THE EDTECH COLLECTIVE
Instructure Partner Ecosystem

READING EGGS

ESSA Level III Study (2023–24)

Prepared for:
3P Learning

Prepared by Instructure:
Meetal Shah, Ph.D., Lead Researcher

May 2025

EXECUTIVE SUMMARY

3P Learning contracted with Instructure, a third-party edtech research company, to examine the relationship between usage of Reading Eggs and student reading achievement. LearnPlatform designed the study to satisfy Level III requirements (Promising Evidence) according to the Every Student Succeeds Act (ESSA, 2015).

Study Sample and Methodology

This study was conducted with data from the 2023–24 school year and included 952 K–2 students across eight elementary schools in one public school district in Indiana. Researchers conducted analyses by grade level to allow for better interpretability of findings: Kindergarten (34%); Grade 1 (38%); and Grade 2 (28%). In terms of demographics, the total sample included White (60%), Black/African American (16%), multi-race (12%), and Hispanic (11%). In terms of gender, 50% of students were female.

Researchers used two key measures to provide insights into *Reading Eggs* implementation and potential impacts of *Reading Eggs* on student reading outcomes: *Reading Eggs* usage data and Amira's benchmark assessment Amira Reading Mastery (ARM) scores. Researchers used a variety of quantitative analytic approaches. First, researchers conducted descriptive statistics to describe participant characteristics and support implementation analyses. Researchers then conducted linear regressions to examine how use of *Reading Eggs* related to student reading achievement from fall 2023 to spring 2024. In addition, researchers calculated standardized effect sizes (Hedge's *g*) and the WWC improvement index to determine the magnitude of changes in student achievement.

Main Research Findings

Main Research Findings	
	Kindergarten
+	Kindergarten students who completed more *Reading Eggs* lessons had higher spring 2024 Amira Reading Mastery scores. **High vs Low Usage**: for a Kindergarten student at the 50th percentile, using at least 54 total lessons would result in them moving to the 70th percentile on average (i.e., a 20 percentile point improvement; $p < .001$). **Moderate Low Usage:** for a Kindergarten student at the 50th percentile, using between 25 and 53 total lessons would result in them moving to the 64th percentile on average (i.e., a 14 percentile point improvement; $p = .001$).
	Grades 1
+	Grade 1 students who completed more Reading Eggs lessons had higher spring 2024 Amira Reading Mastery scores. **High vs Low Usage:** for a Grade 1 student at the 50th percentile, using at least 54 total lessons would result in them moving to the 60th percentile on average (i.e., a 10 percentile point improvement; $p = .003$). **High vs Moderate Usage:** for a Grade 1 student at the 50th percentile, using at least 54 total lessons would result in them moving to the 62nd percentile on average (i.e., a 12 percentile point improvement; $p = .001$).

Conclusions

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

TABLE OF CONTENTS

INTRODUCTION

3P Learning recognizes that early elementary teachers often lack the capacity to meet the unique literacy needs of all students as providing effective supplemental self-paced literacy supports is often costly, time consuming, and fails to engage students in their own learning. *Reading Eggs* provides students aged 3–7 years with a comprehensive online literacy curriculum using thousands of ready-made and self-paced lessons, activities, and resources.

As part of their ongoing efforts to demonstrate the efficacy of *Reading Eggs*, 3P Learning contracted with Instructure, a third-party edtech research company, to examine the relationship between usage of *Reading Eggs* and student achievement. After co-developing an updated logic model (see Appendix A) for Reading Eggs (Scanlan & Henschel, 2022), researchers designed the study to satisfy Level III requirements (Promising Evidence) according to the Every Student Succeeds Act (ESSA) standards. Implementation of the *Reading Eggs* program among K–2 students did not yield a large enough comparison sample of non-users for an ESSA Level II design.

The following research questions guided this study:

Implementation

1. Overall, how many Reading Eggs lessons were completed by students during the 2023–24 school year?
2. Among Reading Eggs users, what were the usage patterns?

Student Outcomes

After controlling for students' prior reading achievement, gender, race, and grade,

3. How were different *Reading Eggs* usage patterns related to students' spring 2024 reading achievement?
    a. Which usage pattern(s) of *Reading Eggs* had the greatest impact on students' spring 20224 reading achievement?

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

STUDY DESIGN AND METHODS

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

Study Design

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

Setting and Participants

This study was conducted with data from the 2023–24 school year and included 952 K–2 students across eight elementary schools in one public school district in Indiana. Researchers conducted analyses by grade level to allow for better interpretability of findings: Kindergarten (34%); Grade 1 (38%); and Grade 2 (28%). In terms of demographics, the total sample included White (60%), Black/African American (16%), multi-race (12%), and Hispanic (11%). In terms of gender, 50% of students were female.

Measures

This study included the following measures to provide insights into *Reading Eggs* implementation and evidence about the potential impacts of *Reading Eggs* on students' reading outcomes.

Reading Eggs Usage Metrics. Researchers utilized 2023-24 student-level usage (i.e., total lessons completed). These usage data informed the extent to which students used *Reading Eggs* during the school year and whether students' use of Reading Eggs related to learning outcomes on the Amira Reading Mastery (ARM) assessment.

Standardized Student Assessments. The Amira Reading Mastery (ARM) score is a norm-referenced metric that reflects a student's reading level relative to their grade. This score helps educators assess students' reading proficiency. Each student receives a score indicating their grade (e.g., 2 or 3) and the month of instruction (e.g., 1 or 5). For instance, a third-grade student reading at the 50th percentile in December would have a score of approximately 3.5, representing Grade 3 and the fifth month of the school year. Researchers used ARM score as an overall measure of reading achievement at two time points: pretest (i.e., fall 2023) and posttest (i.e., and spring 2024). Each grade-level sample was analyzed separated as the ARM scores are not vertically scaled.

Data Analysis

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

Researchers then conducted linear regressions to examine how use of *Reading Eggs* related to student reading achievement from fall 2023 to spring 2024. In addition, researchers calculated standardized effect sizes (Hedge's *g*) to determine the magnitude of changes in student achievement.

IMPLEMENTATION

The charts below highlight *Reading Eggs* use during the 2023–24 school year based on 3P Learning's internal usage data (Table 1; details in Appendix B). Overall, K–2 students completed an average of 32 *Reading Eggs* lessons (SD = 23).

Table 1. Average *Reading Eggs* student usage by grade

Sample	*n*	Mean (Lessons Completed)	SD	Minimum	Maximum
Kindergarten	327	36	26	1	134
Grade 1	360	33	22	1	100
Grade 2	265	24	15	1	80

Researchers conducted a *k*-means cluster analysis to group students by similar levels of *Reading Eggs* usage based on the number of total lessons completed.

For total lessons completed, K–2 students fell into three usage categories ranging from low usage (mean = 12, range = 0–24 total lessons), to moderate usage (mean = 37, range = 25–53 total lessons), and high usage (mean = 71, range = 54–134 total lessons; Figure 1).

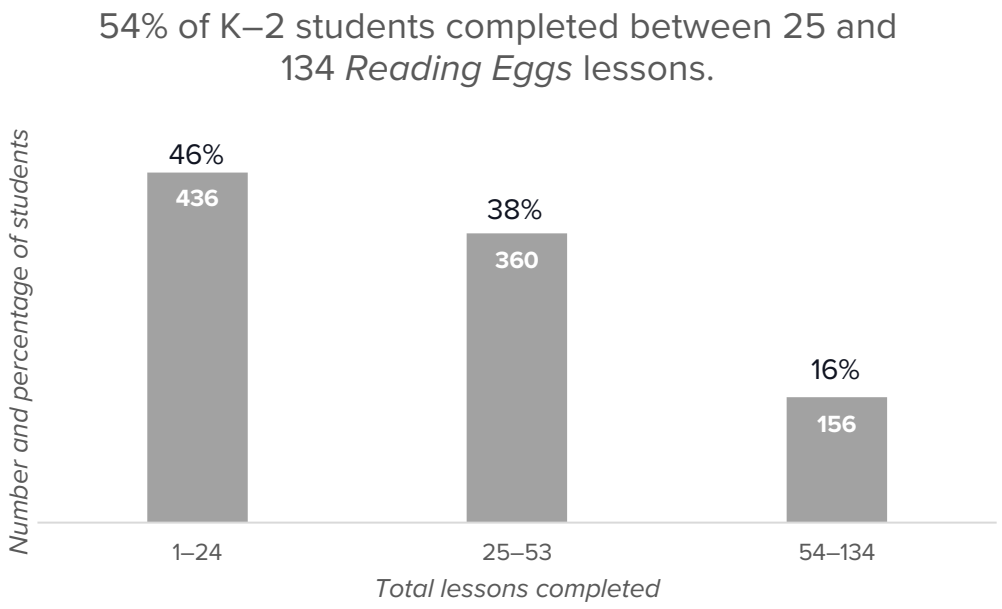


Figure 1. Overall distribution of total lessons completed on *Reading Eggs* by Grade K–2 students (*n* = 952)

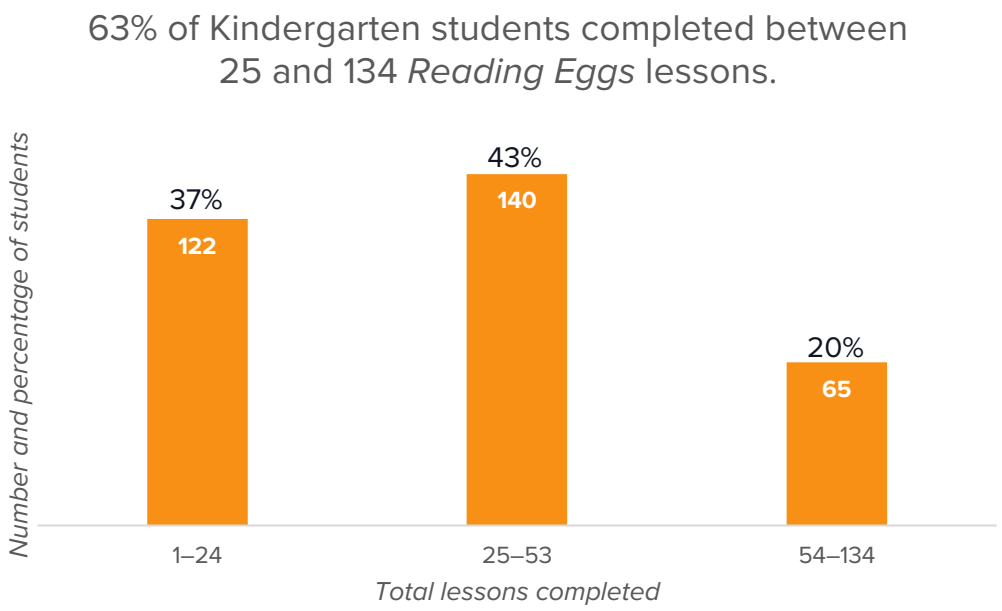


Figure 2. Overall distribution of total lessons completed on *Reading Eggs* by Kindergarten students (*n* = 327)

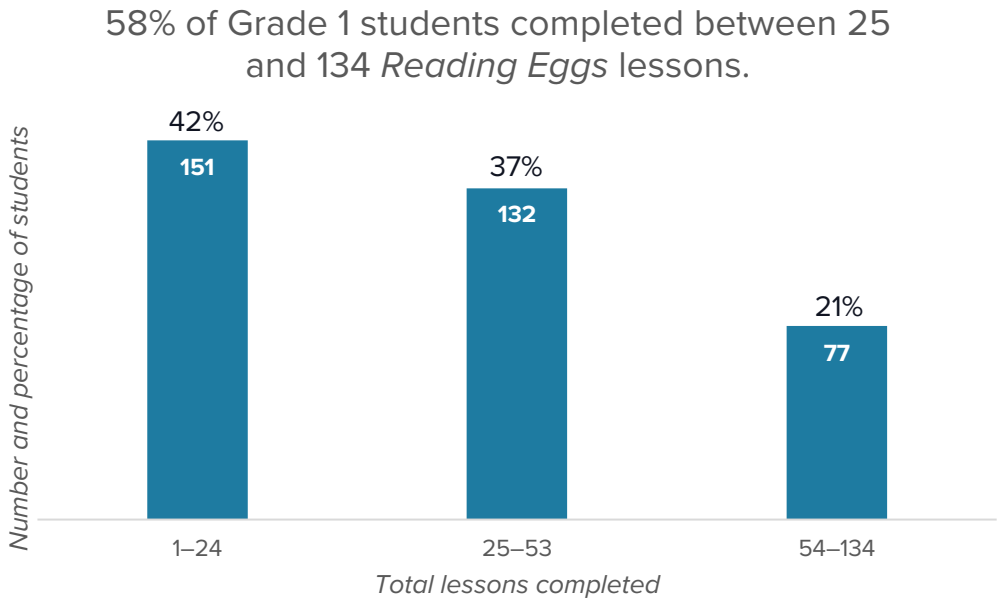


Figure 3. Overall distribution of total lessons completed on *Reading Eggs* by Grade 1 students (*n* = 360)

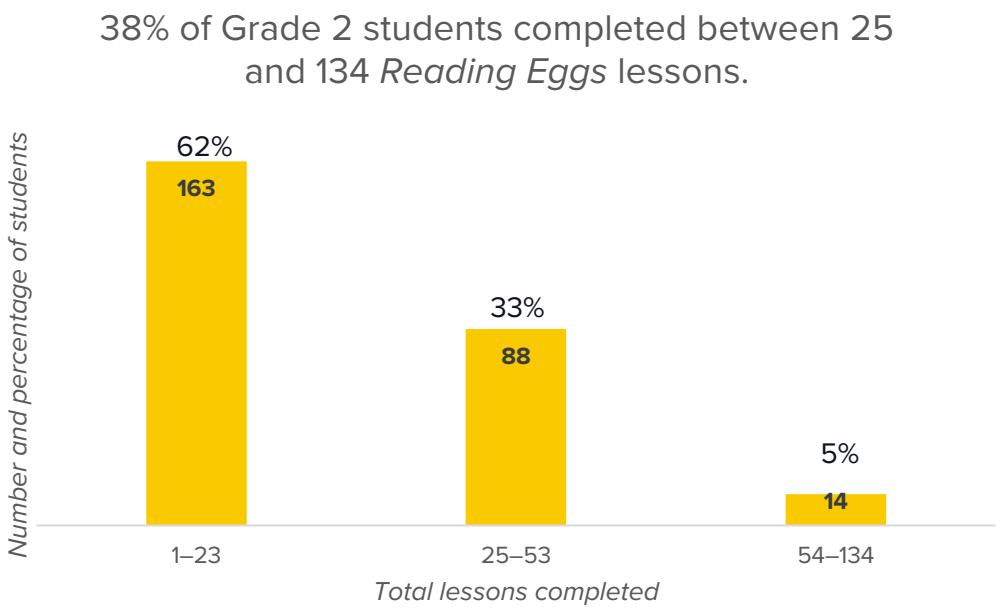


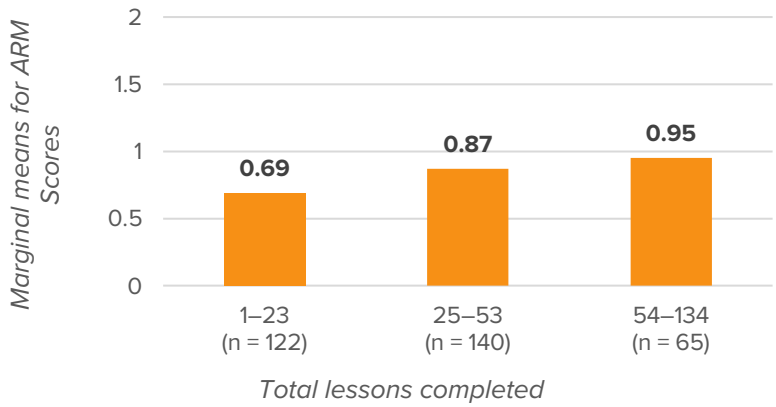
Figure 4. Overall distribution of total lessons completed on *Reading Eggs* by Grade 2 students (*n* = 265)

AMIRA READING MASTERY OUTCOME FINDINGS FOR K–2 STUDENTS

Researchers examined whether greater usage of *Reading Eggs* related to higher end-of-year Amira Reading Mastery scores using linear regression models that included beginning-of-year ARM scores, race, and gender as covariates. This relationship was investigated individually, for each grade. To allow for better interpretability of results, marginal means charts are presented below (see Appendix C for more details about the model and the corresponding Hedges' *g* effect sizes and information about interpreting Amira Reading Mastery Scores by grade).

Association Between Total Lessons Completed and K–2 Students' Outcomes on the Amira Reading Mastery Assessment by Usage Groups

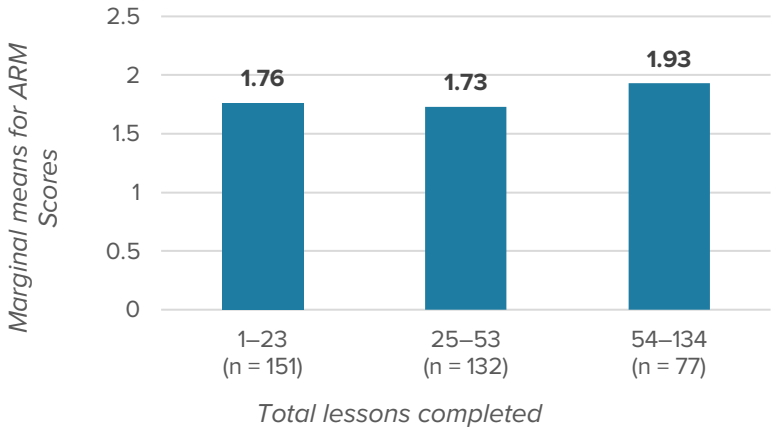
Kindergarten students who completed 25–53 (moderate use) and more than 53 (high use) lessons had **higher** ARM scores than students who completed fewer than 25 lessons (low use). These results were statistically significant at the p <.05 level.



For **high use compared to low use**, the ARM score difference of 0.26 (high: 0.95 vs. low: 0.69), equated to a 20 percentile point improvement. Such that, for a Kindergarten student at the 50th percentile, using *at least* 54 total lessons would result in them moving to the 70th percentile on average (i.e., a 20 percentile point improvement; p < .001).

For **moderate use compared to low use**, the ARM score difference of 0.18 (moderate: 0.87 vs. low: 0.69), equated to a 14 percentile point improvement. Such that, for a Kindergarten student at the 50th percentile, using between 25 and 53 total lessons would result in them moving to the 64th percentile on average (i.e., a 14 percentile point improvement; p = .001).

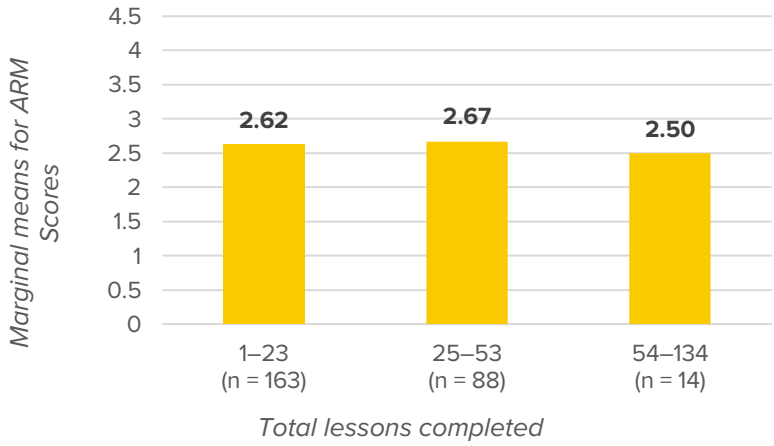
Grade 1 students who completed more than 53 (high use) lessons had **higher** ARM scores than students who completed fewer than 25 lessons (low use) and students who completed 25–53 lessons (moderate use). These results were statistically significant at the p <.05 level.



For **high use compared to low use**, the ARM score difference of 0.17 (high: 1.93 vs. low: 1.76), equated to a 10 percentile point improvement. Such that, for a Grade 1 student at the 50th percentile, using *at least* 54 total lessons would result in them moving to the 60th percentile on average (i.e., a 10 percentile point improvement; $p = .003$).

For **high compared to moderate use**, the ARM score difference of 0.20 (high: 1.93 vs. moderate), equated to a 12 percentile point improvement. Such that, for a Grade 1 student at the 50th percentile, using *at least* 54 total lessons would result in them moving to the 62nd percentile on average (i.e., a 12 percentile point improvement; $p = .001$).

Grade 2 students who completed 25–53 (moderate use) lessons had **higher** ARM scores than students who completed fewer than 25 lessons (low use) and more than 53 (high use). These results were **not** statistically significant at the $p < .05$ level.



CONCLUSIONS AND RECOMMENDATIONS

The findings support an association between *Reading Eggs* usage and improved Amira Reading Mastery scores for Kindergarten and Grade 1 students. This study provides results to satisfy ESSA evidence requirements for Level III (Promising Evidence). Specifically, this study met the following criteria:

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

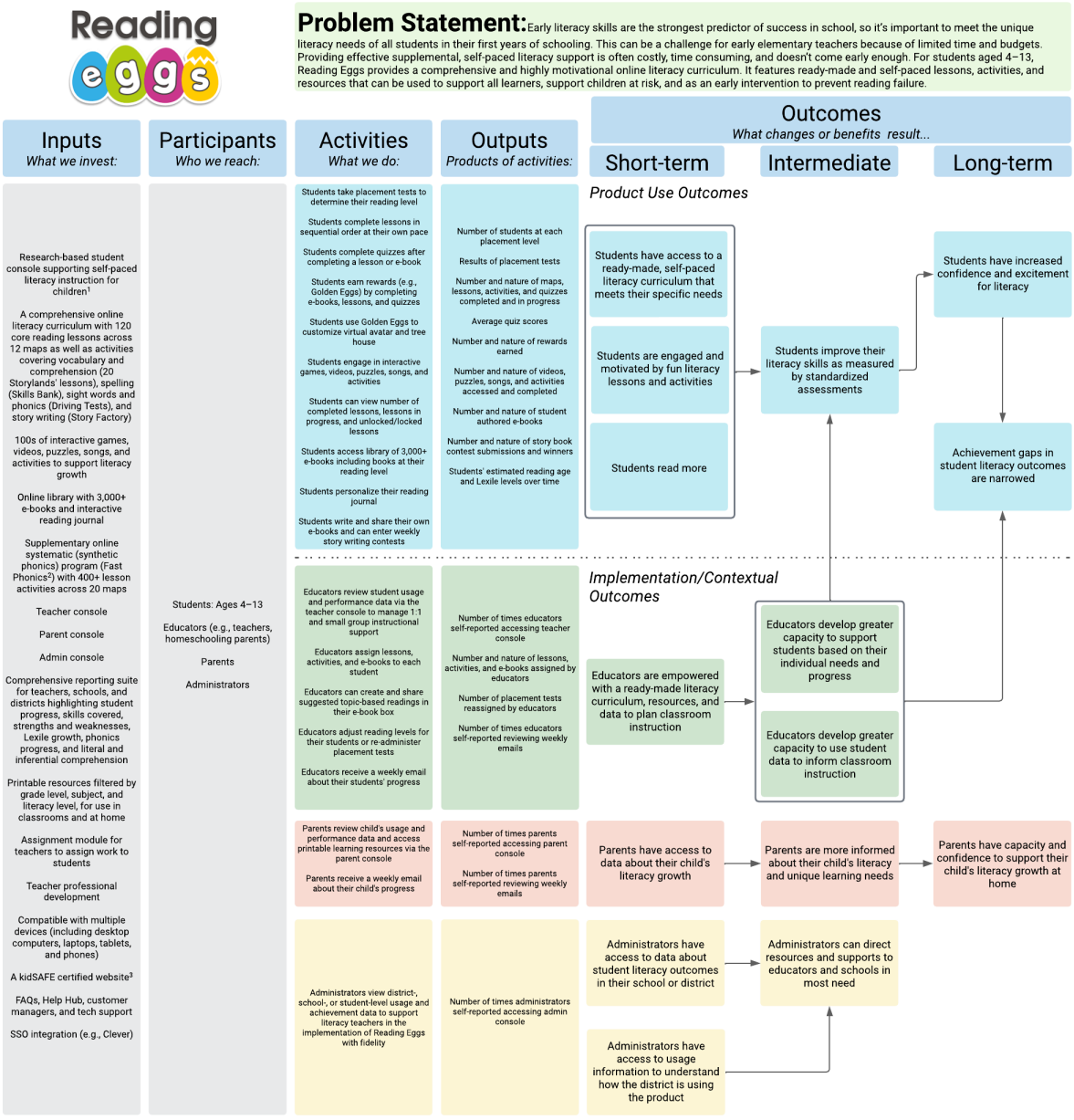
Researchers recommend the following next steps:

- 3P Learning should consider recruiting a comparison district for K–2 students to better understand how early elementary school students who use *Reading Eggs* compare to students using other reading or literacy programs.
- For the Grade 2 sample, it may be valuable to assess whether differences in implementation fidelity, usage patterns, or contextual factors (e.g., instructional time, curriculum alignment) are affecting outcomes for these students. Gathering more detailed qualitative or quantitative data on these factors could help clarify the findings.

REFERENCES

What Works Clearinghouse. (2022). What Works Clearinghouse procedures and standards handbook, version 5.0. U.S. Department of Education, Institute of Education Sciences, National Center for Education Evaluation and Regional Assistance (NCEE). This report is available on the What Works Clearinghouse website at URL

APPENDIX A. READING EGGS LOGIC MODEL



[1]3P Learning recommends that students use Reading Eggs for a minimum of 2 times a week for 20 minutes each session.
[2]Fast Phonics is an online systematic, synthetic phonics program designed for emergent and early readers as well as older students with gaps in their core reading knowledge. Fast Phonics is a supplementary component of Reading Eggs.
[3]A kidSAFE certified website means that the product has been independently reviewed, certified, and/or listed by kidSAFE to meet certain standards of online safety and/or privacy.

Prepared for 3P Learning, November, 2022

APPENDIX B. ADDITIONAL INFORMATION ON PROGRAM IMPLEMENTATION

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

Usage categories: weekly lessons		*n*	Mean	SD
Low	1–24 lessons	122	12	7
Moderate	25–53 weekly lessons	140	38	8
High	54–134 weekly lessons	65	77	20

Table B2. Descriptive statistics for the weekly lessons' usage categories for **Grade 1** sample

Usage categories: weekly lessons		*n*	Mean	SD
Low	1–24 lessons	151	13	7
Moderate	25–53 weekly lessons	132	36	8
High	54–134 weekly lessons	77	67	12

Table B3. Descriptive statistics for the weekly lessons' usage categories for **Grade 2** sample

Usage categories: weekly lessons		*n*	Mean	SD
Low	1–24 lessons	163	14	6
Moderate	25–53 weekly lessons	88	36	7
High	54–134 weekly lessons	14	61	8

APPENDIX C. ADDITIONAL INFORMATION ON AMIRA READING MASTERY OUTCOME FINDINGS FOR K–2 STUDENTS

Association Between Total Lessons Completed and K–2 Students' Outcomes on the Amira Reading Mastery Assessment by Usage Groups

Table C1. Association between **Kindergarten** *Reading Eggs* usage groups and spring 2024 ARM scores

Predictor	Unstd. Beta Coefficient	Standard Error	*t*	*p*-value
Moderate Use vs. Low Use (Hedges' *g* = **0.41*;** stdY = 0.388)	0.18	0.05	3.50	.001
High Use vs. Low Use (Hedges' *g* = **0.57*;** stdY = 0.549)	0.26	0.07	3.95	<.001
High Use vs. Moderate Use (Hedges' *g* = 0.16)	0.08	0.06	1.19	.235
Fall 2023 ARM scores	0.76	0.10	7.37	<.001

Table C2. Association between **Grade 1** *Reading Eggs* usage groups and spring 2024 ARM scores

Predictor	Unstd. Beta Coefficient	Standard Error	*t*	*p*-value
Moderate Use vs. Low Use (Hedges' *g* = -0.05)	-0.03	0.05	-0.67	.500
High Use vs. Low Use (Hedges' *g* = **0.28*;** stdY = 0.274)	0.18	0.06	3.01	.003
High Use vs. Moderate Use (Hedges' *g* = **0.35*;** stdY = 0.326)	0.21	0.06	3.49	.001
Fall 2023 ARM scores	0.73	0.03	21.61	<.001

Table C3. Association between **Grade 2** *Reading Eggs* usage groups and spring 2024 ARM scores

Predictor	Unstd. Beta Coefficient	Standard Error	*t*	*p*-value
Moderate Use vs. Low Use (Hedges' *g* = 0.05)	0.05	0.07	0.66	.509

Predictor	Unstd. Beta Coefficient	Standard Error	*t*	*p*-value
High Use vs. Low Use (Hedges' *g* = -0.12)	-0.12	0.14	-0.84	.400
High Use vs. Moderate Use (Hedges' *g* = 10.20)	-0.16	0.15	-1.09	.277
Fall 2023 ARM scores	1.17	0.05	24.51	<.001