) than the comparison group, indicating a lack of baseline equivalency. Additionally, the treatment and comparison group samples differed in their demographic composition, with more comparison group students coming from historically underserved populations (i.e., the comparison group had a higher percentage of minority students, low income students, English learners, and students with an IEP or Section 504 plan). To adjust for these baseline differences, the average weekly growth rates of the FastBridge measures of interest were used as the outcome measures for the hierarchical linear model (HLM) analysis, thereby controlling for the differences in performance at pretest and differences in the number of weeks between testing administrations across classrooms. Student- and classroom-level characteristics were also included to control for demographic differences between the treatment and comparison groups.

Overall, treatment group students grew significantly faster from winter to spring than the national median in terms of their composite and FastBridge CBMreading scores, while the comparison group growth rate was at or even slightly below the national median. However, differences between the treatment and comparison group students on these FastBridge measures were less conclusive when examining the hierarchical linear model results.

The HLM results revealed no observable statistically significant impact of Lalilo usage on the composite score. Although the treatment group average weekly growth rate was about 0.1 standard deviations higher than the comparison group rate, the effect was not statistically significant. However, a large positive impact was observed on the FastBridge CBMreading subtest, with treatment group students having an average weekly growth rate that was about 1.2 standard deviations higher than the comparison group. This meant that the average treatment group student had a higher CBMreading weekly growth rate than 88%[4] of the comparison group students, after controlling for baseline differences (WWC, 2022). The impact on CBMreading performance was also observed when comparing the bottom quartile of students in terms of their pretest CBMreading performance. The analysis revealed that the average low-performing treatment group student had a higher weekly growth rate than 65% of low-performing comparison group students (that is, an effect size of 0.4 standard deviations), suggesting that Lalilo has an impact even among students needing the most reading support.

Impact of Lalilo on Student Reading Attitudes

In terms of the impact of Lalilo on student attitudes towards reading, most teachers agreed that Lalilo improved their students' general motivation to read; however, treatment group students generally

---

[3] The analytic sample differs for each impact analysis as some students did not complete all measures during both the winter and spring administrations. The associated Phase II report provides the relevant sample sizes.

[4] The percentile corresponding to the effect size was estimated by converting the effect size to Cohen's U3 index, as described in the WWC Version 5.0 Procedures and Standards Handbook. The index is the fraction of comparison group students outperformed by the average intervention group student and is computed based on the proportion of the area under a standard normal curve that is below the value of the effect size.

self-reported the same attitudes towards reading at the beginning and end of the study. Nevertheless, the change in student attitudes toward reading did have a small significant positive effect on the FastBridge composite score rate of change.

Implementation Study

The fourteen recruited first grade teachers and their students also participated in the implementation study. The final analytic sample for this study included data from 228 treatment group students.[5]

Impressions of Lalilo Student Program and Teacher Dashboard

Most teachers reported that they were extremely likely to recommend Lalilo to other teachers. On a scale of 0 to 10, where 0 is "not at all likely," 5 is "neutral," and 10 is "extremely likely," Lalilo received an average rating of 8.7.

Teachers mostly expressed positive opinions about the student program and teacher dashboard. Eleven teachers agreed that the dashboard was easy to navigate and that the data it provided was useful for improving their reading instruction. Teachers also appreciated the instructional complexity and quality of the content in the Lalilo program, the quality of the graphics, the pacing of the program, and its overall quality, and reported that the Lalilo program improved their students' overall reading skills and early literacy/reading skills across specific ELA elements. They also saw value in terms of student engagement (e.g., in some classrooms, students wanted to use Lalilo as a free choice option), believed their students progressed through the activities in the program even when they faced difficulties, and thought that the program improved their students' engagement with text.

Implementation of Lalilo Student Program and Teacher Dashboard

Students used the Lalilo program for 3.5 days per week on average, with a median usage of 4 days, and an average of 40 minutes per week or 11 minutes per day. The program was primarily used a whole-class tool, and students most frequently worked on exercises within the Phonics, Sight/High Frequency Words, Word Families, and Language Study domains. Following implementation of the program, most teachers reported that 11–20 minutes per lesson was the ideal usage time.

Teachers logged into the dashboard site frequently (i.e., an average of 4.3 days per week), but they rarely viewed any of the student, lesson, or standards reports, if at all. For teachers who reported accessing the dashboard in their weekly logs, they typically used it to determine how much time an individual or the class spent on the program, to assign additional Lalilo activities, and to track class and student performance across different skills, specifically phonological and word work skills. Additionally, after implementing the program, half of teachers agreed that the Lalilo data from the dashboard influenced the amount of emphasis that they placed on specific ELA elements during instruction.

Contextual characteristics of the treatment group teachers and their experiences with the Lalilo program may have affected their program implementation and use of the teacher dashboard. On average, teachers who spent more time each week preparing for ELA instruction and less frequently used dashboards prior to the study, logged into the dashboard site more frequently each week.

---

[5] Some treatment group students were included in the implementation study but not in the impact study as they did not complete the relevant impact measures. Details about the study samples are included in the associated Phase II report.

Similarly, teachers who agreed that the dashboard was easy to navigate and that it provided data that was useful for improving reading instruction, and/or experienced no or very few issues with the Lalilo program, also logged into the dashboard site more frequently each week.

Performance Within Lalilo Student Program

Across the implementation period, treatment group students, on average, completed about 89% of Lalilo exercises correctly, with this rate differing by domain and student performance on the Lalilo placement test. The Sight/High Frequency Words domain had a 96% average correctness rate, while the Word Families and Phonics domains had an 85% correctness rate, and the Language Study domain had an 80% correctness rate. Additionally, students who started at higher levels in the Lalilo program (i.e., had better scores on the placement test) tended to have higher average correctness rates across the implementation of the program.

Furthermore, an HLM analysis revealed that average correctness in the first month of usage was a significant predictor of a student's FastBridge composite score growth rate, but not FastBridge CBMreading growth rate. That is, students with a higher initial average correctness improved more per week on the composite measure, on average, than students with a lower initial average correctness.

Impact and Implementation Evaluation of Lalilo

Phase II Report

Aaron Soo Ping Chow, Katy Nilsen, Jazmin Cruz, and Dennis Ciancio

September 2024

Suggested citation: Soo Ping Chow, A., Nilsen, K., Cruz, J., & Ciancio, D. (2024)*. Impact and implementation evaluation of Lalilo: Phase II report. WestEd.*

WestEd is a nonpartisan, nonprofit agency that conducts and applies research, develops evidence-based solutions, and provides services and resources in the realms of education, human development, and related fields, with the end goal of improving outcomes and ensuring equity for individuals from infancy through adulthood. For more information, visit WestEd.org. For regular updates on research, free resources, solutions, and job postings from WestEd, subscribe to the E-Bulletin, our semimonthly e-newsletter, at WestEd.org/subscribe.

Table of Contents

Introduction and Evaluation Overview

The Lalilo English literacy program was developed to enable students to learn and practice foundational pre-reading and reading skills in English. The Lalilo program is constructed with the potential to improve students' emergent and basic reading skills in content areas that include phonics, phonology, sight/high frequency words, word families, independent reading, listening comprehension, and grammar and conventions. The program therefore provides a diversity of exercises across the following five key components of reading (NICHD, 2000): phonological awareness, phonics, fluency, vocabulary, and comprehension. The program format follows a clear trajectory as students progress through the placement activity and learning exercises.

In 2021, WestEd was invited to partner with Renaissance Learning (RL) to examine the impact of Lalilo on academic achievement and reading attitudes, and to help RL develop an understanding of the link between Lalilo implementation, contextual factors, and program impact. Two phases were conducted for the evaluation: the first during the 2022–23 school year[1] and the second during the 2023–24 school year. In both phases, a quasi-experimental design (QED), with first grade students and their teachers, was used. The research design involved two conditions – treatment and comparison. Teachers using the Lalilo program with their students comprised the treatment group. Teachers who did not use Lalilo with their students comprised the comparison group. The comparison group was only available for the second phase of the evaluation.

In both phases, treatment teachers completed an intake survey prior to the study period to determine eligibility and gather initial contextual information. They also completed pre- and post-surveys on reading instruction and assessment, knowledge and resources to teach ELA, school-level interventions, and use of educational technology. Prior to implementation of Lalilo, treatment teachers received professional development (PD) where they viewed videos and resources asynchronously about using Lalilo and about enhancing their foundational literacy instruction, and completed a survey about these videos and resources. In addition, during the study, treatment teachers completed a weekly usage log. Post-implementation interviews with treatment teachers were conducted during Phase I of the evaluation but not during Phase II. Finally, comparison teachers only completed one survey at the beginning of the study to gather contextual information about their classrooms, resources, and instruction.

---

[1] Soo Ping Chow, A., Nilsen, K., Boxerman, J., Cruz, J., & Ciancio, D. (2023). *Impact and implementation evaluation of Lalilo: Phase I brief*. WestEd. https://docs.renaissance.com/R67209

The research questions developed collaboratively by WestEd and Renaissance Learning and addressed in this Phase II report included:

1. What are the effects of Lalilo on student literacy achievement?

2. What are the effects of Lalilo on student ELA motivation and engagement?

3. How was the Lalilo student program and teacher dashboard implemented?

4. How do students perform on the Lalilo in-program exercises and what is the association between Lalilo outcomes, program implementation, and FastBridge outcomes?

5. What are teachers' impressions of the Lalilo student platform, teacher dashboard, impact on student learning and motivation, program content and pacing, as well as teacher resources and professional development as applicable?

6. How do contextual factors affect the implementation and subsequent success of Lalilo?

Below, we present an overview of the study methodology, including sample demographics, participation rates, and study instruments. This is followed by a discussion of the findings on the implementation of the Lalilo program and impact of the program on student reading achievement and attitudes towards reading. Findings are presented for Phase II, although some comparisons are made with the Phase I results.

Method

This section of the report describes the treatment and comparison group samples, and the study instruments used for the evaluation to assess program implementation, impressions, and program impact.

Sample

Twenty-nine first grade teachers, comprising of 14 treatment teachers and 15 comparison teachers, participated in the study. They represented schools from different geographic areas across the country, with 23 of the 29 teachers coming from midwestern states (see Table 1). In these recruited classrooms, 229 students assented to participate in the study's treatment group and 214 students assented to participate in the study's comparison group.

Table 1. School Characteristics of the Treatment and Comparison Groups

Treatment			Comparison		
School Characteristics	# of Teachers	# of Students (%)	School Characteristics	# of Teachers	# of Students (%)
Midwestern Town Area (Public)	4	89 (39%)	Midwestern Suburban Area (Public)[2]	6	96 (45%)
Midwestern Suburban Area (Public)	3	62 (27%)	Midwestern Town Area (Public)	4	67 (31%)
Midwestern Town Area (Private Religious)	3	43 (19%)	Southeast City Area (Public)	2	27 (13%)
Western Rural Area (Public)	3	24 (10%)	Midwestern Urban Area (Charter)	2	18 (8%)
Southwestern Rural Area (Public)	1	11 (5%)	Midwestern Suburban Area (Specialized Public)	1	6 (3%)

The treatment and comparison group student samples differed demographically (see Table 2). The student population in both groups was roughly equally split in terms of gender, with 56% female students and 44% male students in the treatment group, and 48% female students and 52% male students in the comparison group. The racial/ethnic breakdown of the treatment group was relatively homogenous with about 94% White students, while the comparison group had more variation with 71% White students, 13% Hispanic/Latino students, and 9% Black students. Additionally, the comparison group had a higher percentage of low income students, English learners, and students with an individualized education plan (IEP) or Section 504 plan (39%, 3%, and 12%, respectively, in the treatment group, and 46%, 4%, and 27%, respectively, in the comparison group).

---

[2] Consists of schools from two districts.

Table 2. Student Demographic Data

Student Demographic Group	Treatment Group	Comparison
Female	126 (55.8%)	100 (47.6%)
Male	100 (44.3%)	110 (52.4%)
American Indian/Alaskan Native	1 (0.4%)	5 (2.4%)
Asian	1 (0.4%)	4 (1.9%)
Black/African American	0 (0%)	19 (9.1%)
Hispanic/Latino	5 (2.2%)	28 (13.3%)
Native Hawaiian/Pacific Islander	2 (0.9%)	0 (0%)
Two or More Races	5 (2.2%)	5 (2.4%)
White/Caucasian	212 (93.8%)	149 (71.0%)
Low Income	89 (39.4%)	96 (45.7%)
Non-Low Income	137 (60.6%)	114 (54.3%)
English Learner	6 (2.7%)	9 (4.3%)
Non-English Learner	220 (97.4%)	201 (95.7%)
Special Education Student	28 (12.4%)	56 (26.7%)
Non-Special Education Student	198 (87.6%)	154 (73.3%)
Total[3]	229	214

Of the 29 teachers, all were female and nearly all were White, with 26 teachers selecting this race/ethnicity category, one teacher selecting American Indian/Alaskan Native, one selecting

---

[3] Three students in the treatment group and four students in the comparison group did not have demographic data.

Hispanic/Latino, and one preferring not to specify. Further, all teachers reported having a bachelor's degree or master's degree as their highest level of education. In the treatment group, nine teachers had a bachelor's degree and five had a master's degree, while in the comparison group, eight teachers had a bachelor's degree and seven had a master's degree. Finally, teachers in the comparison group had more experience teaching elementary school than teachers in the treatment group. The average number of years teaching elementary school was 9.4 years for the treatment group (median was 8.5 years) and 16.5 years for the comparison group (median was 20 years).

Treatment Specific Measures

Treatment teachers completed a variety of study activities that were used to assess program implementation and impressions. This included an intake survey, a pre-survey and post-survey, a post-PD survey, and short weekly logs. Implementation data were also collected from the Lalilo student program and teacher dashboard and were provided by Renaissance Learning. Data were analyzed for these implementation measures and findings are shown below by research question.

Treatment Intake Survey

All 14 treatment teachers completed the intake survey prior to using Lalilo with their students. On this survey, teachers provided information about their teaching background and classroom environment, which served as contextual data for the study and determined study eligibility.

Treatment Pre-Survey

Thirteen of the 14 treatment teachers also completed a pre-survey at the beginning of the study, which covered their ELA instructional and assessment practices, resource availability and school-level interventions, and use of educational technology in classrooms. Their responses also served as contextual data for the study.

Treatment Post-Professional Development survey

Prior to the implementation of Lalilo, treatment teachers were able to review several asynchronous Lalilo training materials/resources (e.g., prerecorded videos, reports/articles, handouts/guides, blogs) to prepare them for the use of the student program and teacher dashboard in their classrooms. Thirteen of the 14 teachers provided feedback on the asynchronous training materials through the professional development (PD) survey taken prior to implementation.

Treatment Weekly logs

Teachers were instructed to complete short weekly logs about the week's instructional activities for each of the 12 weeks[4] of the study. Log completion varied across teachers; however, there was sufficient data to include these results in this report.

Treatment Post-Survey

After the implementation of Lalilo, all treatment teachers completed a post-survey at the end of the study. Topics included ELA instructional practices and classroom resources like in the pre-survey, but the post-survey also asked about teachers' perceptions and usage of the Lalilo program.

Lalilo Student Program and Teacher Dashboard

Throughout the course of the study, treatment group students used the Lalilo program in a one-to-one format, meaning each student was assigned their own device such as a Chromebook, laptop, or tablet when using the program. Students worked in Lalilo according to the level set for them by the placement test and made subsequent progress through the lessons.

During implementation, various metrics were recorded by the student program and teacher dashboard, including information about the dosage, performance of students, and actions completed by teachers.

Comparison Sample Survey Measure

All 15 comparison teachers completed a survey at the beginning of the study. On this survey, teachers provided information about their current instructional practices in English Language Arts and use of educational technology, in addition to contextual information about their classroom.

Program Impact Measures

In addition to the treatment specific program implementation measures and the comparison survey measure, treatment and comparison group students took the FastBridge earlyReading assessment twice in the 2023–24 school year. Treatment group students also took a modified version of the Elementary Reading Attitude Survey (McKenna & Kear, 1990) before and after the implementation of the Lalilo program. These measures, in addition to the student demographic data delivered by schools, were used to measure the impact of the program on

---

[4] 12 weeks was the recommended implementation period; however, the actual period was shorter for most teachers.

student reading achievement and reading attitudes following the implementation period. Data were analyzed for these impact measures and findings are shown below by research question.

FastBridge earlyReading

FastBridge earlyReading is a suite of assessments that was used to measure the growth in early literacy skills by comparing the treatment and comparison groups' change in student performance across the winter and spring administrations. Seven subtests from this suite were delivered during the winter and spring administrations, and four of these subtests were used to form a composite score that acts as the most reliable measure of literacy achievement.

Elementary Reading Attitude Survey

A modified version of the Elementary Reading Attitude Survey (McKenna & Kear, 1990) was delivered to treatment group students to estimate their reading attitudes before and after Lalilo implementation, with the goal of measuring potential changes in reading attitudes that can be attributed to the Lalilo program. The survey contains 22 multiple choice items and for each item students must select the facial expression that best represents their attitude for that item.

Findings

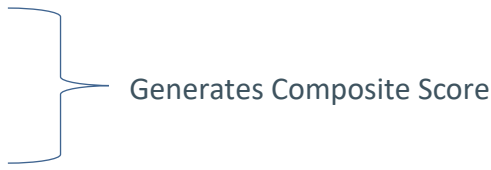## Research Question #1: What are the effects of Lalilo on student literacy achievement?

> **Key Findings:**
>
> 1. Treatment group students displayed significantly higher levels of performance on the FastBridge composite measure and across subtests at pretest than the comparison group, showcasing a lack of baseline equivalency; these differences were controlled for during the impact analyses.
> 2. Students in the treatment group grew significantly faster than the national median in terms of their FastBridge composite and FastBridge CBMreading scores.
> 3. The HLM analysis revealed no observable statistically significant impact on the composite score; however, a large positive impact was observed on the CBMreading subtest.
> 4. When examining the impact of Lalilo usage among the bottom 25 percent of students in terms of their winter FastBridge performance, no impact was observed on the composite score; however, there was again an impact on the CBMreading subtest, showcasing a positive effect of Lalilo on the most low-performing students.

FastBridge earlyReading Overview

As part of the evaluation, treatment and comparison group students were individually administered subtests from the FastBridge earlyReading[5] suite of assessments during the winter and spring of the 2023–24 school year to measure student growth in early literacy/reading skills and examine the impact of Lalilo on student literacy achievement. Students took 7 of the 13 subtests in the assessment, four of which were used to generate the student's composite score. The following subtests were selected for the evaluation as they most closely aligned with the skills targeted during use of the Lalilo student program:

---

[5] For more information about the FastBridge suite of assessments, view the technical manual using the following link: http://support-content.fastbridge.org/FAST_Research/FAST_Technical_Manual_Version_FINAL.pdf

- Word Segmenting
- Nonsense Words
- Sight Words
- CBMreading
- Letter Names
- Letter Sounds
- Word Rhyming

Generates Composite Score

For each of the subtests, identical forms were delivered during the winter and spring screening periods.[6] Following administration of the assessment, students were provided with a composite score, multiple subtest scores (e.g., items/words correct, items/words correct per minute), and composite and subtest benchmark levels that describe student performance.

Below, the results from the composite score and the CBMreading subtest are discussed as these measures are arguably the most valid given the research study and sample (findings for the remaining subtests can be found in Appendix A). The FastBridge earlyReading composite in "first grade demonstrated the highest level of criterion validity, suggesting that it is the best estimate of current and later broad reading performance" (FastBridge Learning, 2018). Similarly, the CBMreading subtest, which is a version of a curriculum-based measurement of oral reading fluency, has strong predictive validity with a correlation of 0.91 with AIMSweb and 0.82 with DIBELS Next (Renaissance Learning, 2024); the Nonsense Words and Sight Words subtests had weaker concurrent and predictive validity correlations with their criterion assessment than CBMreading. CBM assessments are also well-established progress monitoring tools and are considered by researchers to be a useful measure for informing reading instruction (Christ et al., 2012, 2013).

FastBridge earlyReading Data Availability

Within the treatment and comparison groups, 402 students (213 in the treatment group; 189 in the comparison group), took the necessary subtests to receive a composite score in both the winter and spring administrations.[7] The analytic sample for the composite and subtest measures differed substantially due to differences in FastBridge administration between sites. As shown in Table 3, the number of students with two scores (winter and spring) for the Letter Names, Letter Sounds, and Word Rhyming subtests is appreciably less than the number of

---

[6] For FastBridge CBMreading, the same three passages are used for all screening periods, however, the order of the passages changes for each screening administration.

[7] Twenty-four students across two comparison group classrooms took the Decodable Words subtest instead of the Nonsense Words subtest during the winter administration alone. Three students across one comparison group classroom took the Decodable Words subtest instead of the Nonsense Words subtest during both the winter and spring administrations. The Decodable Words subtest can be used as a replacement to the Nonsense Words subtest when creating the composite score.

students for the other subtests (particularly for the comparison group). Given this, the analyses for these three subtests were restricted, however, this was not of primary concern as these subtests were not used to generate the composite score.

There were also differences between the treatment and comparison groups in terms of the number of weeks between winter and spring FastBridge administrations. In general, the FastBridge assessment vendor recommends that there be 10–12 weeks between each screening period so that change can be observed. As shown in Table 3, the average number of weeks for the composite score was about 13.7 weeks for the treatment group (with a range of 5.4 weeks to 17.7 weeks) and 18 weeks for the comparison group (with a range of 12.1 weeks to 24.3 weeks). As a result, in both groups, students, on average, had more time between administrations than the typically suggested amount, allowing for sufficient student growth to be observed. However, given that treatment group students had less opportunities to grow than comparison group students, due to closer administration windows, the impact analyses focused on the average weekly growth of students (i.e., rate of improvement), rather than growth in raw scores.

Table 3. Number of Weeks Between Winter and Spring FastBridge Administrations

FastBridge Measure	Treatment or Comparison Group	N	Mean Number of Weeks Between Administrations (SD)	Minimum Number of Weeks Between Administrations	Maximum Number of Weeks Between Administrations
Composite	Treatment	213	13.7 weeks (3.3 weeks)	5.4 weeks	17.7 weeks
	Comparison	189	18.0 weeks (2.6 weeks)	12.1 weeks	24.3 weeks
Word Segmenting	Treatment	221	13.7 weeks (3.0 weeks)	8.0 weeks	17.7 weeks
	Comparison	198	17.7 weeks (3.0 weeks)	8.4 weeks	24.4 weeks
Nonsense Words	Treatment	220	13.6 weeks (3.2 weeks)	7.3 weeks	17.7 weeks
	Comparison	171	16.9 weeks (2.1 weeks)	8.4 weeks	21.1 weeks
Sight Words	Treatment	221	13.7 weeks (3.1 weeks)	4.6 weeks	17.7 weeks
	Comparison	198	17.7 weeks (3.1 weeks)	8.4 weeks	24.3 weeks
CBMreading	Treatment	216	13.6 weeks (3.3 weeks)	7.3 weeks	17.7 weeks

	Group	N			
	Comparison	189	18.0 weeks (2.7 weeks)	3.3 weeks	24.1 weeks
Letter Names	Treatment	148	11.9 weeks (2.5 weeks)	7.3 weeks	15.1 weeks
	Comparison	10	10.4 weeks (1.2 weeks)	8.4 weeks	12.1 weeks
Letter Sounds	Treatment	148	11.8 weeks (2.6 weeks)	7.0 weeks	15.3 weeks
	Comparison	10	10.4 weeks (1.2 weeks)	8.4 weeks	12.1 weeks
Word Rhyming	Treatment	148	11.9 weeks (2.5 weeks)	7.0 weeks	15.3 weeks
	Comparison	10	10.4 weeks (1.2 weeks)	8.4 weeks	12.1 weeks

Baseline Equivalence Between Treatment and Comparison Groups

As discussed previously, the treatment and comparison groups differed demographically, with the comparison group having a higher percentage of non-White students, low income students, EL students, and students with an IEP or Section 504 plan. The two groups also differed in their baseline (i.e., winter) FastBridge earlyReading performance, with the treatment group having significantly higher levels of performance on the composite measure and across subtests (see Tables 4 and A.1).[8] For example, on the composite measure, the winter score for treatment students was 0.62 standard deviations higher than the winter score for comparison students. Additionally, at baseline, the median composite score of the treatment group sample was at the national median (50th percentile), while the comparison group sample was below the national median (20th–25th percentile).

These performance differences are at least partially explained by the student populations of the participating schools. As shown, the comparison group had a higher percentage of students who were from historically underserved student populations than the treatment group. Given these differences, student and classroom-level characteristics were used as a baseline adjustment strategy during the hierarchical linear model (HLM) analysis.

---

[8] Comparison group students had a higher baseline score on one subtest (Word Rhyming). However, only 10 comparison group students were included in the sample due to minimal completion of the non-composite subtests.

Table 4. FastBridge earlyReading Baseline Equivalence

FastBridge Measure	Treatment or Comparison Group	N	Mean Winter FastBridge Score	Standard Deviation	Effect Size (Cohen's d)
Composite	Treatment	213	60.35	19.28	0.62
	Comparison	189	47.42	22.49	
CBMreading	Treatment	216	53.22	32.85	0.55
	Comparison	189	34.91	34.24	

Analytic Plan

The following two analytic methods were used to examine the impact of Lalilo on foundational literacy skills, as measured by FastBridge earlyReading performance: 1) comparison of the median average weekly growth rates of the treatment and comparison groups with the growth rates of the FastBridge national sample, and 2) comparison of the average weekly growth rates between treatment and comparison groups using hierarchical linear modeling.

Average Weekly Growth Rates

The composite and subtest median average weekly growth rates (or rates of improvement) of the treatment and comparison groups were compared with the growth rates of the FastBridge national sample to examine how both groups grew relative to the national median. These rates were computed for each student by dividing the overall change in performance by the number of weeks between administrations (see Tables 5 and A.2). While the sample of both groups differed from the FastBridge norming sample, according to FastBridge Learning (2019), the "growth rates on FAST measures are more sensitive to instruction than to the demographic composition of a group. Thus, even when the demographic composition of a group differs from the national sample, national growth percentiles still provide a useful context for interpreting progress and instructional effectiveness."

Table 5. Average Weekly Growth Rate Calculations

FastBridge Measure	Composite/Subtest Score Used	Growth Rate Calculation
Composite	Composite Score	Δ Composite Score / Number of Weeks Between Administrations
CBMreading	Total Correct/Minute	Δ Total Correct Per Minute/ Number of Weeks Between Administrations

Notes: The number of weeks between administrations was calculated by dividing the number of days between administrations by seven. This is equivalent to the method used by the FastBridge earlyReading vendor when determining a student's growth rate.

Hierarchical Linear Modelling

Hierarchical linear modelling was used to examine the impact of Lalilo usage on FastBridge performance, while accounting for the nesting structure of the data (students within classrooms) and controlling for different student and classroom-level characteristics[9]. The outcomes were the average weekly growth rates of the FastBridge measures of interest; we elected to use the growth rates rather than the raw scores to adjust for the baseline inequivalence shown previously and the differences in times between testing administrations that may have resulted in treatment and comparison students having unequal opportunities to grow. The resulting model is shown below.

Level 1 (Students): $Y_{ij} = \pi_{0j} + \sum_{a=1}^{A} \pi_{aj} STUDENT_{aij} + \varepsilon_{ij}$

Level 2 (Classrooms): $\pi_{0j} = \beta_{00} + \beta_{01} TREAT_j + \sum_{b=2}^{B} \beta_{0b} CLASS_{bj} + \omega_{0j}$

$$\pi_{aj} = \beta_{a0}, a = 1, \dots, A$$

where Y_{ij} is the student FastBridge outcome for each student i in classroom j; $Student_{aij}$ is a vector of student-level demographic characteristics; $TREAT_j$ is an indicator variable taking a value of 1 for treatment group classrooms and 0 for comparison group classrooms; and $CLASS_{bj}$ is a vector of classroom-level characteristics.

The student and classroom-level covariates were selected using the composite weekly growth rate as the outcome of interest and a forward stepwise approach whereby a single covariate was added at a time and only remained in the model if it significantly improved the fit of the model (according to the model's AIC and BIC statistics and likelihood ratio test results). Following this approach, the model included one student-level characteristic (IEP status) and three classroom-level characteristics (teacher education, frequency screening students for

---

[9] Classroom-level characteristics were reported by the participating teachers during the study survey administrations.

reading difficulties, and frequency providing tier 2 instruction). Student race, gender, socioeconomic status, and EL status were not found to be significant predictors of performance after inclusion in this model (this may have been partially due to the homogeneity of the sample). Additionally, the number of years teaching, length of English Language Arts (ELA) period, time preparing for ELA instruction, frequency using online/digital educational programs or games, frequency using technology-based dashboards, and comfort using online/digital educational programs or games were not significant classroom-level predictors. The covariates selected for this model were also used for the individual subtest models.

Findings

Average Weekly Growth Rates

We compared the median average growth rate of the treatment and comparison students with the growth rates of the FastBridge national sample to compare how the two groups grew relative to the national median. According to FastBridge Learning (2019), "[median growth rates] at least 10 points below or above [the] 50th [percentile] indicate the growth is significantly slower or faster than the national average." For example, a median growth rate at the 60th percentile indicates that the group grew faster than the national average. It should be noted, however, that these growth rates were developed using a normed sample from 2019, that is, before the COVID-19 pandemic. Therefore, the following results may change once updated norms are issued.

Students in the treatment group grew significantly faster than the national median (based on the 2019 normed sample) in terms of their composite and CBMreading scores — the median average weekly growth rate for these measures was between the 60th and 65th percentile (see Table 6). In contrast, the comparison group growth rate was at or even slightly below the national median for these measures (45th–50th percentile).

There was variation in how both groups grew on the remaining individual subtests (see Table A.3). However, as previously noted, we focus primarily on the composite and CBMreading measures as these are expected to be the most valid measures of future reading ability on the assessment. Additionally, there were other considerations that led to the focus on the composite and CBMreading measures.

The results on the Word Segmenting and Word Rhyming subtests may be driven by ceiling effects or the developmental appropriateness of the subtest. At baseline, both the treatment and comparison groups had average scores on the Word Segmenting and Word Rhyming subtests that were close to the ceiling of these subtests (34 and 16, respectively). Students, therefore, had limited growth potential on these two subtests, which likely influenced their subtest growth rates and the usefulness of these measures in estimating relative changes in performance.

Additionally, while the Word Rhyming, Letter Names, and Letter Sounds subtests can be administered to grade one students, they are viewed as more developmentally appropriate for kindergarten students or early or low-performing first grade students and may have been less sensitive to growth than the other subtests. Given that many students in the treatment sample may had already achieved mastery on those more foundational early literacy skills, there would have likely been limited growth potential on these subtests.

Table 6. FastBridge Winter–Spring Average Weekly Growth Rates

FastBridge Measure	Treatment or Comparison Group	N	Mean Winter Score	Mean Spring Score	Average Weekly Growth (Median)	National Growth Percentile
Composite	Treatment	213	60.35	73.17	0.96	60th – 65th
	Comparison	189	47.42	61.17	0.72	45th – 50th
CBMreading	Treatment	216	53.22	76.11	1.70	60th – 65th
	Comparison	189	34.91	59.60	1.31	45th – 50th

Notes: National Growth Percentiles are determined using the FastBridge norms from 2019.

We also examined the median average weekly growth rates by student demographic group (see Table 7). Across demographic groups, the treatment group generally had higher median average weekly growth rates on the composite score than comparison group students, as shown when observing the entire sample.

When disaggregating by gender, we observe that females and males in the comparison group had roughly the same median average weekly growth rates, while in the treatment groups, male students grew faster on average, than their female peers.

When disaggregating by socioeconomic status, we observe that non-low income students have higher weekly growth rates than low income students, in both the treatment and comparison groups. Interestingly, both non-low income and low income treatment group students grew faster than the national sample, suggesting a potential benefit among students from low socioeconomic backgrounds.

Finally, when disaggregating by IEP/Section 504 status, we observe that students with an IEP or Section 504 plan had roughly similar growth rates in the treatment and comparison groups

(both were below the national sample median), while students without these plans had a faster growth rate in the treatment group.

Table 7. FastBridge Winter–Spring Composite Average Weekly Growth Rates by Student Demographic Group

Student Demographic Group	Treatment or Comparison	N	Average Weekly Growth (Median)	National Growth Percentile
Female	Treatment	118	0.92	55th – 60th
	Comparison	89	0.72	45th – 50th
Male	Treatment	93	1.02	60th – 65th
	Comparison	97	0.72	45th – 50th
Low Income	Treatment	76	0.90	55th – 60th
	Comparison	80	0.62	40th – 45th
Non-Low Income	Treatment	135	0.99	60th – 65th
	Comparison	106	0.81	50th – 55th
Special Education Student	Treatment	24	0.47	30th – 35th
	Comparison	53	0.51	30th – 35th
Non-Special Education Student	Treatment	187	1.02	60th – 65th
	Comparison	133	0.81	50th – 55th

Notes: National Growth Percentiles are determined using the FastBridge norms from 2019.

The patterns of the results for the CBMreading subtest, broken out by student demographics (see Table 8), are roughly similar to the composite score results shown in Table 7. Once again, treatment group students generally had faster average weekly growth rates than the comparison group, and non-low income students and students without an IEP or Section 504 plan had a faster growth rates than their peers.

Table 8. FastBridge Winter–Spring CBMreading Average Weekly Growth Rates by Student Demographic Group

Student Demographic Group	Treatment or Comparison	N	Average Weekly Growth (Median)	National Growth Percentile
Female	Treatment	119	1.72	60th – 65th
	Comparison	89	1.40	50th – 55th
Male	Treatment	95	1.69	60th – 65th
	Comparison	97	1.14	40th – 45th
Low Income	Treatment	79	1.63	60th – 65th
	Comparison	80	0.97	30th – 35th
Non-Low Income	Treatment	135	1.85	65th – 70th
	Comparison	106	1.59	55th – 60th
Special Education Student	Treatment	24	1.03	35th – 40th
	Comparison	53	0.86	25th – 30th
Non-Special Education Student	Treatment	190	1.75	65th – 70th
	Comparison	133	1.45	50th – 55th

Notes: National Growth Percentiles are determined using the FastBridge norms from 2019.

Hierarchical Linear Modelling

A two-level HLM model was used to examine the impact of Lalilo usage on the average weekly growth rate on the FastBridge suite of assessments, while accounting for the nesting structure of students within classrooms and controlling for various student and classroom-level characteristics. Overall, there was no observed statistically significant impact of Lalilo usage on the composite score average weekly growth rate (see Table 9); however, there was a large[10] positive impact on the FastBridge CBMreading subtest (i.e., $p < 0.1$). After controlling for various student- and classroom-level characteristics, treatment group students had a slightly

---

[10] Interpretation of the effect sizes was based on this reference: Kraft, M. A. (2020). *Interpreting effect sizes of education interventions*. Educational Researcher, *49*(4), 241-253. https://doi.org/10.3102/0013189X20912798

higher average weekly growth rate of 1.83 words correct per minute than the comparison group; that is, the growth rate of the treatment group was 1.19 standard deviations more than the comparison group on the CBMreading subtest. This meant that the average treatment group student had a higher CBMreading weekly growth rate than 88%[11] of the comparison group students, after controlling for baseline differences (WWC, 2022).

Table 9. Hierarchical Linear Model Results of Composite and CBMreading Measures

	Composite	CBMreading
Lalilo Use	0.068 (0.085) g = 0.090	1.832* (1.006) g = 1.194
IEP/Section 504 Plan	-0.403*** (0.096)	-0.441*** (0.143)
Teacher Education Master's	0.143* (0.080)	-1.471 (0.978)
Screening Frequency 3-5 Times per Year	--	1.786 (2.497)
6+ Times per Year	0.256** (0.114)	1.559 (2.796)
Tier 2 Instruction 2 Times per Week	--	-9.570*** (1.838)
3 Times per Week	1.038** (0.422)	0.065 (1.205)
4 Times per Week	1.035** (0.438)	1.461 (1.606)
Daily	1.057** (0.417)	--
Intercept	-0.242 (0.418)	-0.650 (2.555)
Sample Size	375	378
Number of Teachers	26	27

Notes: * p < 0.1, ** p < 0.05, *** p < 0.01. "Screening Frequency" refers to the frequency screening students for reading difficulties (as reported by the teacher). "Tier 2 Instruction" refers to the frequency providing Tier 2 instruction (as reported by the teacher). Only data from 26 of the 29 recruited teachers were used for the composite score HLM model as one treatment teacher did not complete the pre-survey which provided contextual information about their classroom and one

[11] The percentile corresponding to the effect size was estimated by converting the effect size to Cohen's U3 index, as described in the WWC Version 5.0 Procedures and Standards Handbook. The index is the fraction of comparison group students outperformed by the average intervention group student and is computed based on the proportion of the area under a standard normal curve that is below the value of the effect size.

In addition to using the HLM procedure for the entire analytic sample, we conducted further analyses by restricting the samples to: i) students who had 12 or more weeks between their winter and spring FastBridge administrations, and ii) the bottom quartile of students in terms of their baseline FastBridge performance on the relevant measure.

When restricting the analyses to students who had 12 or more weeks between pre and post-testing (thereby, limiting the sample to students with longer Lalilo usage)[12], the results were very similar to the overall results (see Table A.5). Again, no impact was observed on the composite score; however, using Lalilo appeared to again have a large marginally significant positive effect on performance on the FastBridge CBMreading subtest. The effect on the CBMreading subtest was slightly higher when using the restricted sample (an effect size of 1.32 compared to 1.19). Put differently, among this sample, the average treatment group student had a higher CBMreading weekly growth rate than 91% of the comparison group students.

Similarly, when restricting the analyses to the bottom quartile of students (in terms of their FastBridge performance), Lalilo usage appeared to only have an impact on the CBMreading score (see Table A.6); however, the impact was smaller with this restricted sample (there was an effect size of 0.40). Although there was a smaller impact, the results revealed that Lalilo's effect on CBMreading performance extended to the most low-performing students, with the average low-performing treatment group student having a higher weekly growth rate than 65% of comparison group students.

---

[12] Twelve or more weeks between testing administrations does not necessarily equate to 12 or more weeks of Lalilo implementation.

Research Question #2: What are the effects of Lalilo on student ELA motivation and engagement?

> **Key Findings:**
>
> 1. Most teachers agreed that their students were motivated to use Lalilo, that they progressed through the activities in the program even when faced with difficulties, and that Lalilo improved their students' engagement with text and general motivation to read.
> 2. There was no observed change in student attitudes toward reading on the reading attitudes measure; however, males, low income students, and students with an IEP or Section 504 plan had less positive reading attitudes than their counterparts at both administrations.
> 3. Change in student attitudes toward reading had a small significant positive effect on FastBridge composite score rate of change.

Teacher Reported Reading Attitudes

In the post-survey, most treatment teachers agreed that students were motivated to use Lalilo (12/14), that students progressed through the activities in the program even when they faced difficulties (13/14), and that the program improved their students' engagement with text (10/14).

Above all, 10 of the 14 teachers (almost three-quarters) agreed that Lalilo increased student motivation to read, showcasing a perceived effect of the program on student attitudes in a broader context.

In the Phase I cohort, educators provided more positive responses overall regarding student motivation and engagement after using the program. The differences between the two phases may be partially a result of the different samples of districts and schools recruited. In the first phase, the majority of participating teachers were recruited from a single district, which had several reading specialists participating in the study. This inclusion of reading specialists from the primary recruited district may have helped teachers have a more positive experience when using the program as they may have had support from these reading specialists in implementing Lalilo with students.

Student Reported Reading Attitudes

Even though most teachers reported that students were motivated to use Lalilo and that the program had a positive impact on student engagement with text and overall motivation to read, students generally reported the same attitudes towards reading at the beginning and end of the study, as measured by an adjusted reliable and validated survey of attitudes toward reading (McKenna & Kear, 1990). The survey was adjusted to modify or add questions, and

modify the answer choices (i.e., change Garfield expressions to happy/unhappy facial expressions), due to the low test-retest reliability of the measure during Phase I. In the Phase II analysis, the adjusted survey continued to have questionable test-retest reliability (0.61) for the overall score but strong internal consistency for both administrations (0.91 at pre and 0.94 at post); the internal consistency of the measure during its development by its creators was 0.87 for grade one.

The analysis provided some insight into how the Lalilo intervention impacted student attitudes towards reading and involved calculating means and the differences between pre and post scores (overall and for each question) for the 167 treatment group students who completed the survey before and after Lalilo implementation. Overall, there were no observed changes in student attitudes before and after Lalilo implementation (the pre-score was 71.22 and the post-score was 71.16). Some minor differences were observed at the item level; however, some items had positive changes while others had negative changes, (see Table B.1), suggesting that, overall, students did not have measurable changes in their reading attitudes according to this survey.

The small pre-post differences among first graders can be attributed to various factors. First, a first grader's rapid emotional and attitudinal fluctuations could make their reading motivation more susceptible to short-term changes, thus reducing reliability over time. Second, environmental factors when taking the assessment, such as classroom setting and timing, may have also influenced their responses. For example, the pre-assessment conducted on a rainy day might have skewed their feelings about reading on rainy Saturdays. Finally, the relatively short timeframe between pre- and post-assessments may not have allowed sufficient time for meaningful changes in reading attitudes among first graders to occur.

Although there were no reading attitude differences between pre-Lalilo implementation and post-Lalilo implementation, there were differences between students of different identities at each of the administrations. Males had slightly less positive reading attitudes than females at each administration. Similarly, low income students and students with an IEP or Section 504 plan also had less positive reading attitudes than their peers not in those groups.

Finally, we used the HLM model created in the first research question to examine the impact of the change in student attitudes towards reading on FastBridge performance of the treatment group. While the change in student attitudes had a negligible correlation with FastBridge rate of change on the composite score (0.14), it had a small significant positive effect on the composite score rate of change (an increase of 0.013 points per week), while controlling for IEP status, teacher education, classroom frequency of screening students for reading difficulties, and classroom frequency of providing Tier 2 instruction.

Research Question #3: How was the Lalilo student program and teacher dashboard implemented?

> **Key Findings**
>
> 1. On average, students used the Lalilo program for 3.5 days per week, for an average of 40 minutes per week and 11 minutes per day.
> 2. After the implementation of Lalilo, most teachers reported that ideal usage was 11–20 minutes per lesson.
> 3. Students most frequently worked on exercises within the Phonics, Sight/High Frequency Words, Word Families, and Language Study domains; some usage also occurred within the Comprehension and Vocabulary domains.
> 4. Throughout the study, students validated an average of approximately four lessons and 21 learning objectives per week; there were disparities in engagement levels across reading domains.
> 5. Teachers used various modes of Lalilo implementation, with the primary mode of implementation being as a whole class tool.
> 6. Teachers most frequently used the dashboard to determine how much time an individual or the class spent on the program, to assign additional activities on Lalilo, and to track class and student performance on phonological and word work skills.
> 7. Across the implementation period, teachers logged on to the dashboard site an average of 4.3 days per week. Teachers frequently accessed the dashboard after logging on; however, they rarely viewed any of the student, lesson, or standards reports, if at all.

Student Program Implementation

Dosage of Implementation – Frequency and Duration of Usage

Across the implementation period, the student program data showed that students used the Lalilo program for 3.5 days per week on average, with a median usage of 4 days. Usage ranged from 2.4 days per week to 3.9 days across the implementation period.[13] Additionally, students used the program for an average of 40 minutes per week and 11 minutes per day (a weekly usage range of 19 to 46 minutes across the study). This mostly aligns with the study guidance to use Lalilo for 3–5 sessions per week (about 45–50 minutes total); however, compared to Phase I, there was less usage, with students using the program for an average of 53 minutes per week (about 21 minutes per day).

---

[13] All of the treatment teachers implemented Lalilo for at least seven weeks, with three teachers using the program for more than the recommended 12 weeks of implementation.

These results were in line with the teacher reports from the weekly logs. Across the implementation period, teachers reported using Lalilo an average of 4 days per week and they reported implementing the program in blocks of 11–15 minutes per day, in the majority of logs (60%; see Figure 1). In the post-survey, 10 of the 14 teachers reported that 11–20 minutes per lesson was the ideal usage time (two said 0–10 minutes and two said 21–30 minutes).

Figure 1. Duration of Lalilo Student Program Implementation (Minutes/Day)
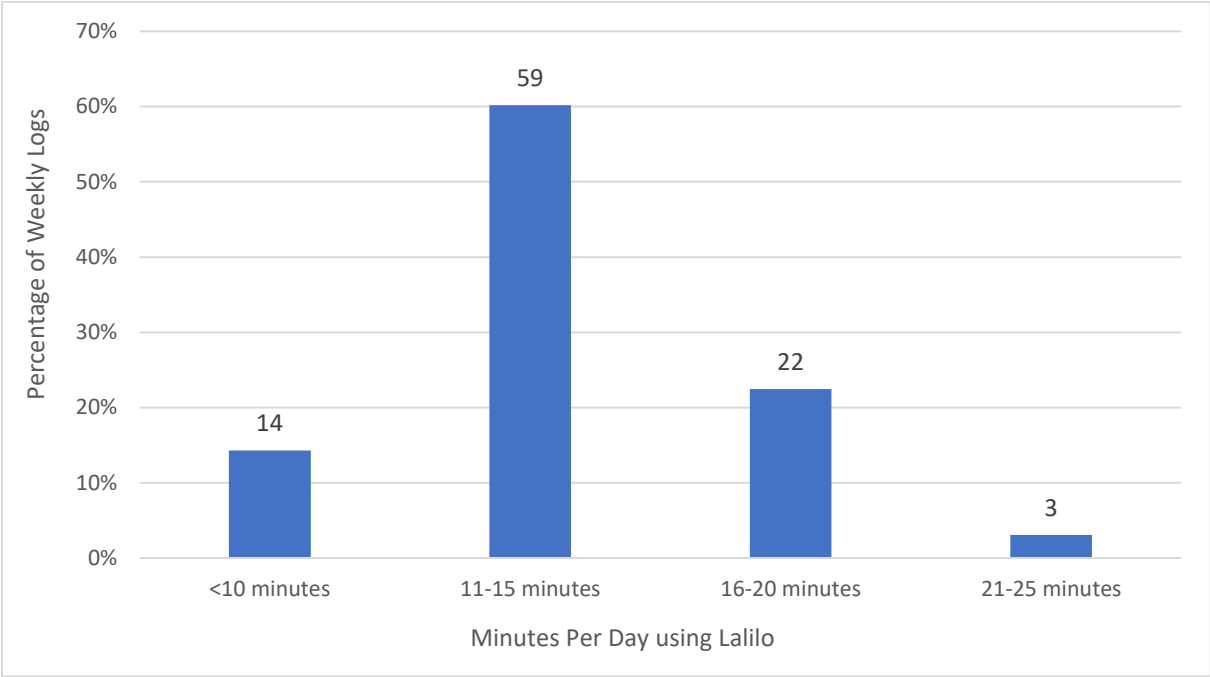
Like in Phase I, the frequency and duration of usage differed by ELA domain (see Table 10). Students were most frequently engaged in completing exercises within the Phonics, Sight/High Frequency Words, Word Families, and Language Study domains; some usage also occurred within the Comprehension and Vocabulary domains. For example, after combining all students and all weeks of implementation, the combined treatment sample worked on exercises within the Phonics domain for a total of 4,217 days. The other usage metrics (i.e., average number of days and minutes per week) showed a similar pattern of results, with the Phonics domain showing the most usage across the study (aside from the Vocabulary domain). While skills in the Vocabulary domain were least frequently targeted, during the weeks when this domain was targeted, students spent the most time on exercises within the Vocabulary domain.

Table 10. Frequency and Duration of Lalilo Usage by Domain

Domain	Total Days (Across Study)	Average Number of Days Per Week (Across Study)	Average Number of Minutes Per Week (Across Study)
Phonics	4,217	2.4	17.8
Sight/High Frequency Words	3,630	2.3	10.9
Word Families	2,528	2.0	10.3
Language Study	1,818	1.7	10.9
Comprehension	1,094	1.2	7.3
Vocabulary	327	2.4	24.4

Dosage of Implementation – Lesson and Learning Objective Validation

On average, students validated approximately four lessons and 21 learning objectives per week across the entire study implementation period, compared to six lessons and 30 learning objectives during Phase I. The number of lessons and learning objectives validated fluctuated from week to week (see Figures 2 and 3). Across the first seven weeks of implementation (i.e., the lowest number of weeks of implementation for a participating teacher), the range for lessons validated for individual students was 1 to 38, while the range for learning objectives validated was 1 to 225.

Figure 2. Average Number of Learning Objectives Validated Per Week (Weeks 1–7)
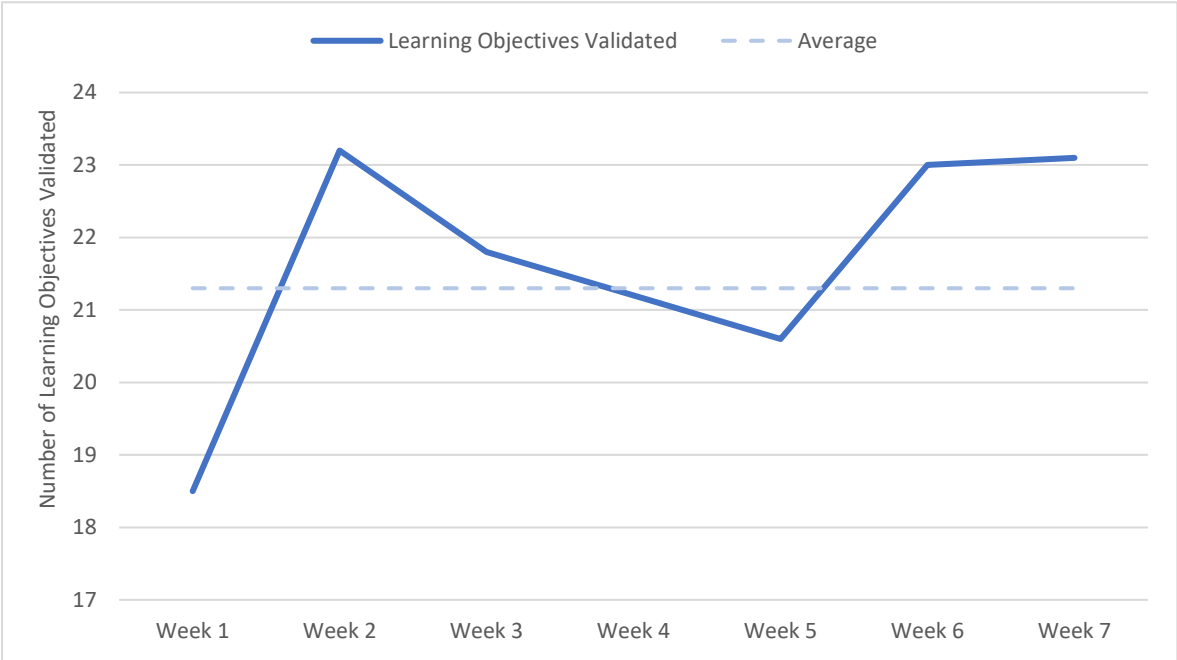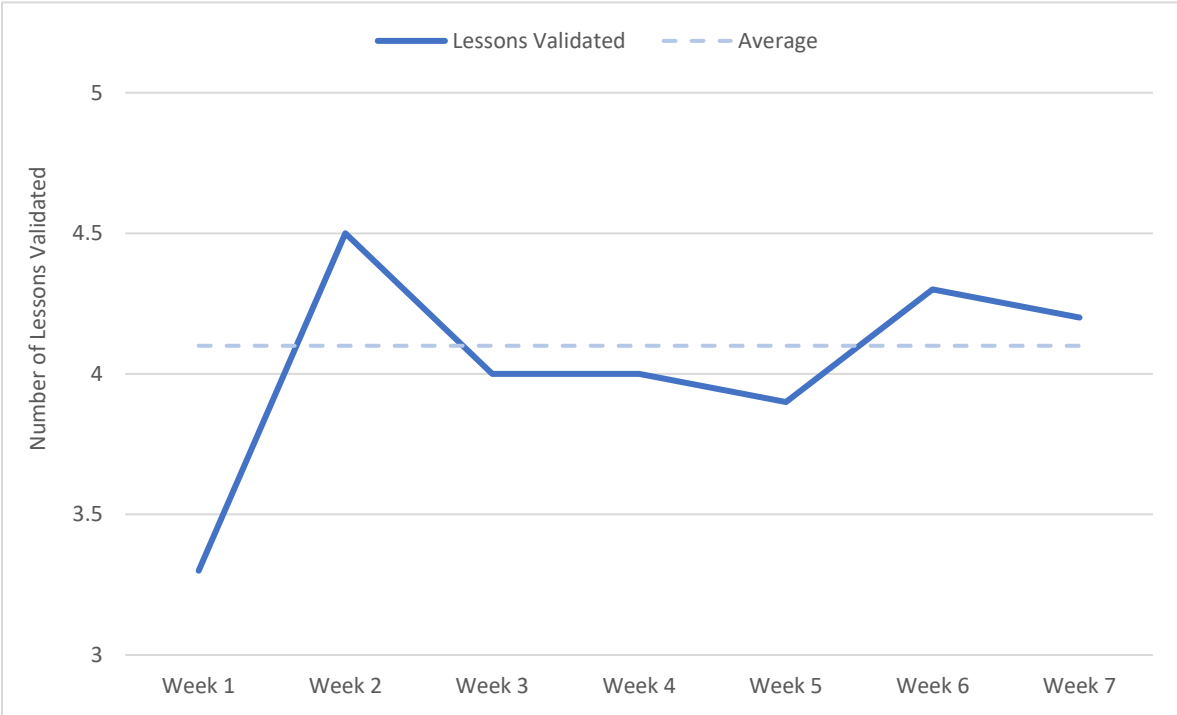

Figure 3. Average Number of Lessons Validated Per Week (Weeks 1–7)

As described previously, the data suggest that the program covered different domains to varying degrees (see Figure 4). While some domains like Phonics and Sight/High Frequency Words show higher engagement in terms of both the average number of lessons validated and the time spent, others like Comprehension and Language Study demonstrated lower engagement. This indicates that students are not equally engaged with all aspects of the program, which may be in part due to differing numbers of lessons for these domains in the scope and sequence and may also be due to the starting level of a student following the placement test. Vocabulary once again showed the highest engagement in the data due to high levels of focus on this area for students who received exercises from this domain.

Figure 4. Average Number of Learning Objectives and Lessons Validated Per Week by Domain
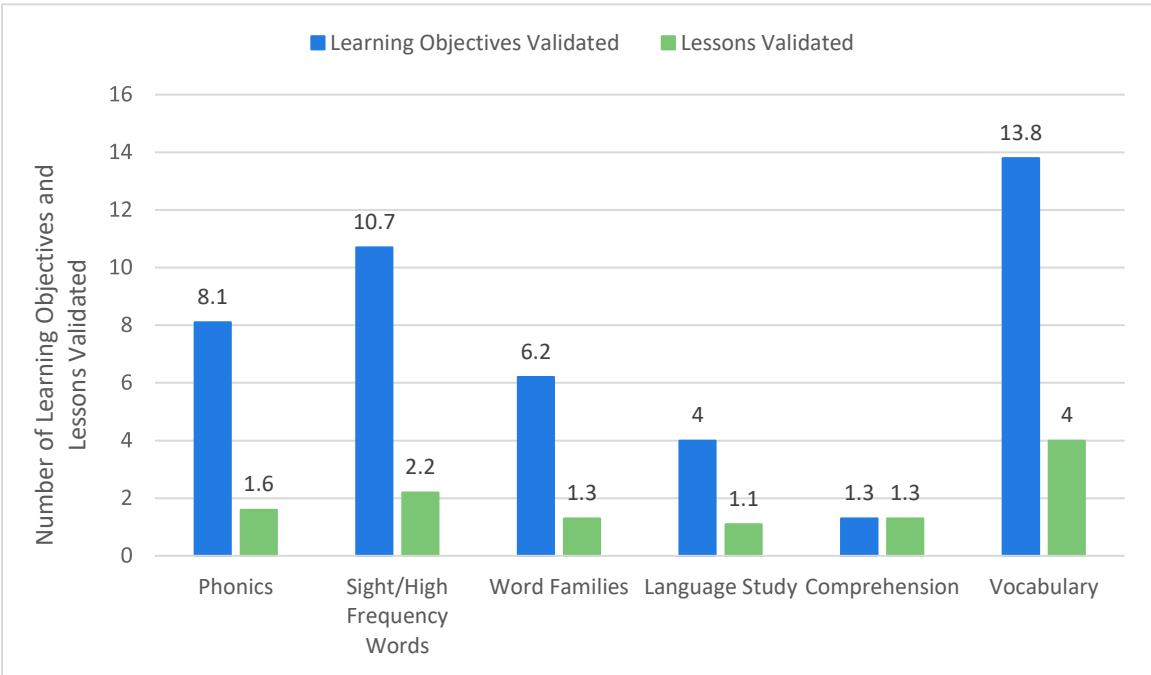
Modes of Implementation

The weekly logs provided insight into how the program was implemented in classrooms. According to teachers, Lalilo was primarily implemented as a whole class tool with one-to-one devices in the classroom or as a free choice option (e.g., students chose to work on Lalilo during free choice time). For instance, in 89% of all weekly logs, teachers reported using Lalilo as whole class tool. In addition to these two primary modes of implementation, teachers structured the use of Lalilo as centers, small group interventions, or homework (see Table 11). These results

are different than Phase I where teachers more frequently reported using the program as a small group intervention or center than as a free choice tool.

Table 11. Modes of Implementation (Frequency Reported in Weekly Logs)

Whole Class	Free Choice	Centers	Small Group Interventions	Homework
88.8% (n = 87)	26.5% (n = 26)	13.3% (n = 13)	8.2% (n = 8)	2% (n = 2)

Note: Teachers were allowed to select multiple answers in each weekly log.

Teacher Dashboard Implementation

Usage According to Teacher Reports

Across the implementation period, teachers most frequently reported, through the weekly logs, using the Lalilo teacher dashboard for 10 minutes or less after logging in (see Table 12), with teachers in 28% of logs reporting that they did not access the dashboard that week.

Table 12. Duration of Lalilo Dashboard Usage

Time Spent Using Dashboard	Percent Use (Across Study)
Did not access dashboard	28%
5 minutes or less	18%
6–10 minutes	29%
11–15 minutes	16%
More than 15 minutes	10%

Teachers reported most frequently using the dashboard to determine how much time an individual or the class spent on the program (see Table 13). Teachers also typically used the dashboard to assign additional activities on Lalilo and track class and student performance across different skills, specifically phonological and word work skills. This less frequent use of

the dashboard to determine performance on different ELA skills is in line with other teacher reports. In the post-survey, only half of teachers agreed that the Lalilo data from the teacher dashboard influenced the amount of emphasis that they placed on specific ELA elements during instruction.

Table 13. Dashboard Usage During Implementation Period

Dashboard Use	Percent Use (Across Study)
Determine individual time spent on program	58%
Determine class time spent on program	51%
Assign activities on Lalilo	37%
Determine class performance across phonological skills	29%
Determine class performance across word work skills	27%
Determine class performance across comprehension skills	5%
Determine individual student performance across phonological skills	26%
Determine individual student performance across word work skills	17%
Determine individual student performance across comprehension skills	6%
Detect students at-risk of reading difficulties	13%
Determine small groups for differentiated instruction	8%
Develop lesson plans	5%
Access educator resources (e.g., user guide)	3%

Note: Teachers were allowed to select multiple answers in each weekly log.

The 11 teachers who reported using the dashboard to assign activities noted that they focused on Phonics, Sight/High Frequency Words, Word Families, and Language Study (e.g.,

punctuation, parts of speech, capitalization). Additional activities for the Fluency, Comprehension, and Vocabulary domains were few, if at all (see Table 14). For example, across 86% of the logs where teachers reported assigning activities, activities from the phonics domain were reportedly assigned.

Table 14. Frequency of Assigning Activities to Domains

Domain	Percent Use (Across Study)
Phonics	86%
Sight/High Frequency Words	51%
Word Families	35%
Language Study	24%
Fluency	5%
Comprehension	0%
Vocabulary	0%

Note: Teachers were allowed to select multiple answers during each weekly log. Only teachers who reported assigning activities (i.e., those who reported this dashboard use in 37% of the logs), as indicated in Table 13, are included.

Usage According to Teacher Dashboard Data

Across the implementation period, teachers logged on to the dashboard platform an average of 4.3 days per week. During the first seven weeks of implementation (i.e., the least number of weeks of implementation for all participating teachers), all fourteen treatment teachers logged on at least once per week (see Figure 5). Additionally, after logging onto the platform, teachers were able to access the teacher dashboard, student report, lesson report, and standards report. Teachers frequently accessed the dashboard after logging on; however, they rarely viewed any of the student, lesson, or standards reports, if at all.

Figure 5. Average Number of Days Teachers Logged on to Educator Site (Weeks 1–7)
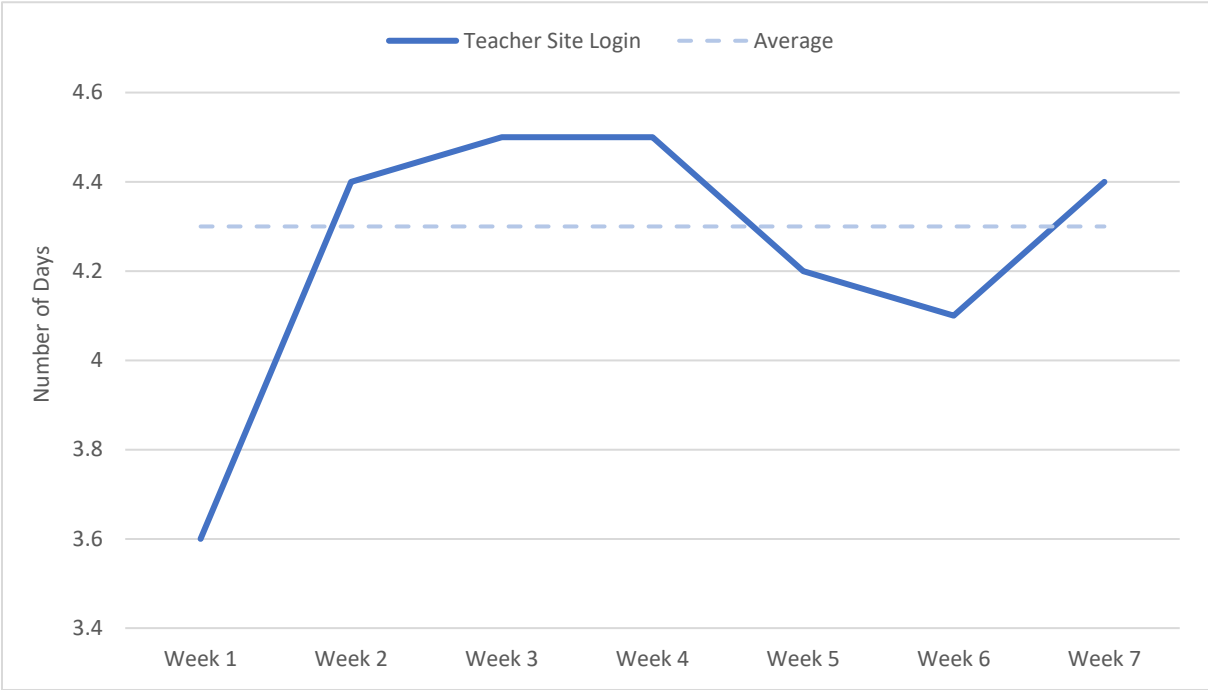
Research Question #4: How do students perform on the Lalilo in-program exercises and what is the association between Lalilo outcomes, program implementation, and FastBridge outcomes?

> **Key Findings:**
>
> 1. Across the entire study, students, on average, completed about 89% of exercises correctly.
>
> 2. Students' average exercise correctness rate did not change much between the first and last months of the study, suggesting that students were being equally challenged in the early and later stages of usage.
>
> 3. Students who scored higher on the placement test also tended to answer more Lalilo exercises correctly out of the total exercises they attempted. In addition, there were differences observed by domain.
>
> 4. Frequency and duration of Lalilo usage were not correlated to the average correctness of the Lalilo exercises completed by the students.
>
> 5. Lalilo average correctness was a significant predictor of rate of improvement on the composite measure of the FastBridge assessment, but not on the CBMreading subtest. Students who had a higher average correctness in the first month of Lalilo usage had a higher rate of improvement on the composite measure.

As students worked in Lalilo, the program recorded average correctness, or the number of exercises that students completed correctly divided by the total number of exercises attempted. Exercise correctness was grouped at the lesson level and averaged across lessons in monthly increments. Due to the adaptive nature of the Lalilo program, it was expected that students would be matched with appropriate content based on their placement test results and performance throughout the implementation period. This means that growth in a student's average correctness may not be observed as the program should appropriately adjust the difficulty of exercises as students progress through it.

As shown in Table 15, students on average completed 89% of exercises correctly, and the average correctness rate differed by literacy domain. The Sight/High Frequency Words domain had the highest correctness rate (96%), followed by Word Families and Phonics (85%), and Language Study (80%). This finding was consistent when reviewing the results from the first and last months of usage. Therefore, minimal growth was observed as students progressed through the program, suggesting that students were being equally challenged in the early and later stages of usage. The largest difference in the average correctness between the first and last months of usage was for the Word Families domain, where the average increased by just over three percent.

Table 15. Average Correctness in Lalilo Program by Domain

Lalilo Domain	Average Correctness Across Study	First Month Average Correctness	Last Month Average Correctness	Change Between First and Last Month
All Domains	88.8%	87.1%	87.4%	0.24%
Sight/High Frequency Words	95.6%	94.2%	95.2%	0.80%
Word Families	85.3%	81.9%	85.0%	3.36%
Phonics	85.1%	84.0%	85.0%	0.68%
Language Study	79.7%	80.1%	78.2%	-0.61%

Note: The Vocabulary domain was not included due to minimal completion of exercises.

Comparing Average Correctness for Students at Different Placement Indices

Students completed a placement test at the onset of using the Lalilo program, which determined where students would begin within the scope and sequence of the program. A higher placement index equates to a higher performance on the placement test and as students work in Lalilo, the program adjusts to target their needs.

As in Phase I, students who scored higher on the placement test also tended to complete more exercises correctly compared to the total number of exercises they attempted (see Table 16). Further, there was variation in how much the first and last month's average correctness differed by placement index; the largest differences were for the two lowest placement indices (a 7.1% difference for students with a placement index of 100 and a 10.9% difference for students with a placement index of 900).

Table 16. Average Correctness in Lalilo Program by Placement Index

Lalilo Placement Index	Average Correctness Across Study	First Month Average Correctness	Last Month Average Correctness	Change Between First and Last Month
100	78.5%	78.1%	85.2%	7.1%
900	82.6%	89.1%	78.2%	-10.9%
1210	84.1%	82.3%	83.4%	1.1%
2010	87.4%	84.0%	84.8%	0.7%
2680	87.6%	86.9%	86.3%	-0.6%
3500	86.8%	85.5%	87.1%	1.6%
3840	89.5%	88.4%	85.8%	-2.6%
4280	90.2%	88.4%	90.8%	2.4%
4550	90.3%	89.1%	88.5%	-0.6%
6000	93.7%	90.9%	92.3%	1.4%
6310	93.8%	94.1%	91.2%	-2.9%
8380	94.6%	93.1%	93.3%	0.2%

Association between Lalilo Average Exercise Correctness and Program Implementation

We examined the association between Lalilo average exercise correctness and program implementation. The findings, like in Phase I, suggested that variations in the frequency and duration of Lalilo usage, as measured by minutes active per month, exercises per month, days active per month, days using the dashboard per month, and days using the student report per

month, did not impact the average correctness of the Lalilo exercises completed by the students.

Association between Lalilo Average Exercise Correctness and FastBridge Performance

We also examined the association between Lalilo average exercise correctness and FastBridge performance to examine whether a student's performance on the Lalilo program predicted performance on the composite and CBMreading FastBridge measures. An HLM analysis revealed that average correctness in the first month of usage was a significant predictor of a student's composite rate of change, with students with a higher initial average correctness improving more per week on the composite measure, on average, than students with a lower initial average correctness. Conversely, average correctness was not a significant predictor of rate of improvement on the CBMreading subtest.

Research Question #5: What are teachers' impressions of the Lalilo student platform, teacher dashboard, impact on student learning and motivation, program content and pacing, as well as teacher resources and professional development as applicable?

> **Key Findings:**
>
> 1. Most teachers viewed the asynchronous Lalilo materials/resources positively, used most of the resources provided to them, but felt the pacing of the videos was too slow.
> 2. Most teachers found the dashboard useful for improving their reading instruction and felt it was easy to navigate, but three teachers disagreed.
> 3. All teachers, with the exception of one, reported they were likely to recommend the Lalilo program to other teachers.
> 4. Teachers generally agreed that Lalilo improved students' reading skills overall and across specific ELA elements.

Professional Development and Resources

All thirteen respondents to the PD survey[14] agreed that the training materials from the early literacy video (How to Enhance Foundational Literacy Instruction) sufficiently prepared them to start implementing Lalilo in their classroom. Most also agreed that the Lalilo resources[15] (e.g., videos, reports/articles, Lalilo handouts/guides, Lalilo blogs) were helpful in preparing them to implement the program and provided them with the information needed to effectively support their students as they began using Lalilo. Of the eleven teachers who viewed all of the resources, including the Lalilo handouts/guides and blogs, all agreed that each of these resources helped them better interpret the data from the teacher dashboard to inform their reading instruction.

Additionally, teachers expressed positive opinions about the support the asynchronous videos provided. Twelve of the 13 teachers agreed, prior to Lalilo implementation, that the how-to videos (How to Get Started with Lalilo, How to Interpret Lalilo Data, and How to Enhance Instruction with Lalilo) helped them better apply data from Lalilo to their lesson planning (one teacher slightly disagreed).

While teachers mostly approved of the Lalilo asynchronous materials/resources, they expressed mixed opinions about the pace of the videos, like in Phase I. Nine of the 13 teachers (about 69%) agreed that the pace of the video was too slow (see Figure 6), compared to just under half during Phase I. Conversely, teachers did express high engagement levels during the early

---

[14] One teacher did not complete the survey. Note that the PD survey was provided prior to Lalilo implementation.
[15] Two teachers reported that they did not view the Lalilo handouts/guides or blogs.

literacy videos (3/13 slightly agreed that they had high engagement, 7/13 agreed, and 3/13 strongly agreed).

Figure 6. Teachers' Opinions on Whether Pace of Lalilo How-to-Videos Was Too Slow
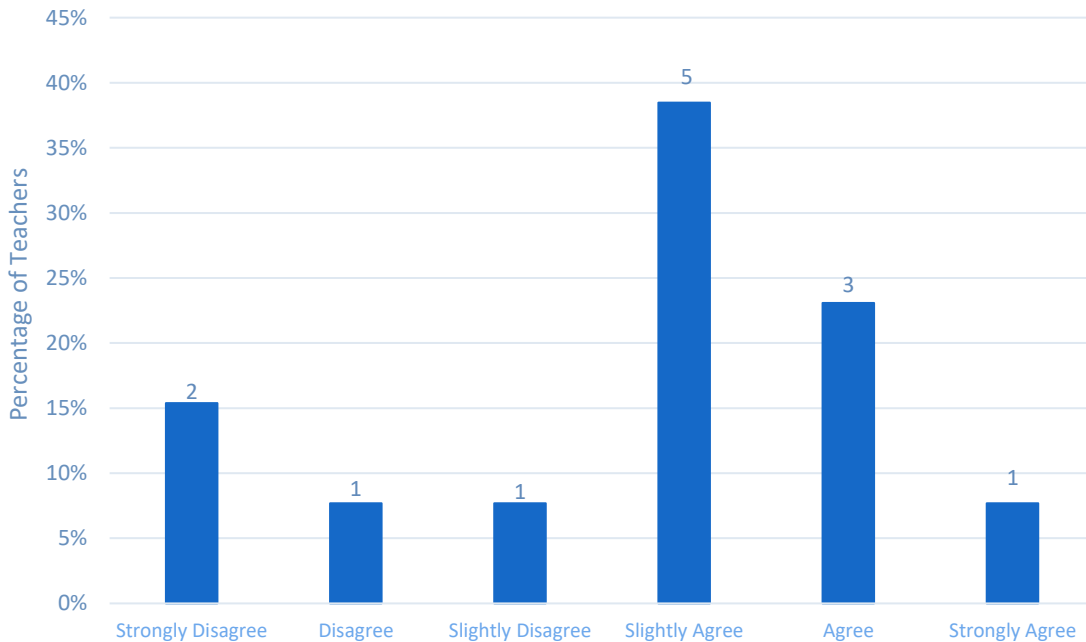
Note: Figure 7 contains responses from 13 of the 14 teachers; one teacher did not respond to this question in the PD survey.

Teacher Dashboard

Following the implementation period, 11 of the 14 treatment teachers expressed positive opinions about the dashboard data. Each of the eleven teachers slightly agreed (one teacher), agreed (four teachers), or strongly agreed (six teachers) that the dashboard provided them with data that was useful for improving their reading instruction (three disagreed that the data was useful for this purpose). Eleven teachers also agreed that the teacher dashboard was easy to navigate, while three teachers disagreed with this statement (see Figure 7).

Figure 7. Ease in Navigating Teacher Dashboard

Lalilo Student Platform & Impact on Student Learning

Like in Phase I, teachers also appreciated the instructional complexity and quality of the content in the Lalilo program, the quality of the graphics, the pacing of the program, and its overall quality.

> Most teachers reported they were extremely likely to recommend Lalilo to other teachers. On a scale of 0 to 10, where 0 is "not at all likely," 5 is "neutral," and 10 is "extremely likely," Lalilo received an average rating of 8.7 (nine teachers provided a 10 rating, two provided an 8 rating, two provided a 7 rating, and one provided a 2 rating).[16]

All teachers, with one exception, agreed that their students improved their overall reading skills by using Lalilo – one teacher slightly disagreed.

Many teachers also attributed Lalilo to their students' improvement in specific elements of ELA instruction like in Phase I. There were slightly higher levels of agreement among teachers that Lalilo improved their students' phonemic awareness skills, sight word reading skills, knowledge of word families, and listening comprehension skills compared to their fluency skills, where 21%

---

[16] The teacher who provided a rating of 2 noted that she had limited time with the program due to the onboarding process. This likely significantly impacted her reported rating.

of teachers (three) slightly disagreed that Lalilo had an impact on this ELA element (see Figure 8).

Figure 8. Teachers' Perspectives on Impact of Lalilo on ELA Elements
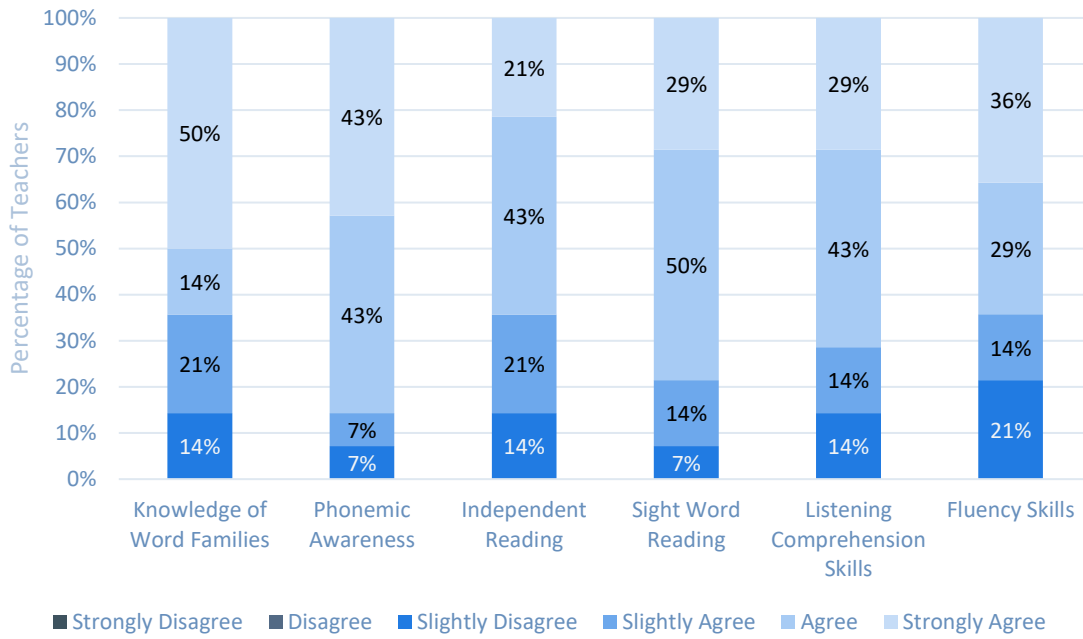
Research Question #6: How do contextual factors affect the implementation and subsequent success of Lalilo?

Key Findings:

1. Treatment teachers varied in their emphasis on ELA elements, ELA preparation time, and use of digital educational programs and dashboards prior to the study, as well as had different perceptions of the Lalilo program and dashboard usability at the conclusion of the study.

2. Examination of differences in teachers' contextual factors revealed some possible links to differences in Lalilo implementation. Greater usage was generally seen for students of teachers who spent more time preparing for ELA instruction, reported longer ideal Lalilo usage times, and fewer technical issues.

Contextual Classroom and Teacher Differences Among Treatment Sample

Teachers in the treatment group had differences in their instructional practices and classroom experiences that may have contributed to the findings regarding the implementation of Lalilo and the success observed in the program. Firstly, teachers differed in how much emphasis they placed on specific ELA elements. Overall, teachers reported emphasizing foundational literacy skills over higher-level reading skills, when asked prior to Lalilo implementation. Teachers most frequently reported emphasizing phonics, phonemic awareness, reading fluency, and sight word reading skills, with these elements receiving an average of at least 4.4 out of 5 (where 1 is "not emphasized" and 5 is "emphasized"). Reading comprehension, listening comprehension, speaking and listening, reading informational texts, word families, and reading literature (e.g., stories, drama) were less emphasized (see Table 17). This pattern of results was also observed after Lalilo implementation, with some minor changes in how much skills were emphasized.

Table 17. Emphasis on ELA Elements Pre- and Post-Lalilo Implementation

ELA Element	Average Emphasis Rating Pre-Implementation (on a scale of 1 to 5)	Average Emphasis Rating Post-Implementation (on a scale of 1 to 5)
Phonics and decoding	4.8	4.6
Phonology/phonemic awareness	4.6	4.4
Reading fluency	4.5	4.4

Sight word reading	4.4	4.4
Reading comprehension	3.8	3.6
Listening comprehension	3.7	4.0
Speaking and listening	3.7	3.9
Informational texts such as texts in social studies and science	3.5	3.3
Word families	3.2	3.6
Literature such as stories, drama, and poetry	3.2	3.2

Teachers also varied in the amount of time they reported preparing for ELA instruction each week (see Table 18); however, none reported spending 121 or more minutes per week. Thirty-eight percent of the teachers (i.e., 5 out of the 13 teachers who completed the pre-survey) reported spending 31 to 60 minutes each week preparing for ELA instruction, and another 38% reported spending 61 to 90 minutes per week.

There was also variation in the frequency with which teachers used digital educational programs and technology-based dashboards prior to the study. Before Lalilo implementation, ten treatment teachers reported using digital educational programs daily in their classrooms and three reported using them weekly. All teachers also reported feeling somewhat comfortable using digital educational programs, with 11 of the teachers reporting feeling comfortable or extremely comfortable. Teachers used technology-based dashboards less frequently than the digital educational programs and some did not express the same level of comfort. Of the 13 teachers who completed the pre-survey, five reported using dashboards daily, five reported weekly usage, one reported monthly usage, and two reported using them rarely. In addition, ten of the teachers reported feeling at least slightly comfortable using these dashboards, while three reported feeling slightly uncomfortable.

Table 18. Contextual Differences Between Treatment Group Teachers Pre-Lalilo Implementation

Contextual Characteristic	Frequency (Number of Teachers)	Percentage
Time Per Week Preparing for ELA Instruction Prior to Study		
30 or fewer mins/week	2	15.4%
31–60 mins/week	5	38.5%
61–90 mins/week	5	38.5%
91–120 mins/week	1	7.7%
121 or more mins/week	0	0%
Frequency Using Digital Educational Programs Prior to Study		
Daily	10	76.9%
Weekly	3	23.1%
Monthly	0	0%
Rarely	0	0%
Never	0	0%
Comfort Using Digital Educational Programs Prior to Study		
Extremely Comfortable	5	38.5%
Comfortable	6	46.2%
Slightly Comfortable	2	15.4%
Slightly Uncomfortable	0	0%

Uncomfortable	0	0%
Extremely Uncomfortable	0	0%
Frequency Using Technology-Based Dashboards Prior to Study		
Daily	5	38.5%
Weekly	5	38.5%
Monthly	1	7.7%
Rarely	2	15.4%
Never	0	0%
Comfort Using Technology-Based Dashboards to Inform Lesson Planning/Differentiated Instruction Prior to Study		
Extremely Comfortable	5	38.5%
Comfortable	3	23.1%
Slightly Comfortable	2	15.4%
Slightly Uncomfortable	3	23.1%
Uncomfortable	0	0%
Extremely Uncomfortable	0	0%

Notes: One treatment teacher did not complete the pre-survey.

Additionally, as discussed in prior sections, teachers were asked, in a post-survey, for their opinions about ideal Lalilo usage time, the ease in navigating the teacher dashboard, and about technical issues experienced during implementation. Most teachers (10 of 14) reported that 11 to 20 minutes was the ideal Lalilo usage time per lesson, with two reporting 0 to 10 minutes and two reporting 21 to 30 minutes; no teachers reported more than 30 minutes (see Table 19). Further, 11 of the teachers at least slightly agreed that it was easy to navigate the teacher dashboard and that it was useful in providing data to improve reading instruction; however, two teachers disagreed, and one strongly disagreed. Finally, 11 teachers had very few or no

technical issues related to Lalilo during implementation and three teachers reported that they had some issues, but that they did not significantly impact their use of the program.

Table 19. Contextual Differences Between Treatment Group Teachers Post-Lalilo Implementation

Contextual Characteristic	Frequency (Number of Teachers)	Percentage
Ideal Lalilo Usage Time		
0–10 minutes per lesson	2	14.3%
11–20 minutes per lesson	10	71.4%
21–30 minutes per lesson	2	14.3%
31–40 minutes per lesson	0	0%
41–50 minutes per lesson	0	0%
51 minutes or more per lesson	0	0%
Ease in Navigating Teacher Dashboard		
Strongly Agree	6	42.9%
Agree	3	21.4%
Slightly Agree	2	14.3%
Slightly Disagree	0	0%
Disagree	2	14.3%
Strongly Disagree	1	7.1%
Usefulness of Dashboard in Providing Data to Improve Reading Instruction		
Strongly Agree	6	42.9%

Agree	4	28.6%
Slightly Agree	1	7.1%
Slightly Disagree	0	0%
Disagree	2	14.3%
Strongly Disagree	1	7.1%
Incidence of Lalilo Technical Issues		
We did not have any technical issues with Lalilo.	6	42.9%
We had very few technical issues with Lalilo.	5	35.7%
We had some technical issues, but they did not significantly impact our use of Lalilo.	3	21.4%
We had technical issues that significantly impacted our use of Lalilo.	0	0%

Implementation and Lalilo Success Differences by Contextual Characteristics

We examined whether the average implementation of the Lalilo student program and the success of students within Lalilo differed depending on the contextual characteristics of the teachers and their classrooms. The goal of the analysis was to provide some context around the results and to examine how teacher characteristics may influence the delivery of the program.

Some differences were observed in the usage of the teacher dashboard between teachers with different instructional practices prior to Lalilo usage. On average, teachers who reported spending more time preparing for ELA instruction, prior to Lalilo implementation, logged into the teacher dashboard site more frequently each week (see Table 20). However, there was a less clear pattern when observing the number of days active on the dashboard and the student report; teachers who reported spending 31 to 60 minutes per week and 61 to 90 minutes per week had the most usage in these areas (this is likely partially explained by the fact that most teachers reported these ranges for their time preparing for ELA instruction).

Additionally, teachers who reported less frequently using dashboards prior to the study logged into the dashboard site more times per week, on average, than teachers who had greater pre-Lalilo dashboard usage. This may be a result of teachers with less experience using digital

educational programs and online dashboards wanting to gain experience to improve their use of the program's data for instructional purposes.

Table 20. Differences in Teacher Dashboard Usage by Time Per Week Preparing for ELA Instruction

Time Preparing for ELA Instruction	Days Logging In		Days Active on Dashboard		Days Active on Student Report	
	Mean	SD	Mean	SD	Mean	SD
30 or fewer mins/week	3.38	1.20	1	0	0	0
31–60 mins/week	4.21	1.33	2.31	1.15	1.25	0.5
61–90 mins/week	4.5	0.93	3.41	1.64	1	0
91–120 mins/week	4.8	0.79	1	0	0	0

Additionally, experiences with the Lalilo program also appeared to affect the frequency with which teachers used the dashboard. On average, teachers who disagreed that the dashboard was easy to navigate and that it provided data that was useful for improving reading instruction logged into the dashboard site less frequently and had a lower number of days active on the dashboard and student report than those who agreed about the ease of navigability and usability of the data. Furthermore, teachers who reported having some technical issues, as opposed to no issues or very few issues, had less dashboard usage as well.

The average number of learning objectives and lessons validated each week also somewhat differed by teacher and classroom characteristics. On average, students of teachers who reported spending more time preparing for ELA instruction had more validated learning objectives and lessons each week than students of teachers who reported spending less time preparing for ELA instruction (see Table 21), except for students of teachers who reported spending 91 to 120 minutes per week. Similarly, students of teachers who reported that the ideal Lalilo usage time was 0 to 10 minutes or 11 to 20 minutes per lesson had similar levels of learning objectives and lessons validated, however, students of teachers who reported an ideal usage time of 21 to 30 minutes had higher validation rates. Finally, students of teachers who reported having some technical issues with Lalilo had less validated learning objectives than students of teachers who reported very few or no technical issues, but they had more validated lessons.

Table 21. Learning Objectives and Lessons Validated by Contextual Characteristics

Contextual Characteristic	Learning Objectives Validated		Lessons Validated	
	Mean	SD	Mean	SD
Time Per Week Preparing for ELA Instruction				
30 or fewer mins/week	13.75	10.75	2.95	2.01
31–60 mins/week	21.03	20.07	4.44	4.05
61–90 mins/week	25.50	15.25	4.40	2.91
91–120 mins/week	16.37	10.03	3.23	2.22
Ideal Lalilo Usage Time				
0–10 minutes per lesson	21.17	13.61	3.89	2.62
11–20 minutes per lesson	20.24	14.66	3.82	2.72
21–30 minutes per lesson	25.98	22.65	5.26	4.51
Incidence of Lalilo Technical Issues				
No technical issues with Lalilo	21.85	15.99	3.91	2.93
Very few technical issues with Lalilo	21.83	14.07	4.15	2.81
Some technical issues but no significant impact	19.31	19.89	4.25	3.98

When examining differences, by contextual characteristics, in the average total time per week and time per day using the Lalilo program, the patterns reflect those shown above. On average, students of teachers who reported spending more time preparing for ELA instruction, longer

ideal Lalilo usage times, and fewer technical issues spent more time per week and per day, on average, than other students (see Table 22).

Table 22. Total Time Spent Per Week and Average Time Per Day By Contextual Characteristics

Contextual Characteristic	Total Time Per Week		Average Time Per Day	
	Mean	SD	Mean	SD
Time Per Week Preparing for ELA Instruction				
30 or fewer mins/week	26.19	15.01	9.64	4.28
31–60 mins/week	44.08	30.96	13.20	6.04
61–90 mins/week	46.35	23.31	11.65	4.47
91–120 mins/week	30.52	13.74	7.56	2.68
Ideal Lalilo Usage Time				
0–10 minutes per lesson	38.12	17.61	9.93	4.07
11–20 minutes per lesson	38.12	23.36	10.67	4.55
21–30 minutes per lesson	53.91	32.66	14.65	5.97
Incidence of Lalilo Technical Issues				
No technical issues with Lalilo	41.10	25.88	11.00	5.22
Very few technical issues with Lalilo	43.70	22.37	11.85	4.67
Some technical issues but no significant impact	31.84	26.16	10.31	4.88

Conclusion

WestEd conducted an impact and implementation study of the Lalilo program with 29 first grade teachers and their students across eleven sites, with 14 teachers in the treatment group and 15 teachers in the comparison group. Analysis of the impact data revealed that there was a lack of baseline equivalency as treatment and comparison group students differed demographically and as treatment group students displayed significantly higher levels of performance on the FastBridge earlyReading composite measure and across subtests at pretest (i.e., prior to Lalilo implementation) than the comparison group. Similarly, treatment group students at pretest were at the national median in terms of their composite score while the comparison group was below the national median. To adjust for these baseline differences, the average weekly growth rates of the FastBridge measures of interest were used as the outcome measures for the hierarchical linear model analysis, thereby controlling for the differences in performance at pretest and differences in the number of weeks between testing administrations across classrooms. Also, student- and classroom-level characteristics were included to control for demographic differences between the treatment and comparison groups.

Overall, treatment group students grew significantly faster than the national median in terms of their FastBridge composite and CBMreading scores. The hierarchical linear model results revealed no observable statistically significant impact on the composite score; however, a large positive impact was observed on the CBMreading subtest, with treatment group students having an average weekly growth rate that was about 1.2 standard deviations higher than the comparison group. This meant that the average treatment group student had a higher CBMreading weekly growth rate than 88% of the comparison group students. The positive impact was also observed when focusing on the bottom quartile of students in terms of their pretest FastBridge performance.

Like in Phase I, analysis of the teacher self-reported data revealed that teachers highly valued the program. They saw value in terms of student engagement (e.g., in some classrooms, students wanted to use Lalilo as a free choice option), believed their students progressed through the activities in the program even when they faced difficulties, and thought that the program improved their students' engagement with text. However, even though teachers reported that students were motivated to use the program, this did not necessarily translate to changes in students' self-reported motivation to read. This may be due in part to the difficulty shifting student motivation to read in only twelve weeks. However, although students generally reported the same attitudes towards reading at the beginning and end of the study, the change

in student attitudes toward reading had a small significant positive effect on the FastBridge composite score rate of change.

The teacher dashboard implementation data surfaced that teachers logged into the dashboard site frequently (i.e., an average of 4.3 days per week), but they rarely viewed any of the student, lesson, or standards reports, if at all. For teachers who reported accessing the dashboard in their weekly logs, they typically used it to determine how much time an individual or the class spent on the program, to assign additional activities on Lalilo, and to track class and student performance across different skills, specifically phonological and word work skills. This less frequent use of the dashboard to determine performance on different ELA skills is in line with other teacher self-reports. In the post-survey, only half of teachers agreed that the Lalilo data from the teacher dashboard influenced the amount of emphasis that they placed on specific ELA elements during instruction. To increase teachers' use of Lalilo for differentiated instruction and remediation, one recommendation is to offer more professional development on the dashboard and its features. Some teachers in Phase I discussed the value of live professional development sessions in particular, so more sessions might help support teachers' progress monitoring and small group instruction to improve students' reading skills.[17]

The student program implementation data surfaced that students used the Lalilo program for 3.5 days per week on average, with a median usage of 4 days, and that the program was used for an average of 40 minutes per week or 11 minutes per day. Additionally, students were most frequently engaged in completing exercises within the Phonics, Sight/High Frequency Words, Word Families, and Language Study domains and typically used the program as a whole class tool. Across the study, students had an 89% average exercise correctness rate in the Lalilo program. This rate differed by domain, with some students having higher levels of performance on some literacy domains. The rate also differed by student placement test performance.

Finally, as in Phase I, teachers mostly expressed positive opinions about the teacher dashboard and student program. Eleven teachers agreed that the dashboard was easy to navigate and that the data it provided was useful for improving their reading instruction. Teachers also appreciated the instructional complexity and quality of the content in the Lalilo program, the quality of the graphics, the pacing of the program, and its overall quality, and reported that the Lalilo program improved their students' overall reading skills. Furthermore, most teachers reported that they were extremely likely to recommend Lalilo to other teachers.

---

[17] In Phase I, teachers received live professional development during their use of Lalilo on the dashboard, but this did not happen in Phase II (i.e., all professional development happened before implementation where they viewed videos and resources asynchronously).

References

Christ, T. J., Zopluoglu, C., Long, J. D., & Monaghen, B. D. (2012). Curriculum-based measurement of oral reading: Quality of progress monitoring outcomes. *Exceptional Children, 78*, 356-373.

Christ, T. J., Zopluoglu, C., Monaghen, B. D., & Van Norman, E. R. (2013). Curriculum-based measurement of oral reading: Multi-study evaluation of schedule, duration, and dataset quality on progress monitoring outcomes. *Journal of School Psychology, 51*, 19-57. https://doi.org/10.1016/j.jsp.2012.11.001

Christ, T. J. and Colleagues. (2018). *Formative assessment system for teachers technical manual*. FastBridge Learning.

FastBridge Learning. (2019). *FastBridge Learning: Benchmarks and norms interpretation and use guidelines* (4th ed.). FastBridge Learning and Illuminate Education.

Kraft, M. A. (2020). Interpreting effect sizes of education interventions. *Educational Researcher, 49*(4), 241-253. https://doi.org/10.3102/0013189X20912798

McKenna, M. C., & Kear, D. J. (1990). Measuring attitude toward reading: A new tool for teachers. *The Reading Teacher, 43*(8), 626-639.

National Institute of Child Health and Human Development. (2000). *Report of the National Reading Panel. Teaching children to read: An evidence-based assessment of the scientific research literature on reading and its implications for reading instruction: Reports of the subgroups* (NIH Publication No. 00-4754). U.S. Department of Health and Human Services, National Institutes of Health.

Renaissance Learning. (2024). Technical manual: Psychometric evidence of the FastBridge universal screening & progress monitoring system. https://renaissance.widen.net/view/pdf/oc4ikwwfu5/PEFastBridge.pdf?t.download=true&u=zceria

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). https://ies.ed.gov/ncee/WWC/Docs/referenceresources/Final_WWC-HandbookVer5_0-0-508.pdf

Appendix A

Table A.1. FastBridge earlyReading Baseline Equivalence

FastBridge Measure	Treatment or Comparison Group	N	Mean Winter FastBridge Score	Standard Deviation	Effect Size (Cohen's d)
Word Segmenting	Treatment	221	32.47	2.94	0.39
	Comparison	198	31.17	3.72	
Nonsense Words	Treatment	220	22.15	13.68	0.51
	Comparison	171	15.93	9.67	
Sight Words	Treatment	221	54.29	24.01	0.76
	Comparison	198	35.14	26.14	
Letter Names	Treatment	148	63.45	15.56	0.53
	Comparison	10	55.3	8.92	
Letter Sounds	Treatment	148	51.07	13.46	0.91
	Comparison	10	38.7	14.47	
Word Rhyming	Treatment	148	14.01	2.79	-0.44
	Comparison	10	15.2	0.92	

Table A.2. Average Weekly Growth Rate Calculations

FastBridge Measure	Composite/Subtest Score Used	Growth Rate Calculation
Word Segmenting	Total Correct	Δ Total Correct / Number of Weeks Between Administrations
Nonsense Words	Total Correct/Minute	Δ Total Correct Per Minute/ Number of Weeks Between Administrations
Sight Words	Total Correct/Minute	Δ Total Correct Per Minute/ Number of Weeks Between Administrations
Letter Names	Total Correct/Minute	Δ Total Correct Per Minute/ Number of Weeks Between Administrations
Letter Sounds	Total Correct/Minute	Δ Total Correct Per Minute/ Number of Weeks Between Administrations
Word Rhyming	Total Correct	Δ Total Correct / Number of Weeks Between Administrations

Notes: The number of weeks between administrations was calculated by dividing the number of days between administrations by seven. This is equivalent to the method used by the FastBridge earlyReading vendor when determining a student's growth rate.

Table A.3. FastBridge Winter–Spring Average Weekly Growth Rates

FastBridge Measure	Treatment or Comparison Group	N	Mean Winter Score	Mean Spring Score	Average Weekly Growth (Median)	National Growth Percentile
Word Segmenting	Treatment	221	32.47	32.58	0	30th – 50th
	Comparison	198	31.17	32.39	0.06	60th
Nonsense Words	Treatment	220	22.15	27.60	0.45	60th – 65th
	Comparison	171	15.93	21.66	0.31	50th
Sight Words	Treatment	221	54.29	71.09	1.23	60th – 65th
	Comparison	198	35.14	54.75	1.04	55th
Letter Names	Treatment	148	63.45	67.51	0.38	15th – 20th
	Comparison	10	55.30	57.50	0.58	20th – 25th
Letter Sounds	Treatment	148	51.07	53.70	0.22	30th – 35th
	Comparison	10	38.70	37.40	0.05	20th – 25th
Word Rhyming	Treatment	148	14.01	15.02	0	N/A
	Comparison	10	15.20	15.20	0	N/A

Notes: National Growth Percentiles are determined using the FastBridge norms from 2019.

Table A.4. Hierarchical Linear Model Results of Word Segmenting, Nonsense Words, and Sight Words Measures

	Word Segmenting	Nonsense Words	Sight Words
Lalilo Use	-0.034* (0.019) g = -0.202	-0.026 (0.080) g = -0.050	0.026 (0.103) g = 0.028
IEP/Section 504 Plan	0.037* (0.022)	-0.135 (0.094)	-0.290** (0.118)
Teacher Education Master's	-0.038** (0.018)	0.201*** (0.075)	-0.016 (0.098)
Screening Frequency 3-5 Times per Year	-0.049 (0.061)	0.125 (0.258)	-0.033 (0.323)
6+ Times per Year	-0.063 (0.063)	0.329 (0.270)	0.125 (0.338)
Tier 2 Instruction 2 Times per Week	-0.007 (0.051)	-0.788*** (0.203)	-1.429*** (0.271)
3 Times per Week	-0.042* (0.022)	-0.143 (0.088)	0.238** (0.117)
4 Times per Week	0.072** (0.032)	0.041 (0.145)	-0.011 (0.169)
Daily	--	--	--
Intercept	0.115* (0.061)	0.242 (0.261)	1.228*** (0.327)
Sample Size	392	364	392
Number of Teachers	27	24	27

Notes: * p < 0.1, ** p < 0.05, *** p < 0.01

Table A.5. Hierarchical Linear Model Results of Composite and CBMreading Measures – 12+ Weeks Between FastBridge Administrations

	Composite	CBMreading
Lalilo Use	0.074 (0.090) g = 0.105	2.056* (1.230) g = 1.321
IEP/Section 504 Plan	-0.331*** (0.097)	-0.396*** (0.140)
Teacher Education Master's	0.125 (0.085)	-1.598 (1.210)
Screening Frequency 3-5 Times per Year	--	2.050 (2.739)
6+ Times per Year	0.263** (0.115)	1.689 (3.032)
Tier 2 Instruction		
2 Times per Week	--	-9.774*** (2.006)
3 Times per Week	0.992** (0.394)	-0.079 (1.441)
4 Times per Week	1.014** (0.408)	1.533 (1.746)
Daily	1.050*** (0.388)	--
Intercept	-0.240 (0.389)	-0.874 (2.815)
Sample Size	315	318
Number of Teachers	23	23

Notes: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table A.6. Hierarchical Linear Model Results of Composite and CBMreading Measures– Bottom Quartile in Winter Performance

	Composite	CBMreading
Lalilo Use	0.115 (0.132) g = 0.180	0.321* (0.181) g = 0.396
IEP/Section 504 Plan	-0.472*** (0.127)	-0.471*** (0.165)
Teacher Education Master's	-0.073 (0.147)	-0.038 (0.189)
Screening Frequency 3-5 Times per Year	--	0.054 (0.547)
6+ Times per Year	0.302 (0.238)	0.592 (0.605)
Tier 2 Instruction		
2 Times per Week	--	0.661 (0.753)
3 Times per Week	0.296 (0.593)	0.043 (0.196)
4 Times per Week	-0.087 (0.628)	-0.416 (0.270)
Daily	0.097 (0.589)	--
Intercept	0.414 (0.594)	0.845 (0.557)
Sample Size	104	104
Number of Teachers	23	23

Notes: * p < 0.1, ** p < 0.05, *** p < 0.01

Appendix B

Table B.1. Student Reading Attitudes Survey Results

Survey Question	Mean (Pre)	Mean (Post)	Change in Attitudes (Post-Pre)	Test-Retest Reliability
How do you feel when you read a book on a rainy Saturday?	2.94	3.13	0.19	0.22
How do you feel when you read a book in school during free time?	3.29	3.15	-0.14	0.40
How do you feel about reading for fun at home?	3.28	3.40	0.13	0.25
How do you feel about getting a book for a present?	3.35	3.32	-0.03	0.33
How do you feel about spending free time reading?	3.13	3.02	-0.10	0.40
How do you feel about starting a new book?	3.47	3.55	0.08	0.32
How do you feel about reading during summer vacation?	2.98	2.80	-0.19	0.34
How do you feel about reading instead of playing?	2.43	2.47	0.04	0.39
How do you feel about going to a bookstore or library?	3.65	3.59	-0.05	0.31
How do you feel about reading different kinds of books?	3.48	3.45	-0.03	0.27
How do you feel when the teacher asks you questions about what you read?	3.28	3.26	-0.02	0.17

How do you feel about filling out workbook pages and worksheets for reading tasks?	3.12	2.93	-0.19	0.32
How do you feel about reading in school?	3.41	3.49	0.07	0.31
How do you feel about reading your school books?	3.39	3.36	-0.03	0.32
How do you feel about learning from a book?	3.62	3.54	-0.08	0.23
How do you feel when it's time for reading at school?	3.34	3.31	-0.04	0.38
How do you feel about the stories you read at school?	3.44	3.51	0.07	0.33
How do you feel when you read out loud in class?	2.83	2.99	0.16	0.37
How do you feel about looking up what words mean?	3.29	3.20	-0.08	0.19
How do you feel about taking a reading test?	3.13	3.26	0.13	0.33
How do you feel about asking for help reading a book if you get stuck?	3.21	3.25	0.04	0.23
How do you feel about reading lots of books?	3.19	3.19	0.00	0.48
TOTAL	71.23	71.16	-0.07	0.61

Notes: Each item uses a 4-point Likert scale.

Appendix C

Table C.1. Comparison Teacher Contextual Survey (Open-Ended Responses)

Survey Question	Responses
Describe your learning environment where reading instruction is happening (for example, use of word wall, group-time to practice blending, etc.).	• I am an online teacher so students are at home and our live lessons are held in Class. It is a Zoom platform created for schools. I have 2 whole group reading lessons a week where we review high frequency words, spelling patterns, and a comprehension skill for the week. I have 2 small group interventions each week where we use the UFLI program to work on filling skill gaps. • I teach online. I teach language arts in whole group and small group reading groups. • Whole group at smartboard for letter-sound and word reading fluency, small group decoding/blending to read practice • SmartBoard, whole group instruction then independent work time. • whole group explicit instruction in phonemic awareness and phonics. Dry erase boards, magnetic letters, decodable reading independent application practice. • Whole group with a smart board, sometimes white boards and phonics puzzles, and small leveled groups using manipulatives • whole group reading instruction- phonemic awareness, phonics skills small group instruction- working with groups at similar skill levels using explicit PA, science of reading, phonics, fluency and comprehension skills. • large group and small group time • We use the UFLI curriculum. We also have a large collection of decodables and leveled texts for students to read during independent reading. I read several books aloud each day. I teach small groups daily for 50 minutes. My entire reading block last around 2 hours. • small group and whole group, centers, teacher table, para working with kids • We generally have stations. The stations consist of: Read to the teacher, word work, read to self and iPad app practice

	- Word Wall, station teaching, small groups, Independent work, skills block, literacy modules - Word walls, vowel town, whole group phonics practice/reading/listening. - Word wall, phonemic awareness large group with small group reteaching, large group phonics, Journeys large group instruction, Pathways Spelling and Pathways Small group instruction. - I do have a para full time in the classroom with me. She helps with those students that are not on grade level or identified as special education. We do have small group phonics in the morning where we work on reading words, completing word work to support the skill being taught, and working with the para to complete skill pages or reading. In the afternoon, we have small groups where we work on skills such as sound spelling, Vowel Town, reading words, sentences, and passages. My para will work on reading a decodable story as well as complete a skills activities to support the phonics skill. We do have a word wall and rule posters for other reading skills in the classroom.
How long is your Reading/ELA period?	- 40 minutes - 45 minutes - 60 minutes - 75 minutes - 60–90 minutes - 90 minutes (x5) - 90 minutes, split up in the day - 60 minutes whole group and 30 minutes small group - 105 minutes - 120 minutes (x2)
What core curriculum or instructional materials are you using to teach reading/ELA this year?	- We have an online curriculum created by Pearson for Connections schools. - Pearson Online Curriculum and UFLI - Sonday Essentials, Journeys - UFLI, Heggerty - Heggerty (phonemic awareness); UFLI (phonics) - EL Education - Functional Phonics and Journeys (x3) - Journeys curriculum, Heggerty, handwriting without tears - Journeys, Pathways, Heggerty - Pathways to Reading, rooted in reading, Heggerty's - Rooted in Reading, Pathways to Reading - Journeys Core Reading Curriculum Pathways for Small and Large Group Phonemic Awareness and Large Group Instruction and Small Group Instruction

	• Journeys and Rooted in Reading

Notes: All 15 comparison group teachers responded to the contextual survey.

Table C.2. Treatment Teacher Intake Survey (Open-Ended Responses)

Survey Question	Responses
Describe your learning environment where reading instruction is happening (for example, use of word wall, group-time to practice blending, etc.).	- Centers, group-time, phonics wall, phonics readers, small group work. - whole group, one on one instruction, segmenting and blending - Heggerty, anchor charts, whole group decoding and encoding, whole group comprehension, small group centers, one on one reading and fluency - Group time, some phonics. I do, we do, you do - Use of sound wall, group-time to practice sounds and blending, RTI small group time, sight words, Accelerated Reader program, Journeys reading program - Sight words posted, small and large group activities. Learning centers. Heggerty. - Use of word wall, group-time, RTI time where we have small group instruction of SIPPS program, practice sounds and blending. 10 minutes-phonemic awareness 30 minutes-phonics instruction-review, lesson, dictation, transfer 20 minutes practice - During reading instruction we do review, new skill direct instruction, dictation, and transfer. - Phonemic Awareness warm up, Reading Horizon lesson, sight word practice, worksheets, transfer worksheets - We use Orton-Gillingham during reading instruction. We do sound cards as a class, we do blending together as a class, we do sight word instruction as well as dictation for all sounds. - Orton Gillingham strategies and resources (multi-sensory, blending, sounds, dictation), reading applications, decodable readers, whole group phonics, whole group comprehension, whole group writing, small groups based on reading level/abilities - Whole group Orton-Gillingham phonics lessons including card pack, sand trays, blending board, red word practice, dictation, sentences, and syllabication. - classroom with whole group and small groups, sound wall, independent work time with different activities including word work, sight words, and writing - OG instruction is used whole group for phonics

Notes: All 14 treatment group teachers responded to the intake survey.

Table C.3. Treatment Teacher Professional Development Survey (Open-Ended Responses)

Survey Question	Responses
What did you like about the videos or resources?	• I really liked the reports that it gives on each students, and the explanation of what the things mean. • They answered questions that I had about how to better understand the data that I would be getting on each of my students. • The pace of the videos was good. They weren't too fast to listen to. • Visual of all the resources within the program. • Easy to follow and great information. • I like getting to know the program a little more before starting the study. • That I can refer back to them. • leveled instruction • I like that [trainer] showed how to utilize the website.
What did you not like about the videos or resources?	• It does not show how to add students to my classroom. I have already done that part, but it would be nice to see it. I also have been using this in my classroom already, so I'm not sure how I can switch my class now (which my free trial ran out on) to the class for this project. • They were very informative. • I wasn't sure if I needed to read all the blogs but I did read through some. • Nothing (x2) • They were a little lengthy for something that could have gone quicker. The website is pretty straightforward to navigate.
Please describe any additional training or resources that would have helped you to better start implementing Lalilo with your students. The information you share may help in informing the synchronous training sessions later in the study.	• I have already started. I do like how there are additional resources I'll be able to get and use once I have access to the project. • Not sure at this time/nothing (x3) • I think your resources were great! • I believe that LETRS was instrumental in helping me to better implement Lalilo. I also like that other teachers at my school use Lalilo-I can discuss it with them.
Please describe any additional training or resources that would have helped you to better use the data in the teacher dashboard to inform your instruction. The information you share may help in	• It would be nice if this data and the STAR CBM, reading, and early literacy data would all be combined to give a snap shot of my students. • Not sure at this time/nothing (x3)

informing the synchronous training sessions later in the study.	• I feel like once I get the program on the iPads I will feel more confident being able to use it hands on.

Table C.4. Treatment Teacher Pre-Survey (Open-Ended Responses)

Survey Question	Responses
How long is your Reading/ELA period?	• 60 minutes (x4) • 90 minutes (x7) • 120 minutes (x2)
What core curriculum or instructional materials are you using to teach reading/ELA this year?	• HMH Houghton Mifflin Harcourt • OG and HMH (x3) • HMH into reading (x2) • Journeys HMH and SIPPS • Journeys • Journeys, SIPPS • Savvas MyView Literacy • Reading Horizons (x3)
Is there any additional information you'd like to add about your reading instruction this year?	• I am currently taking part in the Science of Reading so with that I am learning quite a bit more than I did before. • No

Notes: Thirteen of the 14 treatment group teachers responded to the pre-survey.

Table C.5. Treatment Teacher Weekly Log #1 (Open-Ended Responses)

Survey Question	Responses
How did you first introduce the Lalilo program to students?	- Letters home, the script and survey, and then the placement activity - I introduced the program as a fun reading tool to help the kids show what they already know and to learn new concepts that will make them even better readers! - Students have used Lalilo previously (in the Fall of this school year) with the free version of Lalilo. Students were excited to give data to the people who run Lalilo and were told it is to learn new reading skills. - We have used it previously this year. It was one of their options during iPad station. I told them it would adapt to their skills. - I showed them the demo. - The students had already been playing the free version of Lalilo since the beginning of the year. - I told the students that I had a game for them to play that would help them become better readers. - I just sat them down and we talked about it. - I explained it was a game they could play to help them read & spell. They access it on their iPads. - I read the script from the Lalilo program. Let them know it was a program to help them build their reading and phonics skills. - I introduced Lalilo as a program to help reteach, preteach, and practice reading and writing skills. - We went on through the Renaissance sight. They are instructed that they need to stay focused and work the entire time. I showed them the link that takes them there. They are also instructed to wear their headphones and work quietly. - Did the pre test. - I introduced Lalilo in a whole group setting. I put Lalilo on our smart board and taught them how to login. They already knew how to log onto renaissance because they do it everyday for AR. I talked to them about how it's our new reading program that is taking the place of iRead in our daily routine.
Was there anything that was particularly confusing for students in the set up?	- They all seemed to understand Lalilo well. I have one student who is struggling, however, she is significantly academically lower than the other students. - The class/school code is a bit finicky for some students. It also asks them to ask an adult to approve the microphone

	usage each time. They also had some confusion figuring out the directions and the content.
	• No. The students were able to get on it and work right away. They actually say that it's easy to use. They like it better than the iRead that we used to use at the beginning of the school year.
	• No they got on pretty easily.
	• No (x10)
If you encountered issues affecting students' use of Lalilo, please explain.	• Lalilo often has no sound or will not allow them to push/select some of the buttons. To fix it you have to close completely out of the app and re-login.
	• They didn't put their chromebook on the charger so their battery died.
If students worked on other skills outside of Lalilo, please explain what the reading/ELA skills were and the context they used them in.	• Story structure, narrative sentences, -ar and -or words
	• OG foundational skills in whole group and small group, comprehension whole group
	• whole group, small group
	• none
	• New phonics skill was y as vowel long e
	• Comprehension skill was story elements
	• Writing skill was writing a narrative/personal story
	• We are learning silent e words this week and they read decodables and word lists. They also are using Read Live as a supplement for extra practice.
	• Reading Horizons, Reading Mastery, Read Live.
	• We worked on Magic e in phonics. We did this whole group, small group, independently, and in partners. We also worked on identifying text features in a nonfiction story. This was done whole group.
	• Digraphs
	• Sight words
	• Decodable Passages
	• Describing the order of events for our story of the week
	• Tier 1 instruction was teaching another sound for c and g.
	• We do heggerty daily as well. We also use hmh reading series.
	• Long i words in spelling. Vowel Teams

Notes: All 14 treatment group teachers completed the first weekly log.

Table C.6. Treatment Teacher Weekly Logs #2–12 (Open-Ended Responses)

Survey Question	Responses
If students worked on other skills outside of Lalilo, please explain what the reading/ELA skills were and the context they used them in.	• We worked with er, ir, ur words all week in our reading curriculum. We sang songs, write er/ir/ur words, and also read er/ir/ur words. • phonics skill y as a vowel e, comprehension- character, setting, problem, and solution, sight words • whole group/small group phonics, sight word, and fluency practice • Review of OG skills magic e, -dge, -tch, y as vowel long e/long I • Writing a personal narrative using CUPS • Story elements and story retell • ai/ay, were, from, as, they • oa, oe, their, there • We worked on long i sound. • We worked on reviewing previously learned skills. This included magic e, words ending with -nk/-ng, -ck words, and -tch words. This was done whole group, small group, and independently. We also learned some new sight words, this was done whole group, small group, and independently. • Review of previously learned skills (-ed, -tch, -dge, magic e), whole group, small group, independent time • We used Boomcards, dictation, and decodable reading passages. • Tier 1: Bossy r • Tier 3: Various skills depending on the group • -Silent e words core lesson • -Reading Mastery in leveled small groups • -Leveled individual fluency work • soft c/soft g, don't, more, house • Soft c/g • Writing a personal narrative • Soft c & soft g-whole class • intervention groups • Soft c & soft g, whole group, small group, independent tasks • Reading Mastery, Reading Horizons, epic • Reading Horizons and Reading Mastery • Reading Horizons- Whole group, Reading Mastery-Small group • We worked on long I patterns (ie, igh, and y) through our hmh reading program module 8 week two. We also do heggerty daily. • Heggerty and Module 8 week 3 in our HMH reading series. • again, animal, ee/ea

| | We continue to use Heggerty, (tap it, map it, graph it, and zap it), we use the hmh into reading series, and I also do phonics/word work in centers. Phonics & comprehension lessons Phonics (vowel teams ee, ea) , comprehension (main idea & detail) Vowel Teams Vowel team ea, ee Sight words again, animal, show We also worked on sight words and vowels this week. Vowel team ee/ea, independent time, whole group, small group, interventions Vowel teams oa and oe, sight words, comparing and contrasting Vowel teams-whole class tier 3 reteach/extension-small groups vowel team ai/ay, whole group, small group, interventions, independent work vowel teams-tier 1 various first grade skills-tier 3 opinion writing-tier 1 Vowel teams-whole group Intervention groups worked on a variety of skills depending on the group Vowel teams-whole group Intervention groups-varying skills Vowel Team ai, ay Suffix -s, -es Sight words its, mother, us Sight words were taught in our anchor text of the week. We also worked with -ed and -ing endings in words. We worked with -ed and -ing words all week again. We also worked on -er and -est words. Sight words as, they were mother, suffix -s/-es suffix -ed, which, small, four Suffix -ed, whole group, small group, independent time Suffix -ing, worked on it whole group/small group/interventions/independent work We also worked on suffixes (ly, y, ful). suffix -ing, been, would, could, how none (x3) |

Notes: All 14 treatment group teachers completed at least one log, in addition to their first weekly log. Multiple responses may be attributed to the same teacher.

Table C.7. Treatment Teacher Post-Survey (Open-Ended Responses)

Survey Question	Responses
What did you like best about using Lalilo with your students?	- being able to assign lessons - It's a great review for them to fill in any holes. - I liked being able to assign phonics skills and sight word practice for the words we were working on each week. I also like the ability to assign grammar practice. - I liked the individualized activities and related rewards for passing through each of the different stages. - Their desire to do it. - Being able to assign specific tasks - I like that Lalilo's sequencing is different and reteaches and teaches skills. - It was wonderful to use Lalilo assignments/lessons that correlated with our reading curriculum. - The students themselves loved the program, and would even want to get on it at choice time. - I really liked the fact that I could assign certain skills to my students to work on and when they saw they were able to recognize it both on and off the screen. The fact that my students were able to watch their own progress through the worlds left out the guessing game of how they were doing individually. - I liked the different worlds, how they could earn badges. - My students loved getting to new levels and earning badges. I loved being able to assign assignments that supported what we were learning during that week. - The students were always engaged and never complained about it. - I loved the option to assign specific skills to students, and how easy it was to track their progress. I also really liked the simplicity of the program as far as navigation/logging in/following the content.
What did you like least about using Lalilo with your students?	- It became unmotivating after a few weeks. - I don't like that they can sign in as someone else. - I would like more control over their journey. Sometimes I felt like they were able to choose to do other stuff besides what I had specifically assigned them. I wanted them to complete those activities first before practicing other skills. - Having to depend on their tablets and the internet to be working in order to use it. - I wish there was a timer for how long they should/could play - I don't like not being able to assign lessons.

	• Sometimes the internet at our school is weak so it would kick students off Lalilo. But overall there is nothing I did not like about Lalilo. • We enjoyed the program. • There really wasn't anything that I didn't like. I think it is a very good program. • I wish there were more stories to implement the skills learned. • I had one student finish Lalilo completely really quickly and then he couldn't access it anymore. • The time spent on the program did not correctly allocate as the week went on. It took nearly 80 minutes to get the recommended 45 minutes. This made it difficult to allow enough time per day. I also didn't like how it would make them start over if they didn't finish the entire session in one sitting. • Nothing
Was there anything other than technical issues that kept your students from working on Lalilo?	• Sometimes the speaker feature would not work for the students (they would have to do it multiple times to move on even if it was correct) • They could only use 15 minutes of Lalilo or the program would kick them out. They also did not like how the bar on the side would not always move up. • Time constraints • motivation • No (x7)
Was it hard for you to include Lalilo in your lesson plans? Please explain.	• Not usually, I found that doing it during morning work was effective. • No- It fit right in with our reading workshop time. • It would be good to do every other day or so to keep student interest level high, every day got a little repetitive for our age group. • No, I use Lalilo as an early finisher option. • No, it was a great tool and resource. I loved it and so did the kids! • No, because it benefited and engaged the students. • no. I assigned it as part of their morning work so that we were sure that we wouldn't miss working on it. • Yes, only because I was trying to have them use the program for the recommended amount of time but it took way longer for them to reach that time so I had to continually extend the Lalilo time. • No (x6)
Which ELA elements were influenced the most and how? Please explain.	• sw skills

	- Phonics skill practice and sight word praxtice - Lalilo teacher dashboard allowed me to see what skills they were excelling in and where they needed help. I could make small groups based off this information to focus in on what the students need. - phonics, word families - Phonics and determining what skills each student needed to work on. - phonemic awareness and decoding - Phonological awareness and phonics
How could the teacher dashboard be improved?	- Easier way to view/print student data. The ability to share student data with students in a student-friendly way. - Maybe have an overview page that has more data at face value rather than having to click through different tabs. - I feel that it could possibly give you a graph or something that allows you to see the skills as a whole per the classroom. Such as on one page show me which students need to work on short vowels, or other skills. - The dashboard was easy to use and navigate. - I couldn't access the full version, so I wasn't able to use any data. - I do not know, the basic version only allowed me to see 2 students and only like one data score but I was not allowed to click the data and even see what their score was..... - I did not have access to the dashboard. - I really don't have any suggestions. - no improvements needed - not sure at this time
Please describe any other technical issues your students experienced and note how frequently they occurred (i.e., Never, Rarely, Monthly, Weekly, Daily).	- Rarely - Never (x4) - uncharged iPads - The bar on the side of their screen would not move them up. - We had some technical issues due to our internet provider. - Our school internet was at times weak, but nothing with Lalilo. - The only technical issue was on our schools end. Such as the internet being down.
Is there anything else you would like to share about the technical issues your students experienced, if any?	- speaking into microphone was sometimes an issue - None (x5)
Please explain your recommendation rating in the previous question.	as a free choice, individually assigning tasks

	• I have already recommended that our entire building start using this program next year. It will help with the ELA skills overall. • I already recommend Lalilo to other teachers. • It was a wealth of information from the reports for me. • I loved the convenience and ease of extra practice for students. • Lalilo is a wonderful reading resource that gives individualized help to each student while keeping them engaged and motivated to keep reaching their goals! • I like how it moves with the students and can be monitored • The kids were engaged in the lessons and eager to get to complete them! • The kids loved it and were learning from it. • My students and I equal love Lalilo. I believe it is a great program and benefits their development of foundational skills greatly. • I really like the program as well as my students. Watching many of them progress in their reading by using this program really allowed them have a growing love in reading. • I think the program has great lessons and a wide range of lessons. This is GREAT for differentiating and honing in on specific skills. However, I found my students weren't super motivated to use the program. • I liked it, but I don't know that it had a strong link to better reading. It was a good way for kids to relax while still learning. The kids really enjoyed it. • I could not fully use lalilo so I do not know much about it.
What are best practices you would suggest to other teachers who are considering using Lalilo in their classroom? Best practices could focus on the implementation model (whole class, rotations, etc.), use of the teacher dashboard, or another topic.	• assigning, responsive • It's a great support. • Whole class during a time where you're working with small groups. • I was using it whole class, but could easily be used for centers and with small groups. • whole class and centers • small group/independent time • I would suggest teachers to use it as a rotation or early finisher work. • Having Lalilo has a center in center rotations! • I liked using Lalilo as a whole class and for small group work. • I used this as a whole class with the students on their individual devices, as well as in centers. It was very effective either way. I would also be on my dashboard and watch their active minutes to make sure they were working on it. I also used another program called securly to watch their screens.

	- I would suggest that they use this not only as an individual but as whole class implemation because even your higher readers can benefit working on the program. - I would recommend assigning specific content as students need it, and I would recommend using it 2x a week in a whole group setting. - Use for targeted skill intervention and fluency and comprehension practice.
How could the Lalilo program be improved?	- motivational aspects - Allow students to keep progressing through different grade levels. Let them go further then second grade.. I have a really advanced reader so he was done with the program in no time. - I think the content is great. I would suggest maybe adding more "game" elements to keep students more engaged if the program is being used over an extended period of time. - It would be nice to have something on the dashboard to show if they're on task or not. - Get more out of the basic version. - Have continued curriculum to meet students who are really high! - timer for use - The students loved when they would pass a world! I enjoyed watching their excitement when they passed their goal. Thank you! - Not sure (4) - No improvement needed
Is there any other information about your experience with using Lalilo that you would like to share?	- My students loved advancing through the levels and they loved the books. - I really enjoyed the process and have been recommending to other teachers in other grade levels. - User friendly, students enjoyed it, win-win for everyone! - Having access to the paid version really helped me better understand all the program has to offer!! - It is great!! Great program! - None (x4)

Notes: All 14 treatment group teachers responded to the post-survey.