



Learning Experience Design (LXD) Research & Consulting

a division of Charles River Media Group, LLC

95 Phonics Core Program™ Classroom Kit Grades K-3
Follow Up Efficacy Study with Acadience® Reading K-6: ESSA Level 1
Conducted by Rachel Schechter, Ph.D. and [Learning Experience Design \(LXD\) Research](#), with
statistical support from Lynch Research Associates
November 7, 2022

Introduction

There is a growing concern that core reading curricula for the elementary years have not improved reading scores in the US ([The Condition of Education 2020](#)). Reporters such as Emily Hanford ([APM Reports](#)) have shined a light specifically on the need for explicit, systematic, and sequential phonics instruction for every child. In response to this identified need, 95 Percent Group, LLC created [a new phonics core curriculum](#) that can replace the phonics instructional lessons provided with other core reading curricula (typically the first 20 minutes of the reading block).

The 95 Phonics Core Program (95PCP) is a whole-class, Tier I program designed for students in grades K-3 to address and prevent reading gaps using explicit, structured phonics instruction for 20 minutes per day. Instruction is based on a scope and sequence with 25 lessons for kindergarten and 30 lessons for each of Grades 1-3. The 95 Percent Group hired LXD Research, a third-party independent evaluator, to conduct a research study on the impact of the 95PCP.

Recognizing the importance of teachers' own experiences navigating this period of change in reading curricula, LXD Research centered teacher feedback in the research design. The treatment teachers' implementation stories, gathered through extensive and immersive interviews, focus groups and classroom observation, suggest that even though it took teachers some time to adjust to using the 95PCP, the teachers saw growth in their students' literacy skills which they attributed to the year of 95PCP lessons. The student assessment data then validated what the teachers observed firsthand in their own classrooms. The data from the initial study (Fall 2021 to Spring 2022) showed the 95PCP had a positive, significant impact on student achievement for all grades (K-2). This study follows up with those students to understand how skills learned during Year 1 impacted beginning of year scores in Year 2.

Research Description

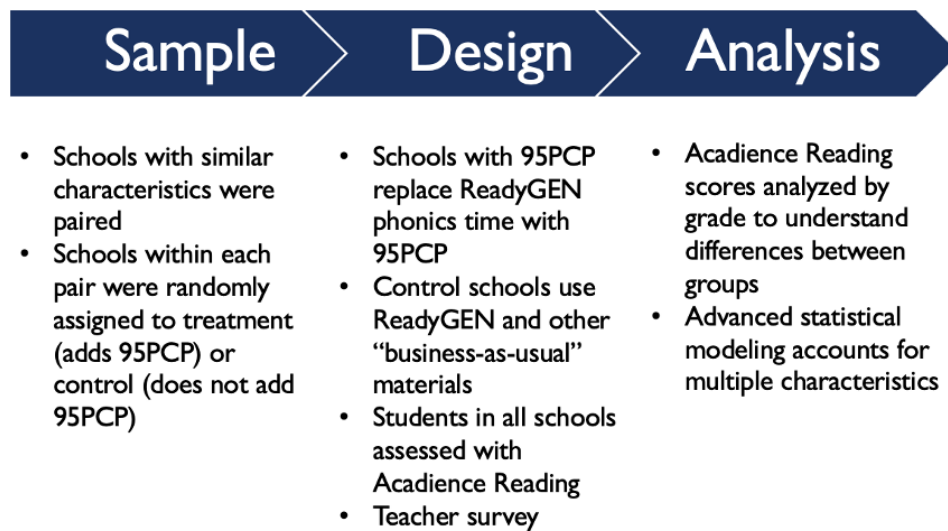
The research study has one primary goal: to examine the impact of the 95PCP on student literacy achievement in Grades K-3. Additional goals of the study are to examine program implementation information and feedback from educators about the program's quality and ease-of-use, as well as their

perceptions about the impact of the 95PCP on teaching and learning. To meet these goals, LXD Research designed a mixed-methods study with a random assignment of participants into conditions at the school level (ESSA Level 1 – Strong). This paper follows up on the first year of implementation with grades K-2.

The recruited research partner is in a large, suburban school district in Missouri. There are over 6,000 students in grades K-3 across 17 elementary schools in the district. The district serves a population in which 23% of students traditionally¹ qualified for free lunch and between 5-12% of students per grade are English Language Learners (ELL). LXD Research engaged with the district to conduct a set of qualitative and quantitative data collection activities throughout the school year: Acadience Reading K-6 assessment administration (at the beginning, middle, and end of the year), a teacher survey, interviews with school literacy coaches, and observations of classes.

Research Questions and Methods

Figure 1. Design Description



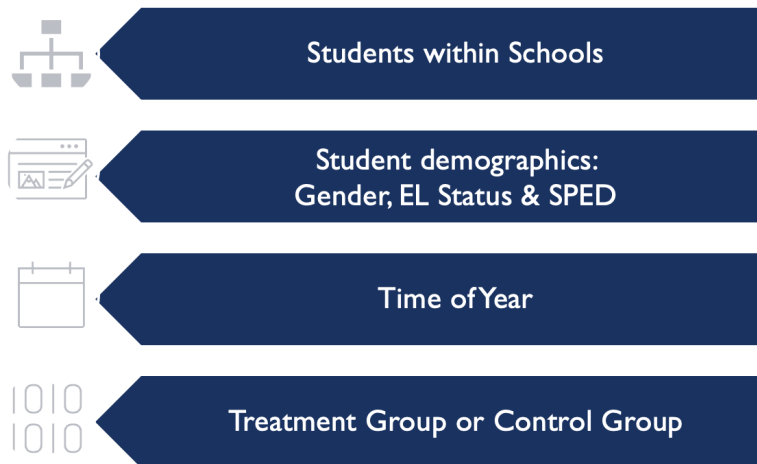
Research Questions for Follow-Up Study

- How does the 95PCP affect K-2 student achievement on formative assessments (of phonics, specifically) in schools that implement the program compared to schools that do not implement the program, comparing Fall 2021 to Fall 2022?
- How does the impact of the 95PCP vary by a student’s beginning of year (BOY) benchmark status (i.e., do students *Below Benchmark* at BOY achieve similar growth as students who were *At Benchmark* at BOY)?

¹ As of Fall 2020, all students receive free lunch, so documentation on free-lunch status is not available at the student level.

An advanced statistical modeling analysis was conducted to understand what the data show in support of answering these questions. This analytical technique helps account for known differences that could impact student outcomes.

Figure 2. Levels in Statistical Modeling



Sample Description

A total of 3443 students in grades 1-3 from 16 schools participated in an Randomized Control Trial (RCT) examining the effects of the 95 Percent Group’s Phonics Core Program in the Fort Zumwalt school district during the Fall of 2022. Of these students, 1883 were in the treatment group and 1560 were in the control group.

Among the 3403 students who had complete data from the Beginning of Year (BOY) in Fall of 2021, 344 students did not have BOY data available in Fall of 2022, signaling an attrition rate of approximately 10%. This attrition was equally likely to occur in the treatment and control groups ($\chi^2=0.63, p=.43$).

Matched BOY Fall 2021 and BOY Fall 2022 Sample

The following analyses include the 2677 students (see Figure 1 for a breakdown of students who were excluded from the initial sample of 3403) who had a Beginning of Year (BOY) assessment for Fall of 2021 and Fall 2022. Students are grouped into 3 grade levels: Rising First Grade (students in Kindergarten during Fall 2021 and currently in First Grade during Fall 2022), Rising Second Grade (students in First Grade during Fall 2021 and currently in Second Grade during Fall 2022), and Rising Third Grade (students in Second Grade during Fall 2021 and currently in Third Grade during Fall 2022). Table 1 describes the number of students in the matched sample by grade level and treatment or control group status.

Figure 3. Matched Fall 2021 and Fall 2022 Student Sample

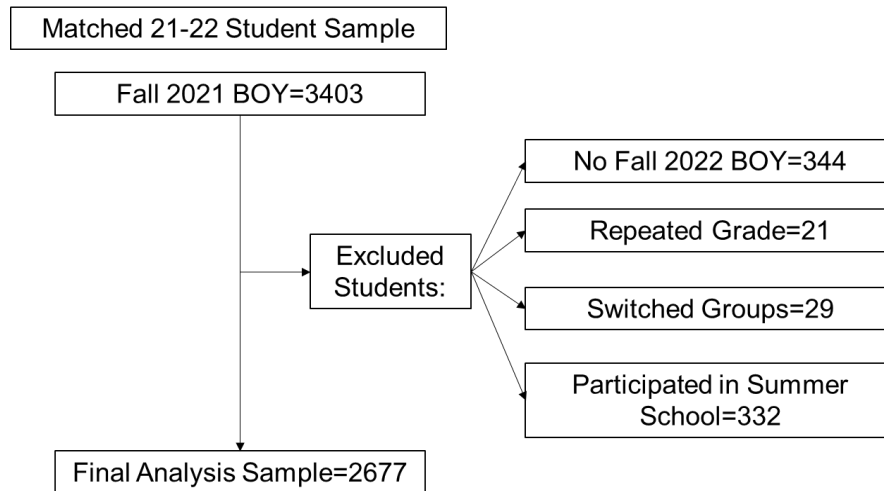


Table 1. Sample sizes for matched sample by grade level and treatment/control group

			BOY 2021	BOY 2022	Matched Sample
Grade Level	School Group	# of Schools	# of Students	# of Students	# of Students
Rising 1st	Control	7	516	470	397
	Treatment	7	612	569	466
	Total	14	1128	1039	863
Rising 2nd	Control	7	532	478	412
	Treatment	7	577	547	456
	Total	14	1109	1025	868
Rising 3rd	Control	7	535	493	419
	Treatment	7	631	611	527
	Total	14	1166	1104	946

Baseline Results

We employed Chi-Square analyses to compare students in the treatment (N=1449) and control (N=1228) groups in regard to gender, English Language Learner (ELL) and special education status (SPED). Results suggested there were no statistically meaningful differences between the treatment and control groups in regard to gender, ELL, and SPED (see Table 2).

Table 2. Sample descriptives for the treatment and control group for the Matched Sample (N=2,677)

Group	Male	SPED	ELL
Control	50%	11%	7%
Treatment	49%	12%	6%

When considering differences between the treatment and control groups by grade level, there were no statistically meaningful differences in regard to gender, SPED or ELL status among Rising 1st grade, Rising 2nd grade or Rising 3rd grade students (see Table 3).

Table 3. Sample descriptives for treatment and control group by grade for the Matched Sample

Grade	Group	Male	SPED	ELL
Rising 1st	Control	52%	8%	8%
	Treatment	48%	7%	5%
Rising 2nd	Control	48%	11%	8%
	Treatment	49%	11%	7%
Rising 3rd	Control	52%	14%	6%
	Treatment	50%	17%	5%

Analytic Approach

Three level hierarchical linear regression models (HLMs) with time (level 1) nested within students (level 2) nested with schools (level 3) were employed to examine growth in composite and subscale scores. All models contained a series of covariates including gender (“Gender”; 1=male, 0=female, ELL status (“ELL”; 1=ELL, 0=non-ELL), SPED status (“SPED”; 1=SPED, 0=non-SPED) an indicator of time (“Time”; 1=Beginning of year (BOY) Year 1, 2=BOY Year 2), an indicator of whether the student was in the treatment or control group (“group”; 1=Control, 2=Treatment), and an interaction between time and group calculated as the product of Time*group (“Tigr”). Table 4 lists the outcomes that were assessed within each age group.

We explored main effects of treatment vs control group by considering the significance of the interaction between time and group (“Tigr”). A significant interaction term would suggest that the slope (i.e., growth) in composite scores is different for the treatment versus control groups. We also

looked at growth in composite scores separately based on students' BOY benchmark scores. All analyses were conducted separately by grade using the statistical software package R 4.2.0.

Table 4. List of Assessments by Grade Level

Grade Level	BOY Y1-BOY Y2 Assessments
Rising 1st	Composite scores, LNF
Rising 2nd	Composite scores, CLS, WWR
Rising 3rd	Composite scores, ORF, ERR, RETELL, RETELL QR, ORF accuracy scores

Results for Student Reading Outcomes

After the first year of use, the data show that a positive impact of the 95 Phonics Core Program was measurable for all three grades.

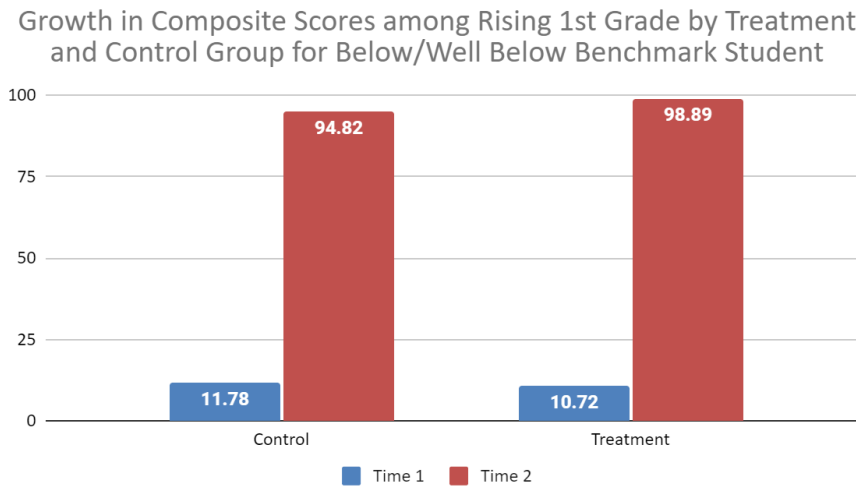
Positive impacts were found for the following Acadience Reading measures:

- Rising First Grade: Composite for students who started below grade level
- Rising Second Grade: Nonsense Word Fluency CLS (Correct Letter Sounds) and Whole Words Read (WWR)
- Rising Third Grade: Composite

Rising First Grade

Because the distribution of the composite and LNF scores were positively skewed within the Rising First grade sample, we elected to use a poisson distribution to examine changes in scores over time. There was not a significant effect of treatment on composite or LNF scores, suggesting that students in the treatment and control group demonstrated similar growth. We looked separately at growth in composite scores among students who were below or well below benchmark at baseline. Among students who were below or well below benchmark at BOY, students in the treatment group demonstrated more growth in composite scores than students in the control group (IRR=1.15, $p < .001$, $f^2 = .00$; see Figure 1a). Complete output for each model can be found in [Appendix 1](#). Results of t-tests (and their associated effect sizes) comparing growth in composite scores between the treatment and control groups can be found in [Appendix 2](#).

Figure 4. Below or or Well Below benchmark students: Rising First Grade students in the treatment group demonstrated significantly more growth in composite scores than students in the control group



Rising Second Grade

Within the Rising Second Grade sample, there was a significant effect of treatment on CLS scores and WWR scores; students in the treatment group demonstrated more growth in CLS scores than students in the control group ($B=6.44$, $p=.001$, $f^2=.01$; see Figure 2a) and students in the treatment group demonstrated more growth in WWR scores than students in the control group ($B=2.30$, $p=.005$, $f^2=.01$; see Figure 2b). There were no statistically meaningful differences between treatment and control group in regard to growth in composite scores. There were also no significant findings in regard to benchmark status. That is, students tended to demonstrate similar growth in composite scores regardless of benchmark status. Complete output for each model can be found in [Appendix 3](#).

Figure 5a. Rising Second Grade students in the treatment group demonstrated significantly more growth in CLS scores than students in the control group

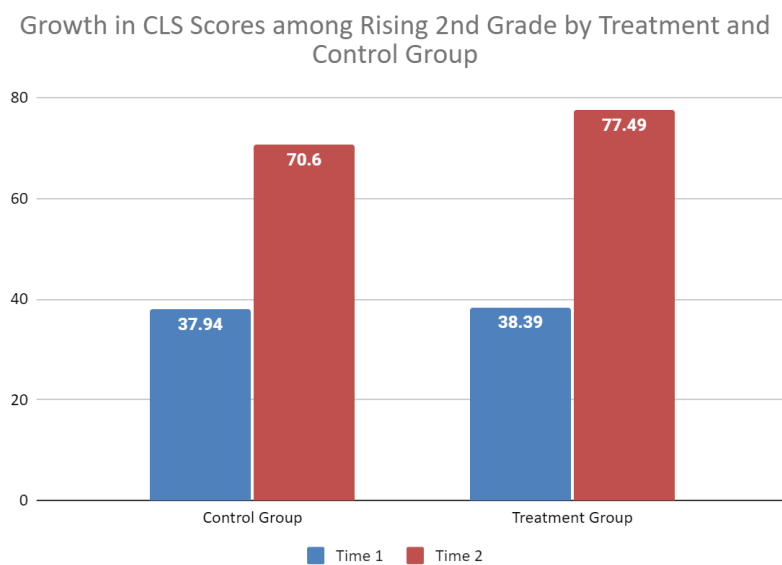
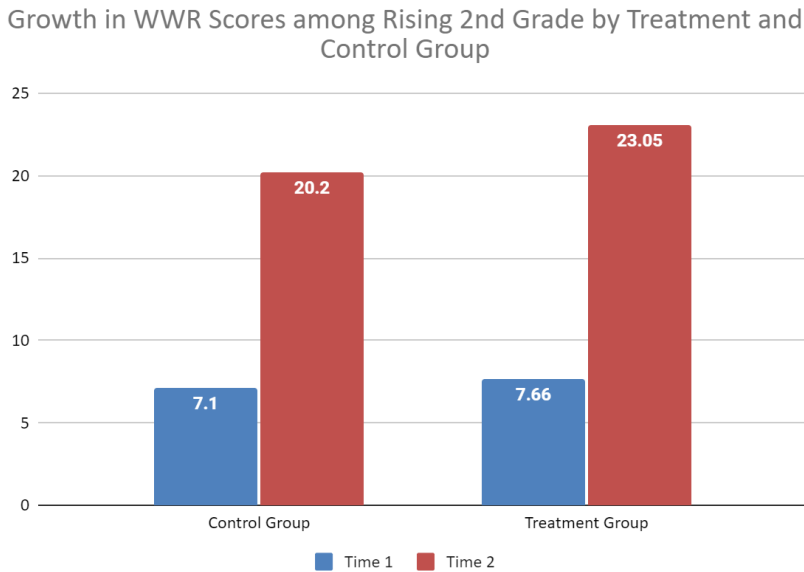


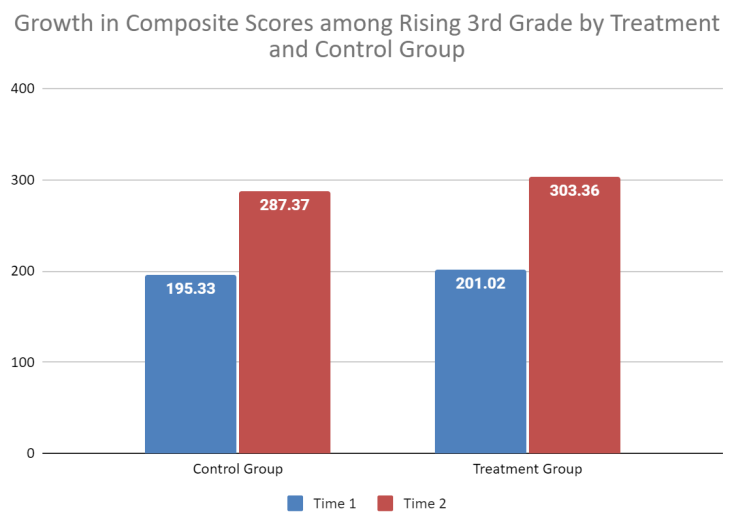
Figure 5b. Rising Second Grade students in the treatment group demonstrated significantly more growth in WWR scores than students in the control group



Rising Third Grade

Within the Rising Third Grade sample, there was a significant effect of treatment on composite scores ($B=10.31$, $p=.009$, $f^2=.01$; see Figure 3a). Students in the treatment group demonstrated more growth in composite scores than students in the control group. There were also no significant findings in regard to benchmark status. That is, students tended to demonstrate similar growth in composite scores regardless of benchmark status. There were no statistically meaningful differences between treatment and control group in regard to growth in ORF, ERR, RETELL, RETELL QR, and ORF Accuracy scores. Complete output for each model can be found in [Appendix 4](#).

Figure 6. Rising Third Grade students in the treatment group demonstrated significantly more growth in composite scores than students in the control group



Conclusion

Over the course of the first year, teachers shared that the change to a new educational program was challenging. To change the way a school teaches reading by using structured phonics can be overwhelming for teachers and learners. The research team heard exactly that from the teachers in the treatment group in this study as the teachers worked to simultaneously learn the content and cadence of 95PCP and teach its lessons, but this was only half of the story. The initial struggle to learn a new way of teaching early literacy was overshadowed by their students' reading growth and the sense, expressed by the comparison group teachers, that 95PCP met an urgent need in their curricular toolbox. 95PCP facilitated alignment between their own burgeoning knowledge of the science of reading (SOR) and their teaching tools.

Educators from this district continued to share stories with the research team after the first year of the study ended. They said that it was clear during summer school which students had received the 95PCP and which students had not. At the start of the year, teachers shared that they had never seen their students so prepared for the start of school. The student assessment data validated what the teachers observed firsthand in their own classrooms. The data show the first year of the 95PCP had a positive, significant impact on student achievement that extended past the summer and into the new school year. There were no noticeable differences in how the program impacted students from different subgroups. For rising first graders, even students who started Below Benchmark continued to see an extra boost in growth from this core supplemental program (that lasted over the summer).

- *Rising first graders in the treatment group demonstrated significantly more growth in Composite scores than students in the control group when grouped by BOY Benchmark Status.*
- *Rising second graders in the treatment group demonstrated more growth in CLS and WWR scores than students in the control group.*
- *Rising third graders in the treatment group demonstrated significantly more growth in Composite scores than students in the control group.*

This is the first of multiple inquiries on how the initial gains of the product sustain and build over multiple years of use. Additionally, new research questions may explore the extent to which teachers' growing knowledge of phonics, whether that be gained through a Science of Reading program such as LETRS or independent study, affects teachers' implementation and fidelity to the use of 95PCP will be incorporated into the year 2 study in future reports. Finally, additional studies that examine the use of a core supplemental program in combination with a high-quality, structured Tier 2 and/or Tier 3 intervention program could support acceleration and growth for all students.

Appendices

Appendix 1: Rising First Grade Results

- Composite score: (IRR=0.99, p=.67) - no significant differences between treatment and control group
- LNF score: (IRR=0.97, p=.12) - no significant differences between treatment and control group

For below or well below students:

- Composite score: (IRR=1.15, p<.001) - significant differences between treatment and control group
-

Composite Score

<i>Predictors</i>	comp 1		
	<i>Incidence Rate Ratios</i>	<i>CI</i>	<i>p</i>
(Intercept)	8.97	7.53 – 10.68	<0.001
Time	3.57	3.42 – 3.72	<0.001
y2female	0.97	0.92 – 1.03	0.348
y2SPED	0.52	0.46 – 0.58	<0.001
y2ELL	0.73	0.64 – 0.82	<0.001
y2group	1.05	0.94 – 1.17	0.370
y2Tigr	0.99	0.97 – 1.02	0.673
Random Effects			
σ^2	0.01		
τ_{00} student_id:y2school_name	0.19		
τ_{00} y2school_name	0.01		
ICC	0.93		
N _{student_id}	863		
N _{y2school_name}	14		
Observations	1725		
Marginal R ² / Conditional R ²	0.680 / 0.978		

Below or Well Below Benchmark Comparisons

The variable of interest is “Tigr,” which represents the interaction between “Time” and “Group” and tells us whether growth in the outcome is different for students in the control versus treatment groups.

<i>Predictors</i>	comp 1		
	<i>Incidence Rate Ratios</i>	<i>CI</i>	<i>p</i>
(Intercept)	1.85	1.41 – 2.43	<0.001
Time	7.08	6.38 – 7.85	<0.001
y2female	0.99	0.89 – 1.09	0.826
y2SPED	0.54	0.46 – 0.63	<0.001
y2ELL	0.74	0.63 – 0.88	0.001
y2group	0.79	0.67 – 0.93	0.006
y2Tigr	1.15	1.07 – 1.22	<0.001
Random Effects			
σ^2	0.02		
τ_{00} student_id:y2school_name	0.21		
τ_{00} y2school_name	0.00		
ICC	0.91		
N _{student_id}	358		
N _{y2school_name}	14		
Observations	715		
Marginal R ² / Conditional R ²	0.839 / 0.986		

<i>Predictors</i>	lnf 1		
	<i>Incidence Rate Ratios</i>	<i>CI</i>	<i>p</i>
(Intercept)	9.12	7.55 – 11.01	<0.001
Time	2.15	2.03 – 2.28	<0.001
y2female	0.96	0.90 – 1.01	0.135
y2SPED	0.64	0.57 – 0.72	<0.001
y2ELL	0.75	0.66 – 0.85	<0.001
y2group	1.08	0.96 – 1.21	0.205
y2Tigr	0.97	0.94 – 1.01	0.117
Random Effects			
σ^2	0.03		
τ_{00} student_id:y2school_name	0.18		
τ_{00} y2school_name	0.01		
ICC	0.85		
$N_{\text{student_id}}$	863		
$N_{\text{y2school_name}}$	14		
Observations	1726		
Marginal R^2 / Conditional R^2	0.409 / 0.911		

Appendix 2: Effect Sizes Based on t-tests

In the table below we report effect sizes (Cohen's d) resulting from dependent samples t-test that compared growth in composite scores in the treatment and control groups. T-tests were run separately for Rising 1st, Rising 2nd and Rising 3rd grade students.

Grade	Condition	Number of students	Average difference in Composite between Fall 2021 and Fall 2022	SD	Significance	Effect Size Cohen's d
Rising 1st	Control	396	84.77	36.37	p=.11	0.11
	Treatment	466	88.64	34.94		
Rising 2nd	Control	412	75.31	64.41	p=.07	0.12
	Treatment	456	83.46	67.22		
Rising 3rd	Control	419	92.05	58.00	p=.010	0.17
	Treatment	525	102.26	62.03		

Appendix 3: Rising Second Grade Results

- Composite score: (B=8.15, p=.07) - no significant differences between treatment and control group
- CLS score: (B=6.44, p=.001) - significant differences between treatment and control group
- WWR score: (B=2.30, p=.005) - significant differences between treatment and control group

For below or well below students:

- Composite score: (B=10.78, p=.07) - no significant differences between treatment and control group
-

Composite Score

<i>Predictors</i>	<i>Estimates</i>	comp 2	
		<i>CI</i>	<i>p</i>
(Intercept)	50.69	22.67 – 78.71	<0.001
Time	67.16	53.06 – 81.27	<0.001
y2female	-10.65	-17.83 – -3.48	0.004
y2SPED	-60.99	-72.63 – -49.36	<0.001
y2ELL	-26.55	-40.29 – -12.80	<0.001
y2group	-7.23	-24.54 – 10.08	0.413
y2Tigr	8.15	-0.64 – 16.94	0.069
Random Effects			
σ^2	2171.88		
τ_{00} student_id:y2school_name	1731.63		
τ_{00} y2school_name	67.28		
ICC	0.45		
$N_{\text{student_id}}$	868		
$N_{\text{y2school_name}}$	14		
Observations	1736		
Marginal R ² / Conditional R ²	0.334 / 0.636		

Below or Well Below Benchmark Comparisons

The variable of interest is “Tigr,” which represents the interaction between “Time” and “Group” and tells us whether growth in the outcome is different for students in the control versus treatment groups.

<i>Predictors</i>	<i>Estimates</i>	comp 2 <i>CI</i>	<i>p</i>
(Intercept)	40.84	6.46 – 75.22	0.020
Time	53.20	34.17 – 72.23	<0.001
y2female	-4.76	-12.27 – 2.76	0.215
y2SPED	-54.96	-65.85 – -44.07	<0.001
y2ELL	-18.16	-31.25 – -5.07	0.007
y2group	-10.62	-31.77 – 10.53	0.325
y2Tigr	10.78	-1.02 – 22.58	0.073
Random Effects			
σ^2	2402.23		
τ_{00} student_id:y2school_name	692.38		
τ_{00} y2school_name	69.10		
ICC	0.24		
N _{student_id}	534		
N _{y2school_name}	14		
Observations	1068		
Marginal R ² / Conditional R ²	0.337 / 0.497		

CLS Scores

<i>Predictors</i>	<i>Estimates</i>	cls 2	
		<i>CI</i>	<i>p</i>
(Intercept)	11.28	-1.09 – 23.64	0.074
Time	26.23	19.98 – 32.48	< 0.001
y2female	-8.00	-11.35 – -4.65	< 0.001
y2SPED	-22.02	-27.46 – -16.58	< 0.001
y2ELL	-6.40	-12.83 – 0.02	0.051
y2group	-6.00	-13.62 – 1.63	0.123
y2Tigr	6.44	2.55 – 10.33	0.001
Random Effects			
σ^2	426.79		
τ_{00} student_id:y2school_name	402.74		
τ_{00} y2school_name	11.53		
ICC	0.49		
$N_{\text{student_id}}$	868		
$N_{\text{y2school_name}}$	14		
Observations	1736		
Marginal R^2 / Conditional R^2	0.316 / 0.653		

WWR Scores

<i>Predictors</i>	<i>Estimates</i>	wwr 2	
		<i>CI</i>	<i>p</i>
(Intercept)	-4.28	-9.36 – 0.81	0.099
Time	10.81	8.25 – 13.37	<0.001
y2female	-2.80	-4.03 – -1.56	<0.001
y2SPED	-6.96	-8.96 – -4.96	<0.001
y2ELL	-2.68	-5.04 – -0.31	0.026
y2group	-1.73	-4.88 – 1.41	0.280
y2Tigr	2.30	0.70 – 3.89	0.005
Random Effects			
σ^2	71.74		
τ_{00} student_id:y2school_name	47.30		
τ_{00} y2school_name	2.39		
ICC	0.41		
$N_{\text{student_id}}$	868		
$N_{\text{y2school_name}}$	14		
Observations	1736		
Marginal R^2 / Conditional R^2	0.325 / 0.601		

Appendix 4: Rising Third Grade Results

- Composite score: (B=10.31, p=.009) - significant differences between treatment and control group
- ORF score: (B=0.97, p=.45) - no significant differences between treatment and control group
- ERR score: (IRR=0.92, p=.06) - no significant differences between treatment and control group
- RETELL score: (B=-1.24, p=.21) - no significant differences between treatment and control group
- RETELL QR score: (B=0.06, p=.41) - no significant differences between treatment and control group
- ORF Accuracy score: (B=-0.74, p=0.46) - no significant differences between treatment and control group

For below or well below students:

- Composite score: (B=10.97, p=.18) - no significant differences between treatment and control group

Composite Score

<i>Predictors</i>	<i>Estimates</i>	comp 3	
		<i>CI</i>	<i>p</i>
(Intercept)	107.91	74.20 – 141.62	<0.001
Time	81.74	69.09 – 94.39	<0.001
y2female	-14.41	-24.89 – -3.92	0.007
y2SPED	-95.76	-110.30 – -81.23	<0.001
y2ELL	-35.59	-58.73 – -12.45	0.003
y2group	-4.61	-25.21 – 15.99	0.661
y2Tigr	10.31	2.56 – 18.05	0.009
Random Effects			
σ^2	1816.57		
τ_{00} student_id:y2school_name	5680.98		
τ_{00} y2school_name	160.40		
ICC	0.76		
$N_{\text{student_id}}$	946		
$N_{\text{y2school_name}}$	14		
Observations	1890		
Marginal R^2 / Conditional R^2	0.324 / 0.840		

Below or Well Below Benchmark Comparisons

The variable of interest is “Tigr,” which represents the interaction between “Time” and “Group” and tells us whether growth in the outcome is different for students in the control versus treatment groups.

<i>Predictors</i>	<i>Estimates</i>	comp 3	
		<i>CI</i>	<i>p</i>
(Intercept)	51.46	0.06 – 102.86	0.050
Time	67.94	42.12 – 93.77	<0.001
y2female	-15.97	-29.65 – -2.30	0.022
y2SPED	-82.67	-97.35 – -67.98	<0.001
y2ELL	-21.33	-45.76 – 3.10	0.087
y2group	-14.28	-45.71 – 17.15	0.373
y2Tigr	10.97	-5.04 – 26.98	0.179
Random Effects			
σ^2	2252.33		
τ_{00} student_id:y2school_name	1989.04		
τ_{00} y2school_name	195.10		
ICC	0.49		
$N_{\text{student_id}}$	274		
$N_{\text{y2school_name}}$	14		
Observations	546		
Marginal R^2 / Conditional R^2	0.423 / 0.707		

ORF Scores

<i>Predictors</i>	<i>Estimates</i>	orf 3	
		<i>CI</i>	<i>p</i>
(Intercept)	48.36	36.64 – 60.07	<0.001
Time	21.62	17.46 – 25.77	<0.001
y2female	-4.96	-9.16 – -0.77	0.020
y2SPED	-30.62	-36.42 – -24.82	<0.001
y2ELL	-10.33	-19.57 – -1.10	0.028
y2group	1.74	-5.37 – 8.85	0.631
y2Tigr	0.97	-1.57 – 3.52	0.454
Random Effects			
σ^2	196.34		
τ_{00} student_id:y2school_name	956.39		
τ_{00} y2school_name	16.37		
ICC	0.83		
$N_{\text{student_id}}$	946		
$N_{\text{y2school_name}}$	14		
Observations	1892		
Marginal R^2 / Conditional R^2	0.183 / 0.863		

ERR Scores

err 3			
<i>Predictors</i>	<i>Incidence Rate Ratios</i>	<i>CI</i>	<i>p</i>
(Intercept)	3.17	2.40 – 4.19	<0.001
Time	1.05	0.91 – 1.21	0.510
y2female	1.08	0.99 – 1.18	0.086
y2SPED	1.45	1.28 – 1.63	<0.001
y2ELL	1.21	1.00 – 1.47	0.052
y2group	1.03	0.87 – 1.22	0.692
y2Tigr	0.92	0.84 – 1.00	0.064
Random Effects			
σ^2	0.27		
τ_{00} student_id:y2school_name	0.33		
τ_{00} y2school_name	0.00		
ICC	0.55		
$N_{\text{student_id}}$	946		
$N_{\text{y2school_name}}$	14		
Observations	1892		
Marginal R^2 / Conditional R^2	0.038 / 0.569		

RETELL Scores

retell 3			
<i>Predictors</i>	<i>Estimates</i>	<i>CI</i>	<i>p</i>
(Intercept)	11.02	4.54 – 17.50	0.001
Time	10.54	7.41 – 13.66	<0.001
y2female	0.79	-0.87 – 2.45	0.352
y2SPED	-12.19	-14.51 – -9.87	<0.001
y2ELL	-7.08	-10.78 – -3.39	<0.001
y2group	3.16	-0.81 – 7.13	0.119
y2Tigr	-1.24	-3.15 – 0.68	0.205
Random Effects			
σ^2	109.65		
τ_{00} student_id:y2school_name	108.87		
τ_{00} y2school_name	4.30		
ICC	0.51		
$N_{\text{student_id}}$	945		
$N_{\text{y2school_name}}$	14		
Observations	1872		
Marginal R^2 / Conditional R^2	0.157 / 0.585		

RETELL QR Scores

retellqr 3			
<i>Predictors</i>	<i>Estimates</i>	<i>CI</i>	<i>p</i>
(Intercept)	1.89	1.42 – 2.36	< 0.001
Time	0.12	-0.11 – 0.35	0.310
y2female	0.02	-0.08 – 0.12	0.745
y2SPED	-0.54	-0.69 – -0.39	< 0.001
y2ELL	-0.29	-0.52 – -0.05	0.016
y2group	0.02	-0.27 – 0.31	0.890
y2Tigr	0.06	-0.08 – 0.20	0.413
Random Effects			
σ^2	0.56		
τ_{00} student_id:y2school_name	0.30		
τ_{00} y2school_name	0.02		
ICC	0.36		
$N_{\text{student_id}}$	925		
$N_{\text{y2school_name}}$	14		
Observations	1777		
Marginal R^2 / Conditional R^2	0.052 / 0.397		

ORF Accuracy scores

<i>Predictors</i>	<i>Estimates</i>	acc 3	
		<i>CI</i>	<i>p</i>
(Intercept)	82.74	76.02 – 89.47	<0.001
Time	7.10	3.92 – 10.29	<0.001
y2female	-3.46	-5.36 – -1.55	<0.001
y2SPED	-23.00	-25.63 – -20.36	<0.001
y2ELL	-8.68	-12.88 – -4.49	<0.001
y2group	2.61	-1.49 – 6.71	0.212
y2Tigr	-0.74	-2.69 – 1.21	0.457
Random Effects			
σ^2	115.40		
τ_{00} student_id:y2school_name	159.69		
τ_{00} y2school_name	4.13		
ICC	0.59		
$N_{\text{student_id}}$	946		
$N_{\text{y2school_name}}$	14		
Observations	1892		
Marginal R^2 / Conditional R^2	0.225 / 0.680		