95 PHONICS CORE PROGRAM™
2021-2022, GRADES K-2 EFFICACY STUDY
JUNE 22, 2022

LXD RESEARCH
95 PERCENT GROUP LLC

95 PHONICS CORE PROGRAM™

EVIDENCE for ESSA STRONG

2021/2022 RESULTS - GRADES K-2

PROGRAM DESCRIPTION

The 95 Phonics Core Program (95PCP) adds an explicit phonics strand to the daily Reading Block to ensure that all K-3 students receive consistent evidence- and research-based phonics instruction to improve outcomes. The program is designed to accommodate both in-person and remote learning.

SCHOOL DESCRIPTION

LOCATION: Missouri

GRADES: K-2

SIZE: 3,257 Students

DEMOGRAPHICS: 75% White | 23% Low Income | 9% EL | 12% SPED

STUDY DETAILS

Sixteen schools in the district were paired by Spring 2021 ELA scores and then **randomly assigned** to treatment and control groups.

CONTROL GROUP

TIER 1 LITERACY CURRICULUM
Core: ReadyGEN
Supplemental:
- **95 Multisyllable Routine Cards**
- Reading A-Z
- Heggerty Bridge the Gap
- Phonics First and other materials

TREATMENT GROUP

CHANGES TO LITERACY TOOLKIT
Added: **95 Phonics Core Program**
Removed: ReadyGEN phonics strand

ASSESSMENT

ACADIENCE READING
Acadience Reading K-6 was conducted at the beginning and end of the 2021-2022 school year.

Analysis used a three-level hierarchical linear regression model, (students in classrooms) controlling for gender, ELL, and SPED. See the executive summary for more details.

COMPARING RESULTS

From Fall 2021 to Spring 2022, students using **95 Phonics Core Program showed higher gains on Acadience Reading than the control group** in all three grades.

While 2014-15 study on ReadyGEN showed growth for first graders on Terranova assessments, it **did not have a control group** to provide ESSA-level evidence for any grades K-2.

Grade	ReadyGEN	95PCP
Kindergarten	○	●
First	◐	●
Second	○	●

○ No evidence   ◐ Limited evidence   ● Strong evidence

EDUCATOR VOICES

"I think it's giving everybody a **common language** and I've just seen a lot of growth in my struggling readers this year with spelling."

"We're ending with **reading levels almost higher than I've ever seen.** They're applying or learning to books or it's coming through in their writing as well."

"**There are certain things they love!** They love the word sorts, you know, they love the pictures, and they love the word chains. Oh my gosh. My kids love the word chains."
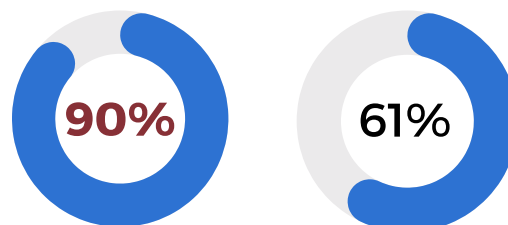
Observers would notice...

SKILL APPLICATION

73%

25%
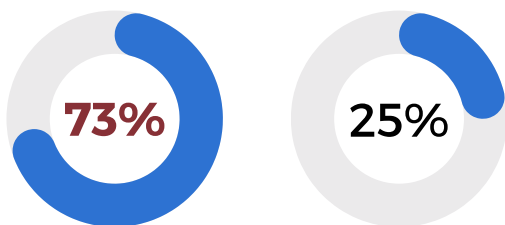
95 PERCENT GROUP   **COMPARISON**

95 Percent Group teachers lead students in skill application during 73% of the lesson compared to only 25% in the comparison classrooms.
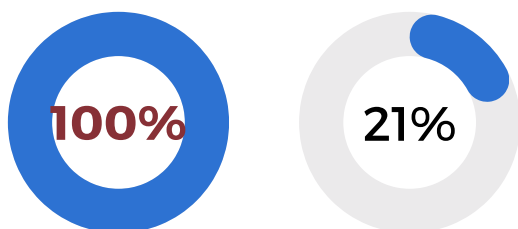
STUDENT ENGAGEMENT

90%

61%

95 PERCENT GROUP   **COMPARISON**

95 Percent Group classrooms had most of the students on task during the lesson compared to 61% in the comparison classrooms.

LESSON CLARITY

100%

21%

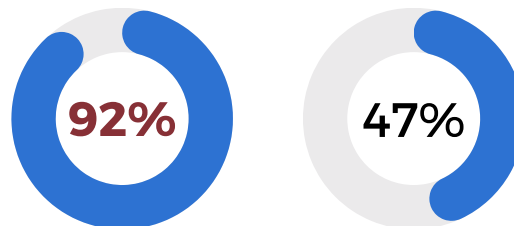**95 PERCENT GROUP**   **COMPARISON**

95 Percent Group had 100% of classrooms where the lesson number was clear compared to 21% in the comparison classrooms.

STUDENT INDEPENDENCE

92%

47%

95 PERCENT GROUP   **COMPARISON**

95 Percent Group had 92% of classrooms where most of the students prepared without help for each lesson segment compared to 47% in the comparison classrooms.

QUOTES FROM LITERACY COACHES

READYGEN ONLY

"There's not a whole lot of background [phonics] information for the teacher, and there's not a whole lot to pull from as far as like a script or anything like that. So when you think about like the Science of Reading and teaching students how to learn, how to read, it's kind of like figuring it out. So it's been a little bit of a struggle."

READYGEN + 95PCP

"The Phonics Core Program has really been a great supplement because ReadyGEN was definitely lacking in [phonics] and it's what the kids and the teachers needed. That's been the missing link for sure."

RESULTS FROM RIGOROUS STATISTICAL MODELS

Models accounted for known differences that could impact outcomes (statistical controls).

Classrooms

Gender, EL & SPED

SOR Training

BOY Benchmark Level

Adding 95 Phonics Core Program boosted gains for all grades.

| Grade	Gains on Composite	Gains on Subtests
Kindergarten	⊕	⊕
First	→	⊕
Second	⊕	⊕

→ Similar gains ⊕ 95PCP higher gains

A CLOSER LOOK AT STUDENTS WHO WERE BELOW BENCHMARK IN FALL 2021

A strong Tier 1 phonics program benefits all students and may particularly benefit students who start the school year Below Benchmark on their universal screener. These students may or may not qualify for Tier 2 supports and therefore not receive differentiated instructional time or materials due to resource constraints. Systematic, explicit instruction could help boost students who are Below Benchmark into At/Above Benchmark categories after just a few months of whole-group lessons. This page examines this group of students, along with the At/Above students (who should remain in this category), from each grade in the study.

Results from replacing ReadyGEN phonics with 95PCP can be noticed through the percentage of students who advance to At/Above Benchmark.

(The % who remained At/Above Benchmark is included in graphs below.)

KINDERGARTEN
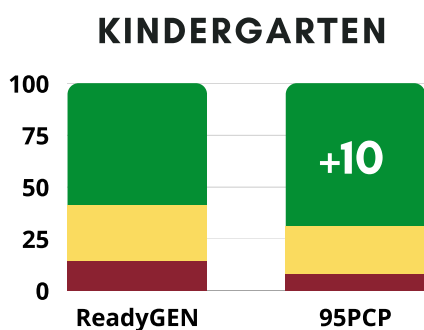
Composite Level at EOY (Overall Reading)
- 🟢 At/Above
- 🟡 Below
- 🔴 Well Below

FIRST GRADE

SECOND GRADE
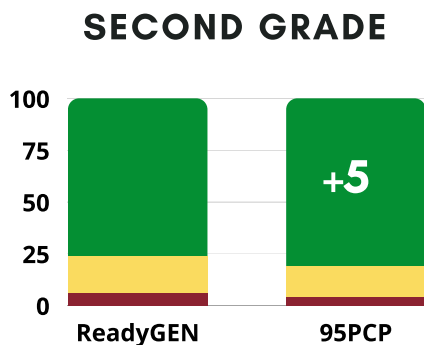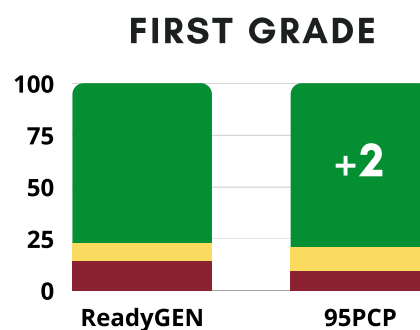
Percent of Students At/Above Benchmark at the End of the Year

Grade	Comparison	Treatment
K	59%	69%
1	77%	79%
2	76%	81%

% out of all students Below or At/Above Benchmark at BOY

ANALYSIS DESCRIPTION

Fall 2022 scores were compared to Fall 2021 to understand how gains from the first year were sustained over the summer. (Students who attended summer school were excluded.) This final assessment period allows for the comparison of 2,677 students using 95 Phonics Core Program to the control group using Acadience Reading for all three grades, now described as Rising 1st, Rising 2nd, and Rising 3rd.

Results from adding 95PCP to ReadyGEN are noticed using composite score gains and percent of students At/Above Benchmark.

In terms of gains, all grades earned at least 5 additional points on the reading composite score from Fall 2021 to Fall 2022 using 95PCP.

In terms of having more students At/Above Benchmark at the start of the year, all grades with 95PCP outperformed other schools. Rising 1st reduced the proportion of students ready for first grade, while 95PCP schools stayed steady.

Gains on Composite Scores from Fall 2021 to Fall 2022
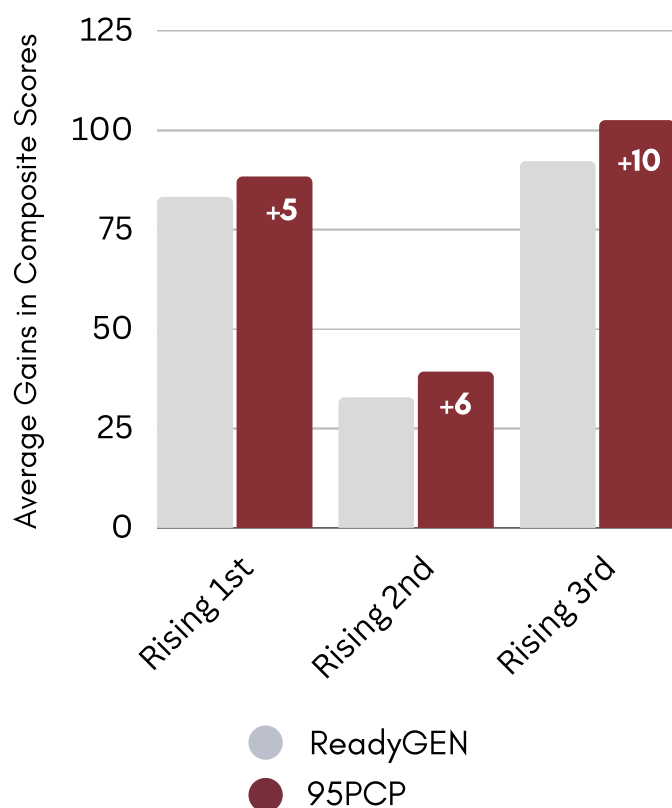
Change in % of Students On/Above Benchmark from Fall 2021 to Fall 2022

95 Phonics Core Program™ Classroom Kit Grades K-3
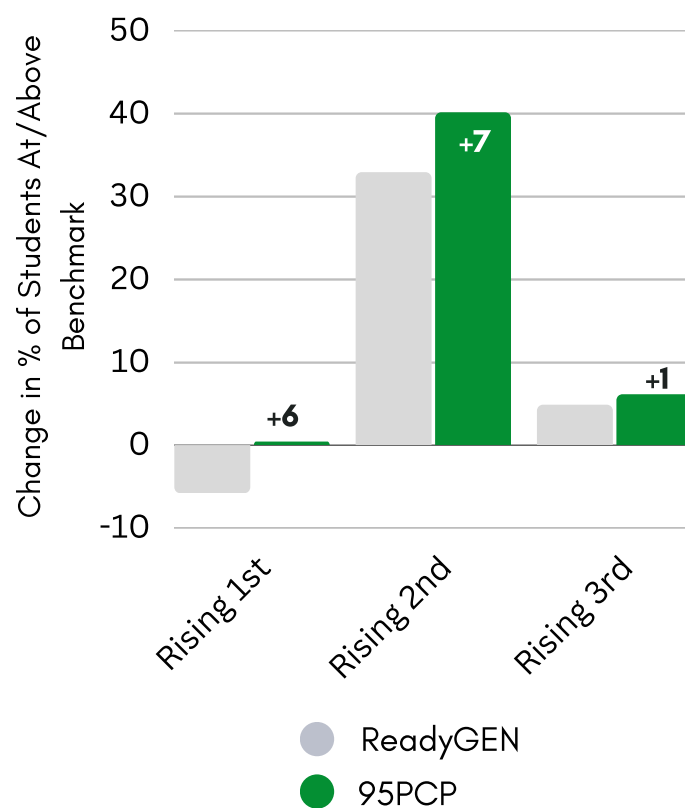
Efficacy Study with Acadience® Reading K-6: Level 1 ESSA Level of Evidence
Conducted by Rachel Schechter, Ph.D. and LXD Research at Charles River Media Group

Executive Summary

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 shone 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.

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 random selection of participants into conditions at the school level (ESSA Level 1 – Strong). This paper focuses on grades K-2 for the first semester of implementation only.

The recruited research partner is in a majority-White school district in Missouri. There are over 6,000 students in grades K-5 across 17 elementary schools in the district. The district serves a population in which 23% of students traditionally[1] qualified for free lunch and between 5-12% of students per grade are English Language Learners (described later as "EL status"). LXD Research is engaging with the district to conduct a set of qualitative and quantitative data collection activities, of which two main activities were completed in Fall and Winter 2021: Acadience Reading K-6 assessment administration

---

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

(at the Beginning of Year, BOY, and the Middle of Year, MOY) and a teacher survey with control schools (the teacher survey for treatment schools will be administered in the Spring of 2022).

Research Questions and Methods
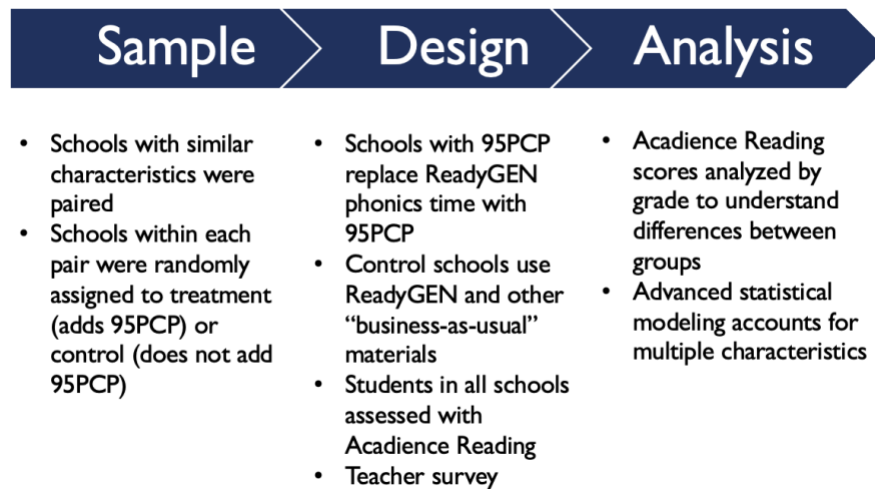
Sample > **Design** > **Analysis**

Sample
- Schools with similar characteristics were paired
- Schools within each pair were randomly assigned to treatment (adds 95PCP) or control (does not add 95PCP)

Design
- Schools with 95PCP replace ReadyGEN phonics time with 95PCP
- Control schools use ReadyGEN and other "business-as-usual" materials
- Students in all schools assessed with Acadience Reading
- Teacher survey

Analysis
- Acadience Reading scores analyzed by grade to understand differences between groups
- Advanced statistical modeling accounts for multiple characteristics

Research Questions

- 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 by MOY?
- How does the impact of the 95PCP vary by grade and student subgroup (gender, English Language Learner status, students in special education)?
- How does the impact of the 95PCP vary by a student's BOY benchmark status (i.e., do students *Below Benchmark* at BOY achieve similar growth as students who were *At Benchmark* at BOY)?
- What, if any, impact does teacher training in the Science of Reading have on student scores?

An advanced statistical modeling analysis was conducted to understand what the data show in support of answering these questions. These techniques help account for known differences that could impact student outcomes.
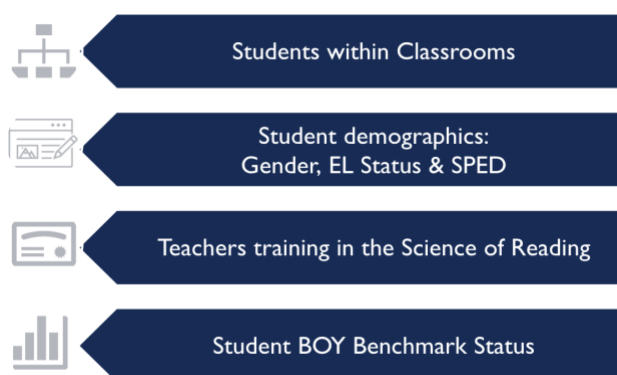
Students within Classrooms

Student demographics: Gender, EL Status & SPED

Teachers training in the Science of Reading

Student BOY Benchmark Status

After only 14 weeks 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:

- Grade K: Composite, FSF (First Sound Fluency), and LNF (Letter Naming Fluency)
- First Grade: Nonsense Word Fluency CLS (Correct Letter Sounds) and WWR (Whole Words Read)
- Second Grade: Composite and ORF (Oral Reading Fluency)

Kindergarten: Treatment group showed more growth from BOY to MOY than the Control group in Composite Scores
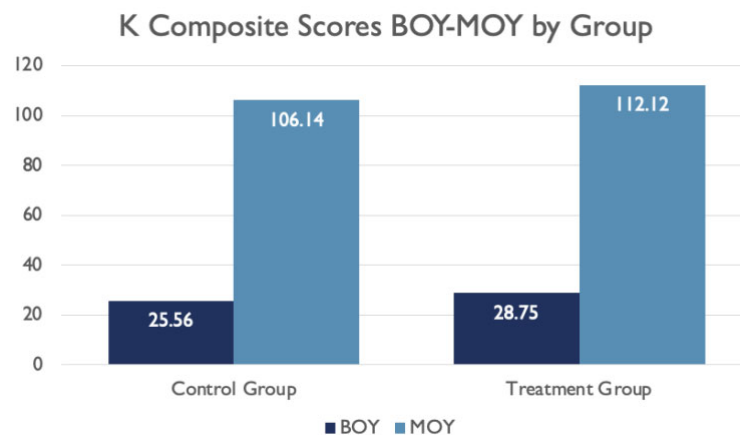
Kindergarten: Treatment group showed more growth from BOY to MOY than the Control group in FSF (First Sound Fluency) and LNF (Letter Naming Fluency) Scores
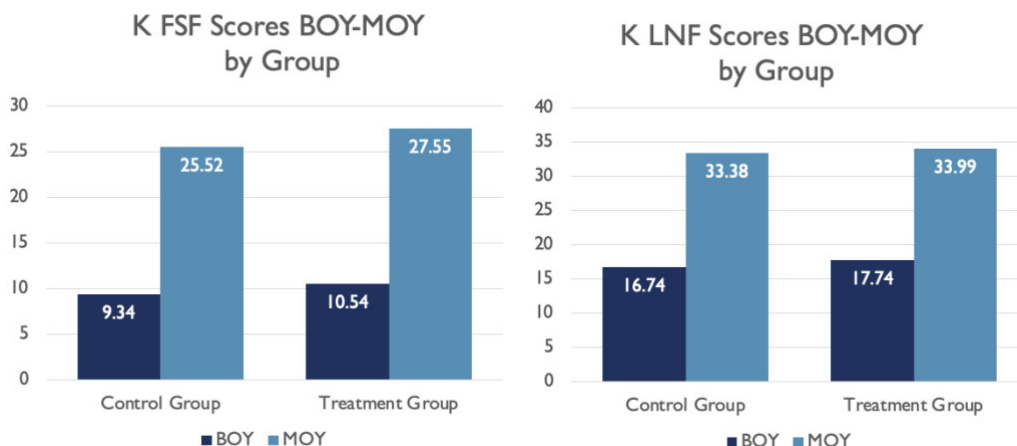
First Grade: Treatment group showed more growth from BOY to MOY than the Control group in Nonsense Word Fluency CLS (Correct Letter Sounds) and WWR (Whole Words Read) Scores

First Grade CLS Scores BOY-MOY by Group

	Control Group	Treatment Group
BOY	33.51	32.9
MOY	53.84	55.73

First Grade WWR Scores BOY-MOY by Group

	Control Group	Treatment Group
BOY	5.82	5.89
MOY	13.2	14.65
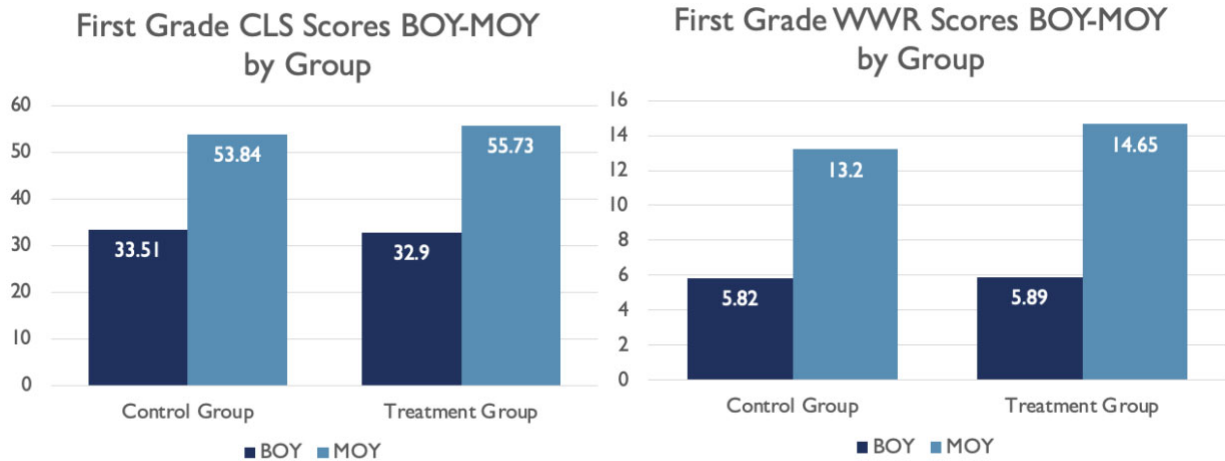
Second Grade: Treatment group showed more growth from BOY to MOY than the Control group in Composite and ORF (Oral Reading Fluency) Scores

Second Grade Composite Scores BOY-MOY by Group

	Control Group	Treatment Group
BOY	178.29	187.11
MOY	232.2	248.19

Second Grade ORF Scores BOY-MOY by Group

	Control Group	Treatment Group
BOY	66.19	71.32
MOY	83.11	92.17
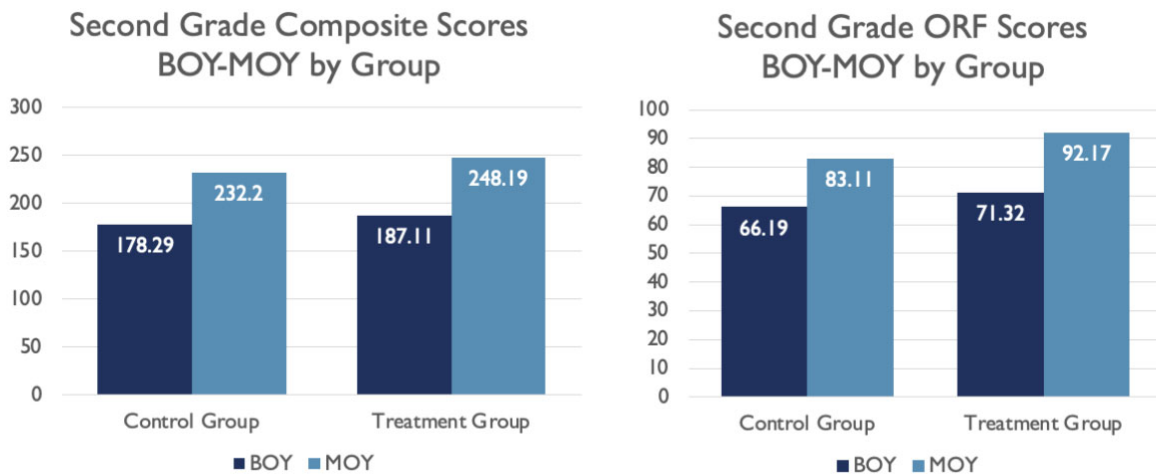
Conclusion & Next Steps

While the 95PCP is intended to be used all year, the positive impacts during the Fall semester (14 weeks) are encouraging and support the ESSA Level 1 evidence for the 95PCP's effectiveness. The analysis of the MOY Acadience data collection begins to answer the central questions of this research study. In particular, the following data analysis findings provide reasonable and positive answers to the research questions in advance of our EOY data collection:

- 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 by MOY?
  - The data show a positive, significant impact on student Acadience scores for all grades (K-2) on multiple subtests, as well as on the Composite score for grades K and 2.
- How does the impact of the 95PCP vary by grade and student subgroup (gender, English Language Learners, students in special education)?
  - Students performed similarly regardless of gender, EL status, or SPED status.
- How does the impact of the 95PCP vary by a student's BOY benchmark status (i.e., do students *Below Benchmark* at BOY achieve similar growth as students who were *At Benchmark* at BOY)?
  - Students performed similarly regardless of BOY benchmark status. In other words, students who were *Below Benchmark* or *Well Below Benchmark* made the same amount of progress as students who started the year *At or Above Benchmark*.
- What, if any, impact does teacher training in the Science of Reading (SOR) have on student scores?
  - The impact of teacher SOR training is not straightforward. For kindergarten, the training seemed to support student growth; in first grade the data show students of teachers in the training made fewer gains; and in second grade there was no apparent impact. Reasonable hypotheses about why first graders made fewer gains may include the impact of time scarcity for teachers simultaneously participating in the SOR course and implementing a new phonics program. There was no detectable impact of teacher training on the learning of students in the control schools.

As the study continues during the Winter of 2022, Administrator interviews will allow for a better understanding of how phonics instruction and intervention has progressed over the course of the first half of the year. These more in-depth conversations will provide additional context from both treatment and control schools on what instruction looks like across grades and within each building. The Spring data collection includes focus groups with educators and the end-of-year (EOY) Acadience data collection.

Table of Contents

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 shone 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. For example, the First Grade Scope and Sequence includes 30 lessons disaggregated into seven topics (introduction, short vowel CVC, consonant blends, consonant digraphs, long vowel silent-e, phonograms, and introduction to second-grade skills). Each lesson focuses on specific phonics skills, provides examples of high-frequency words, and contains information about other skills addressed within the topic. The 95 Percent Group offers a kit for each grade, including a teacher's edition, student workbooks, manipulatives, and a digital presentation. The 95PCP may be offered in-person or virtually. The 95PCP also aligns with assessments and interventions (such as Phonics Lesson Library) offered by 95 Percent Group to ensure consistency.

95 Percent Group partnered with LXD Research to conduct a third-party evaluation of the 95PCP as it was implemented during the 2021-2022 school year in a Missouri school district. All the elementary schools use ReadyGEN as a core reading curriculum and half were randomly selected to use the 95PCP for phonics instruction instead of the ReadyGEN word study materials. Random assignment to conditions ensures the highest level of scientific rigor (ESSA Evidence Level 1).

Evaluation Questions

The evaluation aims to answer the following questions:
1. How does the 95PCP affect student achievement on formative assessments (of phonics, specifically) in schools that implement the program compared to schools that do not implement the program?
2. How does the impact of the 95PCP vary by school, grade, and student subgroup (gender, English Language Learners [EL] status, students in special education [SPED] status)?

3. What is the nature and extent of the 95PCP implementation?
    a. How is the 95PCP typically implemented?
    b. To what extent is the 95PCP implemented with fidelity, and does the program adhere to the Theory of Action?
    c. How do contextual factors affect 95PCP implementation, such as the content and quality of professional development and the characteristics of districts and schools, such as the level of administrator support?
4. What is the nature and extent of literacy program implementation in comparison schools?
5. What are teacher and administrator perceptions about the quality and impact of the 95PCP?
    d. What are teacher and administrator initial reactions to the 95PCP and its associated materials, content, pacing, and professional development?
    e. What suggestions do they have for improvement?
6. What is the association between variations in 95PCP implementation and student outcomes?

Methods

This study uses a mixed-method design that includes quantitative and qualitative data collection. The Fall 2021 and Winter 2022 research activities included the beginning-of-year (BOY) reading assessment, a middle-of-year (MOY) assessment, and a survey of teachers from control schools that focused on phonics instructional strategies and background.

Design

This study uses a mixed-methods approach, including a randomized experimental design complemented by classroom observations, teacher surveys, and administrator interviews. This combination of methods allows researchers to understand how the materials are being used in the classroom, collect teacher feedback on the quality and perceived impact of the program, and evaluate student academic achievement.

School districts were recruited in Spring 2021. In exchange for their participation, district leaders received all 95PCP materials and training at no cost and discounts for any 95PCP materials purchased in the 2022-2023 school year. The control schools used the regular materials that they have used in previous years. Prior to the 2021-2022 school year, the district leaders allowed for the randomization of schools to a treatment (95PCP) or control condition. Schools were organized into pairs using school size and ELA scores from Fall 2020 and Spring 2021, and then a coin toss determined which school in each pair would receive the 95PCP. Students were pretested within the first four weeks of school using Acadience Reading (BOY: September 1-17[2]) and were tested again between December 13-17, 2021 (middle of the year, or MOY, after at least 12 weeks of instruction). The final testing period will be at the end of the year (EOY) in Spring 2022.

Treatment Group: Program Key Features

The 95PCP features instructional practices that differ from the typical reading instruction provided by core curricula. A phonemic awareness and phonics continuum of skills is followed using structured literacy characteristics, described in Table 1.

---

[2] Four students in first grade were tested Monday and Tuesday of the following week (September 20 and 21), due to absences.

Table 1. 95 Percent Group's Literacy Characteristics in 95PCP Lessons

Characteristic	Evident in Lesson Framework
1. Explicit	I Do directly states and defines focus skill and student expectations.
2. Systematic	Intentional language and steps include consistent hand gestures and verbal cues; there is a gradual transfer of responsibility from teacher to student.
3. Sequential	Structure moves from simple to complex in key ways including lesson order, word choice, materials used, and teacher talk.
4. Adequate Modeling	This most prominent feature provides precise language at each level of modeling.
5. Corrective Feedback	Teacher response is reactive to individual student errors.
6. Differentiated Instruction	We Do and You Do sections provide two levels that enable teachers to differentiate instruction to meet students' needs.
7. Scaffolded Instruction	Steps of the I Do, We Do, and You Do allow the teacher to gradually transfer responsibility for learning to the students.
8. Continual Assessment	This occurs through informal observation and monitoring during instruction; the focus skill correlates to the *PSI*.

The 95 Percent Group's version of the gradual release model (Table 2) allows all students to practice every skill using multisensory materials, including a phonics mat and chips. While a paper version of the Phonics Chip Kit is included in the 95PCP, a plastic version is available and sold separately.

Table 2. Gradual Release Model in 95PCP

Modeling Steps	Chip Movement	Speaking
I Do	teacher	teacher
We Do		
Level 1: Accuracy	teacher	teacher and students
Level 2: Fluency	teacher	students
You Do	students	students

The 95PCP phonological awareness and phonics continua are shown in Figures 1 and 2, respectively. There is a clear progression from simpler to more complex skills, following the research-based developmental progression for learning to read. The International Dyslexia Association, for example, describes structured literacy as "Systematic means that organization of material follows the logical order of language. The sequence begins with the easiest and most basic concepts and elements and progresses methodically to the more difficult."

Figure 1. Phonological Awareness Continuum of 95 Percent Group
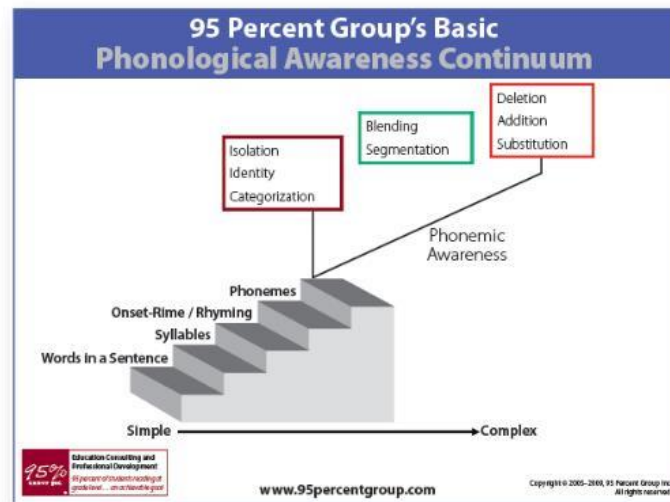
Figure 2. Phonics Continuum of Skills of 95 Percent Group
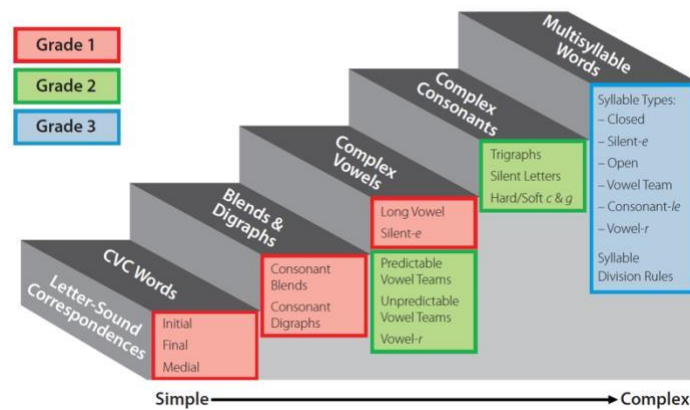
Control Group: Phonics Instruction

The district uses ReadyGEN for their core reading program, which is published by Savvas Learning Company (formerly Pearson). This curriculum has one published study that meets the Level 3 (Promising) ESSA criteria for first grade using the Terranova 3 assessment. The program is described as using the Gradual Release of Responsibility Model, a generative approach to vocabulary instruction, and many language-focused, text-based strategies for teaching reading and writing. The curriculum also includes assessments and online games.

Control Group: Survey Summary

A survey conducted in the Fall of 2021 collected information from the teachers in the control schools about their approach to teaching phonics across all tiers. A total of 83 teachers who teach K-3, the focus grades for the 95PCP, responded; the respondents included at least one representative from each school. Nearly all teachers use ReadyGEN to teach phonics, with Phonics First being the other consistently mentioned program. More than 25% of teachers indicated that ReadyGEN didn't have phonics instruction, or they didn't know if it did. The most-used supplemental phonics program was Reading A-Z, with Heggerty Phonemic Awareness and Phonics First following behind. A small group of teachers mentioned using 95 Percent Group's Multisyllable Routine Cards (Second Grade only), Heggerty Bridge the Gap, Raz-Kids, and Leveled Literacy Intervention materials. The amount of time allocated for phonics instruction varied widely within schools and grades, ranging from no time to more than 25 minutes per day. A survey of the teachers in the treatment schools will be conducted in Spring 2022.

Assessment: Acadience Reading K-6

Acadience Reading is an assessment that helps teachers identify children at risk for reading difficulties and determine the skills to target for instructional support. Acadience assessments are standardized and assess core early literacy skills (Table 3). Because the subtests and their weighting change for each assessment period (see [Acadience User Manual](#)), Composite scores are used to compare reading ability in this report.

Table 3. Acadience Reading Subtests and Skill Coverage

Subtest	Indicators of These Basic Early Literacy Skills
First Sound Fluency (FSF) & Phoneme Segmentation Fluency (PSF)	Phonemic Awareness
Letter Naming Fluency (LNF)	Indicator of risk
Nonsense Word Fluency (NWF)	Alphabetic Principle and Basic Phonics (Correct Letter Sounds and Whole Words Read)
Oral Reading Fluency (ORF) & Retell Fluency (RTF)	Advanced Phonics and Word Attack Skills, Accurate and Fluent Reading of Text (ORF Words Correct Per Minute and Errors); Reading Comprehension (RTF Total and Quality of Response)
Maze	Reading Comprehension

Sample

The 95PCP is being implemented in a majority-White school district in Missouri. The district serves a population in which 23% of students traditionally[3] qualified for free lunch. A total of 3,327 K-2 students in 178 classrooms across 14 elementary schools were included in this analysis, further described in Table 4.

Table 4. Number of Students, Classes, and Schools by Grade and Condition

Grade Level	School Group	# of Classes	BOY # of Students	MOY # of Students	Matched Sample # of Students
K	Treatment	32	612	599	598
	Control	27	517	497	497
	Total	59	1129	1096	1095
1	Treatment	30	578	567	565
	Control	29	532	519	519
	Total	59	1110	1086	1084
2	Treatment	33	631	624	623
	Control	27	536	527	525
	Total	60	1167	1151	1148

Of these students, 1,786 were in the treatment group and 1,541 were in the control group. Students in the treatment and control groups were similar with respect to gender and SPED status. However, students in the control group were more likely to be ELLs (B=-.03, p < .001; see Table 5*).

Table 5. Demographic descriptions for treatment and control group

Group	Male	SPED	ELL*
Control	49%	14%	10%
Treatment	50%	13%	7%

Acadience Reading Beginning-of-Year Scores

The random assignment of schools successfully created similar treatment and control groups in each grade. The significance level for each pair was greater than 0.05, and effect sizes (Hedges' g) were 0.11 or lower for all grades (Table 6).

---

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

Table 6. Acadience Composite Score Results for Beginning of Year (all students with BOY)

Grade	Condition	Number of students	Treatment avg	SD	Significance	Effect Size Cohen's d
K	Treatment	612	31.69	23.98	p=.06	.12
	Control	517	29.04	23.78		
1st grade	Treatment	578	99.24	42.21	p=.81	.01
	Control	532	99.85	40.21		
2nd grade	Treatment	631	162.66	87.59	p=.04	.06
	Control	534	152.51	85.29		

Analytic Approach

This report focuses on exploring the following research questions:

- 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 by MOY?
- How does the impact of the 95PCP vary by school, grade, and student subgroup (gender, EL, and SPED status)?
- How does the impact of 95PCP vary by a student's BOY benchmark status (i.e., do students *Below Benchmark* at BOY achieve similar growth as students who were *At Benchmark* at BOY)?
- What, if any, impact does teacher training in the Science of Reading have on student scores?

To answer these questions, three-level hierarchical linear regression models (HLMs) with time (level 1) nested within students (level 2) nested with classrooms (level 3) were employed to examine growth in Acadience Reading 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=BOY, 2=MOY), 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").

We explored main effects of treatment versus 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 status. Finally, we considered whether formal teacher training ("sort"; 0=No Training, 1=1+ Years Training) in the Science of Reading (not a 95 Percent Group program) moderated growth in Composite scores. All analyses were conducted separately by grade using the statistical software package R 3.6.2.

Results

Kindergarten

Within the Kindergarten grade sample, we examined growth in Composite scores as well as growth in First Sound Fluency (FSF) and Letter Naming Fluency (LNF) scores. Because the scores were highly negatively skewed counts, we elected to use a Poisson distribution to examine changes in scores over time. There was a significant effect of treatment on Composite scores (IRR=.98, p=.044), FSF scores (IRR=.96, p=.032), and LNF scores (IRR=-.97, p=.039) Figures 3a-c show these effects graphically. There were no significant findings regarding benchmark status. That is, students tended to demonstrate similar growth in scores regardless of their benchmark status at BOY. While there was no significant effect of teacher training within the control group, there was a significant effect of teacher training within the treatment group (IRR=.94, p=.005), suggesting that treatment students whose teachers had training in the Science of Reading tended to demonstrate more growth in Composite scores than treatment students whose teachers had no training (Figure 3d). Complete output for each model can be found in Appendix 1.

Figure 3a. Kindergarten students in the treatment group demonstrated significantly more growth in Composite scores than students in the control group
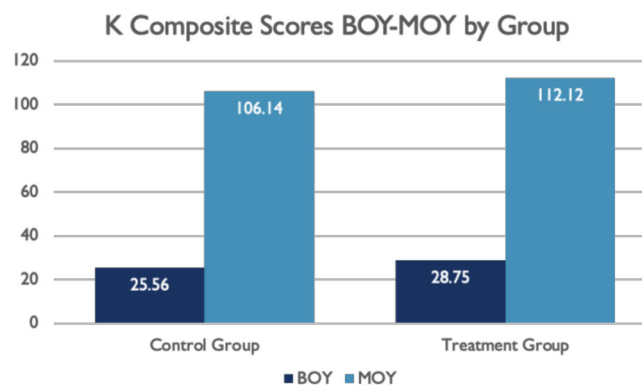
Figure 3b. Kindergarten students in the treatment group demonstrated significantly more growth in FSF scores than students in the control group
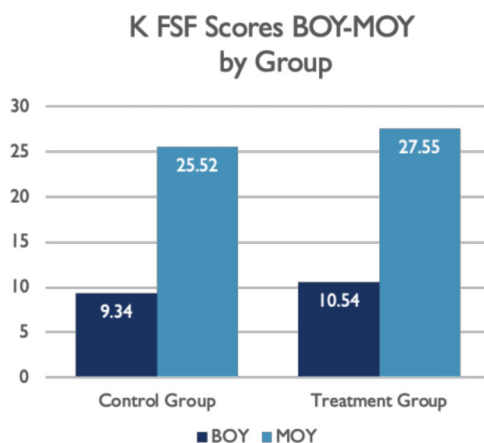
Figure 3c. Kindergarten students in the treatment group demonstrated significantly more growth in LNF scores than students in the control group
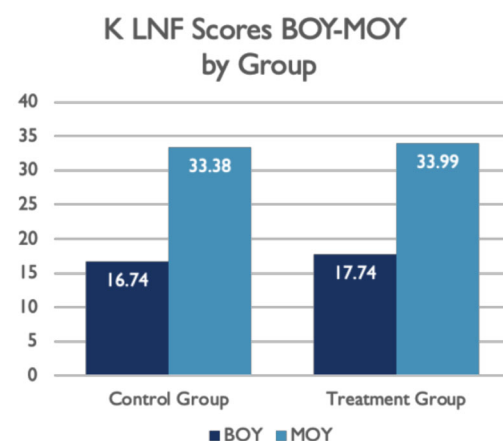
Figure 3d. Kindergarten treatment group students whose teachers had some training in the Science of Reading demonstrated significantly more growth in Composite scores than students whose teachers had no training

K Composite Scores, PCP Group, by Teacher SOR Training

Chart showing:
- No SOR Training: BOY 26.58, MOY 105.64
- 1+ Classes SOR Training: BOY 36.09, MOY 128

Legend: BOY, MOY

First Grade

Within the First-Grade sample, we examined growth in Composite scores as well as growth in Nonsense Word Fluency Correct Letter Sound (CLS) and Whole Words Read (WWR) scores. Students in the treatment group demonstrated significantly more growth in CLS (B=2.50, p=.047) and WWR (B=1.38, p=.019) scores than students in the control group, as shown in Figures 4a and 4b, respectively. There were no statistically meaningful differences between the treatment and control groups regarding Composite scores. There were also no significant findings regarding benchmark status. That is, students tended to demonstrate similar growth in scores regardless of their benchmark status at BOY. While there was no significant effect of teacher training within the control group, there was a significant effect of teacher training within the treatment group (B=-14.06, p=.025), suggesting that treatment group students of teachers who had training in the Science of Reading tended to demonstrate less growth in Composite scores than those whose teachers had no training (Figure 4c). Complete output for each model can be found in Appendix 2.

Figure 4a. First Grade students in the treatment group demonstrated significantly more growth in CLS scores than students in the control group

First Grade CLS Scores BOY-MOY by Group

Chart showing:
- Control Group: BOY 33.51, MOY 53.84
- Treatment Group: BOY 32.9, MOY 55.73

Legend: BOY, MOY

Figure 4b. First Grade students in the treatment group demonstrated significantly more growth in WWR scores than students in the control group

First Grade WWR Scores BOY-MOY by Group

Chart showing:
- Control Group: BOY 5.82, MOY 13.2
- Treatment Group: BOY 5.89, MOY 14.65

Legend: BOY, MOY

Figure 4c. Treatment group students whose teachers had some training in the Science of Reading demonstrated significantly less growth in Composite scores than students whose teachers had no training

Grade 1 Composite Scores, PCP Group, by Teacher SOR Training

No SOR Training: BOY 111.77, MOY 177.69
1+ Classes SOR Training: BOY 94.87, MOY 132.67

■ BOY ■ MOY

Second Grade

Within the Second-Grade sample, we examined growth in Composite scores as well as improvement in Oral Reading Fluency Words Correct Per Minute (ORF) scores, Oral Reading Fluency Error (ERR) scores, Retell Total (RETELL) scores, and Retell Quality (RETELLQR) scores. Students in the treatment group demonstrated significantly more growth in Composite scores (B=7.17, p=.024, f^2=.01) and OFR scores (B=3.93, p < .001, f^2=.01) than students in the control group (Figures 5a-b). There were no statistically meaningful differences between the treatment and control groups regarding ERR, RETELL, or RETELLQR scores. There were also no significant findings regarding benchmark scores or teacher training. That is, students tended to demonstrate similar growth in scores regardless of their benchmark status or whether their teacher had SOR training. Complete output for each model can be found in Appendix 3.

Figure 5a. Second Grade students in the treatment group demonstrated significantly more growth in Composite scores than students in the control group
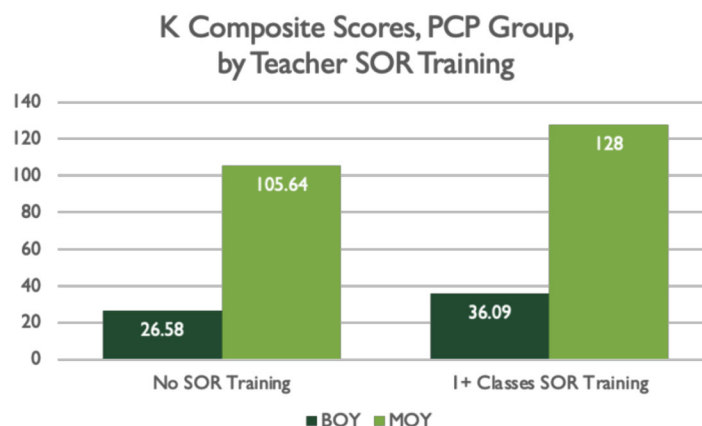
Second Grade Composite Scores BOY-MOY by Group

Control Group: BOY 178.29, MOY 232.2
Treatment Group: BOY 187.11, MOY 248.19

■ BOY ■ MOY
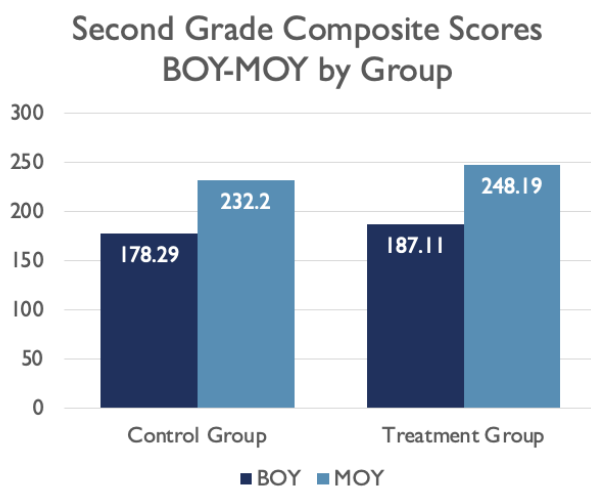
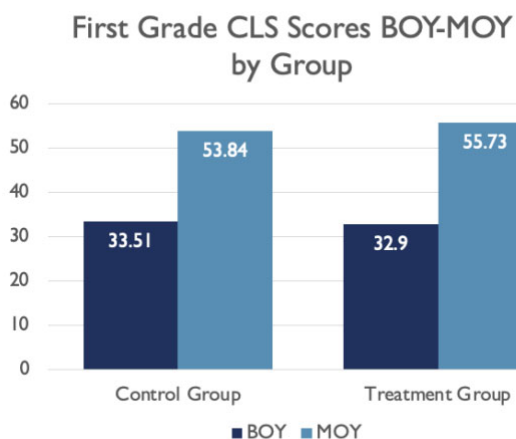*Figure 5b. Second Grade students in the treatment group demonstrated significantly more growth in ORF scores than students in the control group*

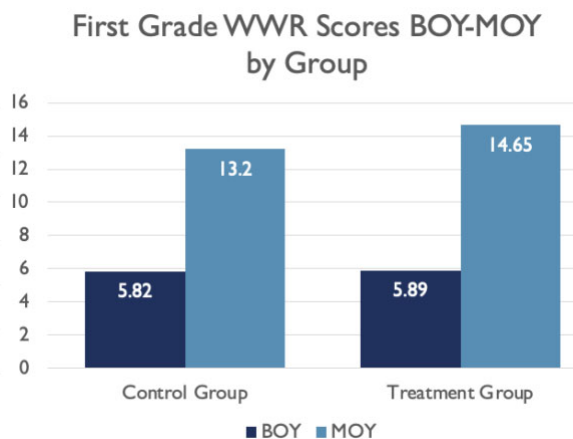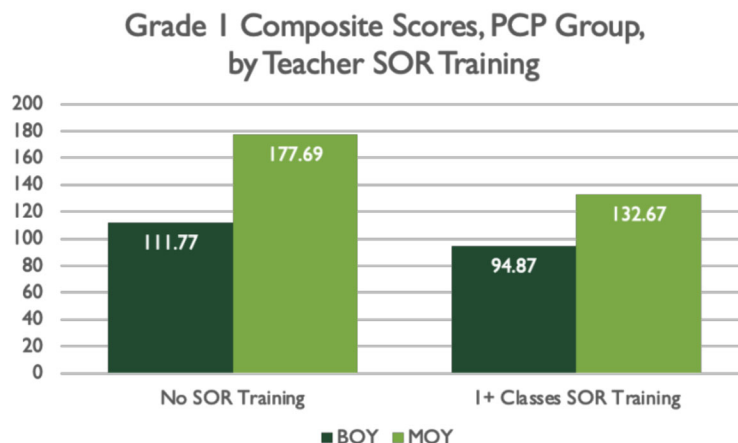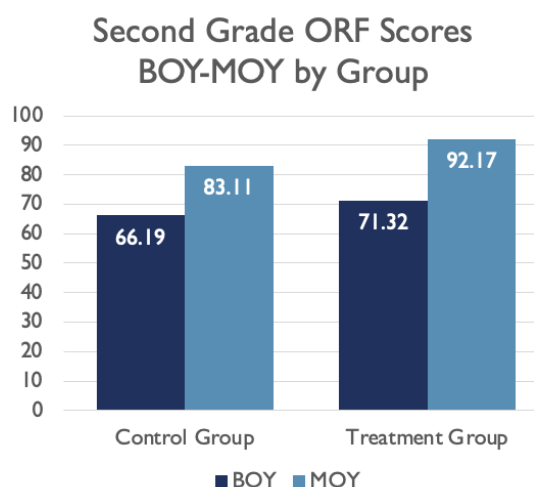**Second Grade ORF Scores BOY-MOY by Group**

Control Group: BOY 66.19, MOY 83.11
Treatment Group: BOY 71.32, MOY 92.17

■ BOY ■ MOY

- 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 by MOY?
  - The data show a positive, significant impact on student achievement on multiple Acadience subtests for all grades (K-2), as well as on the Composite scores for grades K and 2.
- How does the impact of the 95PCP vary by grade and student subgroup (gender, EL, or SPED status)?
  - Students performed similarly, regardless of Gender, EL or SPED status.
- How does the impact of the 95PCP vary by a student's BOY benchmark status (i.e., do students *Below Benchmark* at BOY achieve similar growth as students who were *At Benchmark* at BOY)?
  - Students performed similarly, regardless of BOY benchmark status. In other words, students who were *Below Benchmark* or *Well Below Benchmark* made the same amount of progress as students who started the year *At or Above Benchmark*.
- What, if any, impact does teacher training in the Science of Reading have on student scores?
  - The impact of teacher SOR training is not straightforward. For kindergarten, the training seemed to support student growth; in first grade the data show students of teachers in the training made fewer gains; and in second grade there was no apparent impact. Reasonable hypotheses about why first graders made fewer gains may include the impact of time scarcity for teachers simultaneously participating in the SOR course and implementing a new phonics program. There was no detectable impact of teacher training on the learning of students in the control schools.

As the study continues during the Winter of 2022, Administrator interviews will allow for a better understanding of how phonics instruction and intervention has progressed over the course of the first half of the year. These more in-depth conversations will provide additional context from both treatment and control schools on what instruction looks like across grades and within each building. The Spring data collection includes focus groups with educators and the EOY Acadience data collection.

Appendices

Appendix 1: Kindergarten Results

- Composite score: (IRR=0.98, p=0.044) - significant differences between treatment and control groups
- First Sound Fluency (FSF) score: (IRR=0.96, p=0.032) - significant differences between treatment and control groups
- Letter Naming Fluency (LNF) score: (IRR=-0.97, p=0.039) - significant differences between treatment and control groups

For Below or Well Below Benchmark students:

- Composite score: (IRR=-2.14, p=0.343) - no significant differences between treatment and control groups

For Below Benchmark students:

- Composite score: (IRR=-1.00, p=0.932) - no significant differences between treatment and control groups

Exploring teacher training:

- EXPERIMENTAL GROUP: Composite score: (IRR=0.94, p=0.005) - significant effect of teacher training

- CONTROL GROUP: Composite score: (IRR=1.03, p=0.184) - no effect of teacher training

*Kindergarten: **Composite Score***

Predictors	ark 6 comp K		
	Incidence Rate Ratios	*CI*	*p*
(Intercept)	5.92	4.63 – 7.55	**<0.001**
Time	4.13	3.98 – 4.30	**<0.001**
Gender	1.07	0.99 – 1.16	0.071
SPED	0.54	0.48 – 0.61	**<0.001**
ELL	0.84	0.73 – 0.97	**0.014**
group	1.11	0.96 – 1.28	0.173
Tigr	0.98	0.95 – 1.00	**0.044**

Random Effects

σ^2	0.02
τ_{00} student_id:class_name	0.39
τ_{00} class_name	0.05
ICC	0.97
N student_id	1095
N class_name	59
Observations	2190
Marginal R^2 / Conditional R^2	0.534 / 0.984

Kindergarten: *Above/Below Benchmark Comparisons*

<u>Note:</u> The first column contains data for students *Above Benchmark*, the second column contains data for students *Below Benchmark*. The variable of interest is "Tigr", which represents the interaction between "Time" and "Group" and tells us whether the amount growth on the outcome measure is different for students in the control versus treatment groups.

Predictors	ark 6 comp 1 cap			ark 6 comp 2 cap		
	Estimates	CI	p	Estimates	CI	p
(Intercept)	79.94	62.84 – 97.03	**<0.001**	136.77	123.10 – 150.44	**<0.001**
Time	28.83	20.33 – 37.33	**<0.001**	21.87	15.55 – 28.18	**<0.001**
Gender	-3.58	-9.07 – 1.90	0.200	-1.01	-6.07 – 4.06	0.696
SPED	-43.71	-52.78 – -34.64	**<0.001**	-25.54	-34.78 – -16.30	**<0.001**
ELL	-13.09	-23.39 – -2.79	**0.013**	3.64	-5.55 – 12.84	0.438
group	-3.42	-13.93 – 7.09	0.524	7.00	-1.21 – 15.21	0.095
Tigr	2.57	-2.74 – 7.89	0.343	-2.14	-6.02 – 1.74	0.280
Random Effects						
σ^2	989.67			466.76		
τ_{00}	1541.07 student_id:class_name			1319.37 student_id:class_name		
	67.67 class_name			31.13 class_name		
ICC	0.62			0.74		
N	1079 student_id			963 student_id		
	59 class_name			60 class_name		
Observations	2158			1926		
Marginal R^2 / Conditional R^2	0.147 / 0.675			0.071 / 0.761		

Kindergarten: Above/Below Benchmark Comparisons (Well Below Benchmark students excluded)

Predictors	Incidence Rate Ratios	CI	p
(Intercept)	3.54	2.88 – 4.34	**<0.001**
Time	5.49	4.93 – 6.11	**<0.001**
Gender	1.01	0.99 – 1.04	0.277
SPED	0.99	0.96 – 1.02	0.419
ELL	1.01	0.96 – 1.05	0.802
group	1.00	0.88 – 1.13	0.940
Tigr	1.00	0.93 – 1.07	0.932

Random Effects

σ^2	0.02
τ_{00} student_id:class_name	0.00
τ_{00} class_name	0.00
N student_id	375
N class_name	59
Observations	439
Marginal R^2 / Conditional R^2	0.974 / NA

ark 6 comp K

Predictors	ark 6 comp K		
	Incidence Rate Ratios	*CI*	*p*
(Intercept)	6.69	5.97 – 7.51	**<0.001**
Time	3.98	3.91 – 4.05	**<0.001**
Gender	1.14	1.03 – 1.27	**0.009**
SPED	0.58	0.50 – 0.68	**<0.001**
ELL	0.86	0.70 – 1.04	0.128
sort	1.23	0.98 – 1.55	0.068
sortxtime	0.94	0.91 – 0.98	**0.005**

Random Effects

σ^2	0.01
τ_{00} student_id:class_name	0.37
τ_{00} class_name	0.04
ICC	0.97
N student_id	598
N class_name	32
Observations	1196
Marginal R^2 / Conditional R^2	0.547 / 0.984

	ark 6 comp K		
Predictors	*Incidence Rate Ratios*	*CI*	*p*
(Intercept)	6.54	5.69 – 7.52	**<0.001**
Time	4.01	3.93 – 4.09	**<0.001**
Gender	1.00	0.89 – 1.12	0.944
SPED	0.50	0.42 – 0.60	**<0.001**
ELL	0.83	0.68 – 1.01	0.064
sort	1.32	0.99 – 1.75	0.057
sortxtime	1.03	0.99 – 1.08	0.184

Random Effects

σ^2	0.02
τ_{00} student_id:class_name	0.40
τ_{00} class_name	0.06
ICC	0.97
N student_id	497
N class_name	27
Observations	994
Marginal R^2 / Conditional R^2	0.538 / 0.985

	fsf k		
Predictors	*Incidence Rate Ratios*	*CI*	*p*
(Intercept)	2.89	2.16 – 3.88	**<0.001**
Time	2.85	2.66 – 3.05	**<0.001**
Gender	1.13	1.03 – 1.24	**0.008**
SPED	0.41	0.36 – 0.48	**<0.001**
ELL	0.76	0.64 – 0.90	**0.002**
group	1.18	0.99 – 1.41	0.068
Tigr	0.96	0.92 – 1.00	**0.032**

Random Effects

σ^2	0.06
τ_{00} student_id:class_name	0.55
τ_{00} class_name	0.07
ICC	0.92
$N_{student\_id}$	1095
N_{class_name}	59
Observations	2190
Marginal R^2 / Conditional R^2	0.334 / 0.944

Kindergarten: Letter Naming Fluency (LNF)

Predictors	Incidence Rate Ratios	CI	p
	lnf k		
(Intercept)	7.63	6.14 – 9.47	**<0.001**
Time	2.07	1.96 – 2.19	**<0.001**
Gender	1.01	0.94 – 1.08	0.865
SPED	0.69	0.62 – 0.77	**<0.001**
ELL	0.86	0.75 – 0.98	**0.024**
group	1.10	0.97 – 1.26	0.142
Tigr	0.97	0.93 – 1.00	**0.039**

Random Effects

σ^2	0.04
τ_{00} student_id:class_name	0.34
τ_{00} class_name	0.03
ICC	0.90
N student_id	1095
N class_name	59
Observations	2190
Marginal R^2 / Conditional R^2	0.240 / 0.926

Appendix 2: First Grade Results

- Composite score: (B=4.62, p=0.30) - no significant differences between treatment and control groups
- Nonsense Word Fluency Correct Letter-Sounds (CLS) score: (B=2.50, **p=0.047**) - significant differences between treatment and control groups
- Nonsense Word Fluency Whole Words Read (WWR) score: (B=1.38, **p=0.019**) - significant differences between treatment and control groups

For Below or Well Below Benchmark students:

- Composite score: (B=1.01, p=0.68) - no significant differences between treatment and control group s

For Below Benchmark students:

- Composite score: (B=-.09, p=0.949) - no significant differences between treatment and control groups

Exploring teacher training:

- EXPERIMENTAL GROUP: Composite score: (B=-14.06, p=0.025) - significant effect of teacher training

- CONTROL GROUP: Composite score: (B=-7.16, p=0.267) - significant effect of teacher training

Predictors	ark 6 comp 1		
	Estimates	CI	p
(Intercept)	60.73	34.15 – 87.30	**<0.001**
Time	50.39	36.30 – 64.49	**<0.001**
Gender	-6.19	-13.90 – 1.52	0.115
SPED	-50.58	-63.33 – -37.82	**<0.001**
ELL	-15.80	-30.29 – -1.32	**0.033**
group	-5.34	-21.73 – 11.05	0.523
Tigr	4.62	-4.19 – 13.43	0.304

Random Effects

σ^2	2716.55
τ_{00} student_id:class_name	2659.82
τ_{00} class_name	138.44
ICC	0.51
N student_id	1079
N class_name	59
Observations	2158
Marginal R^2 / Conditional R^2	0.162 / 0.587

First Grade: Above/Below Benchmark Comparisons

Note: The first column contains data for students *Above Benchmark*, the second column contains data for students *Below Benchmark*. 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.

Predictors	ark 6 comp 1 Estimates	ark 6 comp 1 CI	p	ark 6 comp 1 Estimates	ark 6 comp 1 CI	p
(Intercept)	31.63	-6.12 – 69.38	0.100	87.48	74.05 – 100.91	<0.001
Time	99.00	78.51 – 119.48	<0.001	-4.88	-12.61 – 2.85	0.216
Gender	-6.44	-15.41 – 2.54	0.160	2.13	-1.63 – 5.89	0.267
SPED	-5.30	-24.70 – 14.10	0.592	-27.61	-33.16 – -22.07	<0.001
ELL	-4.25	-22.09 – 13.59	0.641	-6.01	-12.77 – 0.75	0.081
group	3.41	-20.27 – 27.08	0.778	-1.11	-9.35 – 7.14	0.792
Tigr	-0.89	-13.77 – 11.98	0.892	1.01	-3.84 – 5.86	0.683

Random Effects

σ^2	2108.20	383.73
τ_{00}	1627.02 student_id:class_name	403.93 student_id:class_name
	129.81 class_name	37.16 class_name
ICC	0.45	0.53
N	635 student_id	753 student_id
	59 class_name	59 class_name
Observations	961	1197
Marginal R^2 / Conditional R^2	0.370 / 0.656	0.110 / 0.586

First Grade: Above/Below Benchmark Comparisons (Well Below Benchmark students excluded)

Predictors	ark 6 comp 1		
	Estimates	*CI*	*p*
(Intercept)	92.48	85.85 – 99.11	**<0.001**
Time	9.91	5.52 – 14.31	**<0.001**
Gender	1.35	-0.05 – 2.76	0.058
SPED	0.35	-2.57 – 3.28	0.814
ELL	1.17	-1.47 – 3.80	0.385
group	0.65	-3.43 – 4.72	0.755
Tigr	-0.09	-2.84 – 2.66	0.949
Random Effects			
σ^2	36.29		
$\tau_{00 \; student\_id:class\_name}$	5.93		
$\tau_{00 \; class\_name}$	0.00		
$N_{student\_id}$	317		
N_{class_name}	59		
Observations	353		
Marginal R^2 / Conditional R^2	0.391 / NA		

Predictors	Estimates	ark 6 comp 1 CI	p
(Intercept)	45.85	28.59 – 63.10	**<0.001**
Time	65.92	57.67 – 74.17	**<0.001**
Gender	-3.65	-14.39 – 7.08	0.505
SPED	-50.28	-68.18 – -32.38	**<0.001**
ELL	-11.92	-32.06 – 8.21	0.246
sort	5.61	-18.34 – 29.57	0.646
sortxtime	-14.06	-26.39 – -1.72	**0.025**

Random Effects

σ^2	2746.57
τ_{00} student_id:class_name	2653.18
τ_{00} class_name	221.95
ICC	0.51
N student_id	561
N class_name	30
Observations	1122
Marginal R^2 / Conditional R^2	0.174 / 0.597

Predictors	ark 6 comp 1		
	Estimates	CI	p
(Intercept)	58.39	40.92 – 75.86	**<0.001**
Time	51.04	41.63 – 60.45	**<0.001**
Gender	-9.53	-20.56 – 1.51	0.091
SPED	-49.46	-67.49 – -31.42	**<0.001**
ELL	-25.44	-45.90 – -4.98	**0.015**
sort	-1.47	-23.37 – 20.44	0.896
sortxtime	7.16	-5.48 – 19.80	0.267

Random Effects

σ^2	2661.62
τ_{00} student_id:class_name	2670.00
τ_{00} class_name	0.00
ICC	0.50
N student_id	518
N class_name	29
Observations	1036
Marginal R^2 / Conditional R^2	0.165 / 0.583

First Grade: Nonsense Word Fluency Correct Letter Sounds (CLS)

Predictors	Estimates	CI	p
		cls 1	
(Intercept)	16.29	7.52 – 25.07	<0.001
Time	17.83	13.87 – 21.79	<0.001
Gender	-3.81	-6.71 – -0.91	**0.010**
SPED	-16.43	-21.24 – -11.63	<0.001
ELL	-2.53	-8.00 – 2.94	0.364
group	-3.11	-8.50 – 2.28	0.259
Tigr	2.50	0.03 – 4.98	**0.047**
Random Effects			
σ^2	214.35		
τ_{00} student_id:class_name	460.10		
τ_{00} class_name	27.22		
ICC	0.69		
N student_id	1079		
N class_name	59		
Observations	2158		
Marginal R^2 / Conditional R^2	0.170 / 0.746		

First Grade: Nonsense Word Fluency Whole Words Read (WWR)

Predictors	wwr 1		
	Estimates	CI	p
(Intercept)	-0.25	-3.93 – 3.43	0.894
Time	6.00	4.16 – 7.84	**<0.001**
Gender	-1.31	-2.39 – -0.23	**0.017**
SPED	-4.72	-6.51 – -2.93	**<0.001**
ELL	-1.02	-3.06 – 1.01	0.325
group	-1.31	-3.58 – 0.96	0.259
Tigr	1.38	0.23 – 2.53	**0.019**

Random Effects

σ^2	46.36
$\tau_{00 \ student\_id:class\_name}$	55.41
$\tau_{00 \ class\_name}$	4.02
ICC	0.56
$N_{student\_id}$	1079
N_{class_name}	59
Observations	2158
Marginal R^2 / Conditional R^2	0.151 / 0.628

Appendix 3: Second Grade Results

- Composite score: (B=7.17, **p=0.024**) - significant differences between treatment and control groups
- Oral Reading Fluency Words Correct Per Minute (ORF) score: (B=3.93, **p<0.001**) - significant differences between treatment and control groups
- Oral Reading Fluency Error (ERR) score: (B=0.41, p=0.18) – no significant differences between treatment and control groups
- Retell Total (RETELL) score: (B=-0.03, p=0.98) – no significant differences between treatment and control groups
- Retell Quality of Response (RETELL QR) score: (B=-0.07, p=0.29) – no significant differences between treatment and control groups

For Below or Well Below Benchmark students:

- Composite score: (B=-3.39, p=0.51) - no significant differences between treatment and control groups

For Below Benchmark students:

- Composite score: (B=0.61, p=0.834) - no significant differences between treatment and control groups

Exploring teacher training:

- EXPERIMENTAL GROUP: Composite score: (B=-1.23, p=0.784) - no significant effect of teacher training

- CONTROL GROUP: Composite score: (B=3.25, p=0.474) - no significant effect of teacher training

Predictors	ark 6 comp 2		
	Estimates	CI	p
(Intercept)	122.73	98.63 – 146.84	**<0.001**
Time	46.74	36.63 – 56.85	**<0.001**
Gender	3.68	-5.07 – 12.42	0.410
SPED	-39.87	-55.86 – -23.89	**<0.001**
ELL	16.54	0.58 – 32.50	**0.042**
group	1.65	-12.84 – 16.15	0.823
Tigr	7.17	0.96 – 13.39	**0.024**

Random Effects

σ^2	1196.62
τ_{00} student_id:class_name	4007.74
τ_{00} class_name	180.01
ICC	0.78
N student_id	963
N class_name	60
Observations	1926
Marginal R^2 / Conditional R^2	0.160 / 0.813

Second Grade: *Above/Below Benchmark Comparisons*

Note: The first column contains data for students *Above Benchmark*, the second column contains data for students *Below Benchmark*. 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.

Predictors	ark 6 comp 2			ark 6 comp 2		
	Estimates	CI	p	Estimates	CI	p
(Intercept)	147.02	126.69 – 167.36	<0.001	76.63	48.44 – 104.83	<0.001
Time	52.75	41.98 – 63.53	<0.001	32.96	17.12 – 48.80	<0.001
Gender	6.24	0.17 – 12.32	0.044	4.42	-4.38 – 13.21	0.324
SPED	-6.75	-19.15 – 5.65	0.286	-24.49	-37.66 – -11.32	<0.001
ELL	15.17	4.33 – 26.01	0.006	-2.29	-19.03 – 14.46	0.789
group	-2.62	-14.78 – 9.54	0.673	-1.59	-19.29 – 16.11	0.860
Tigr	6.86	0.33 – 13.39	0.039	-3.39	-13.58 – 6.79	0.513
Random Effects						
σ^2	926.29			629.75		
τ_{00}	1327.01 student_id:class_name			983.00 student_id:class_name		
	48.05 class_name			64.21 class_name		
ICC	0.60			0.62		
N	802 student_id			309 student_id		
	60 class_name			60 class_name		
Observations	1456			470		
Marginal R^2 / Conditional R^2	0.315 / 0.724			0.146 / 0.679		

Second Grade: Above/Below Benchmark Comparisons (Well Below Benchmark students excluded)

Predictors	ark 6 comp 2		
	Estimates	CI	p
(Intercept)	85.36	71.61 – 99.11	<0.001
Time	42.58	33.94 – 51.22	<0.001
Gender	-1.68	-4.92 – 1.55	0.306
SPED	-6.46	-12.43 – -0.48	**0.034**
ELL	-0.14	-6.21 – 5.94	0.965
group	-1.12	-10.00 – 7.75	0.803
Tigr	0.61	-5.11 – 6.33	0.834
Random Effects			
σ^2	85.01		
τ_{00} student_id:class_name	39.16		
τ_{00} class_name	0.00		
N student_id	183		
N class_name	56		
Observations	216		
Marginal R^2 / Conditional R^2	0.848 / NA		

Second Grade: Teacher Training, Experimental Group

Predictors	ark 6 comp 2		
	Estimates	CI	p
(Intercept)	122.38	107.37 – 137.39	<0.001
Time	61.60	55.91 – 67.29	<0.001
Gender	8.04	-3.74 – 19.82	0.181
SPED	-42.81	-64.21 – -21.42	<0.001
ELL	23.35	-1.99 – 48.68	0.071
sort	2.87	-17.82 – 23.56	0.786
sortxtime	-1.23	-10.00 – 7.54	0.784

Random Effects

σ^2	1290.96
$\tau_{00 \ student\_id:class\_name}$	3968.36
$\tau_{00 \ class\_name}$	242.20
ICC	0.77
$N_{student\_id}$	529
N_{class_name}	33
Observations	1058
Marginal R^2 / Conditional R^2	0.169 / 0.805

Second Grade: Teacher Training, Control Group

	ark 6 comp 2		
Predictors	*Estimates*	*CI*	*p*
(Intercept)	132.38	117.25 – 147.50	**<0.001**
Time	52.56	46.83 – 58.29	**<0.001**
Gender	-1.39	-14.49 – 11.72	0.836
SPED	-36.55	-60.75 – -12.36	**0.003**
ELL	11.82	-8.81 – 32.45	0.261
sort	-12.38	-32.81 – 8.05	0.235
sortxtime	3.25	-5.65 – 12.15	0.474

Random Effects

σ^2	1085.55
τ_{00} student_id:class_name	4071.56
τ_{00} class_name	115.15
ICC	0.79
N student_id	434
N class_name	27
Observations	868
Marginal R^2 / Conditional R^2	0.141 / 0.823

Second Grade: Oral Reading Fluency Words Correct Per Minute (ORF)

Predictors	Estimates	CI	p
		orf 2	
(Intercept)	48.07	38.52 – 57.62	<0.001
Time	12.99	9.62 – 16.35	<0.001
Gender	1.52	-2.36 – 5.40	0.442
SPED	-10.39	-17.49 – -3.30	0.004
ELL	9.90	2.83 – 16.98	0.006
group	1.20	-4.50 – 6.91	0.679
Tigr	3.93	1.86 – 6.00	<0.001

Random Effects

σ^2	132.65
τ_{00} student_id:class_name	842.37
τ_{00} class_name	30.44
ICC	0.87
N student_id	963
N class_name	60
Observations	1926
Marginal R^2 / Conditional R^2	0.106 / 0.882

Second Grade: Oral Reading Fluency Errors (ERR)

	err 2		
Predictors	Estimates	CI	p
(Intercept)	6.78	5.08 – 8.48	<0.001
Time	-2.36	-3.34 – -1.38	<0.001
Gender	0.12	-0.30 – 0.55	0.566
SPED	1.73	0.95 – 2.50	<0.001
ELL	-0.09	-0.86 – 0.69	0.824
group	-0.61	-1.64 – 0.43	0.249
Tigr	0.41	-0.19 – 1.01	0.182
Random Effects			
σ^2	11.23		
τ_{00} student_id:class_name	5.32		
τ_{00} class_name	0.30		
ICC	0.33		
N student_id	963		
N class_name	60		
Observations	1926		
Marginal R^2 / Conditional R^2	0.055 / 0.370		

Second Grade: Retell Total (RETELL)

Predictors	retell 2		
	Estimates	CI	p
(Intercept)	13.44	7.91 – 18.97	<0.001
Time	5.84	2.94 – 8.74	<0.001
Gender	3.31	1.72 – 4.90	<0.001
SPED	-7.48	-10.39 – -4.57	<0.001
ELL	1.31	-1.60 – 4.21	0.377
group	2.33	-1.02 – 5.68	0.173
Tigr	-0.03	-1.81 – 1.76	0.978
Random Effects			
σ^2	98.67		
$\tau_{00\ student\_id:class\_name}$	103.17		
$\tau_{00\ class\_name}$	5.92		
ICC	0.53		
$N_{student\_id}$	963		
N_{class_name}	60		
Observations	1926		
Marginal R^2 / Conditional R^2	0.079 / 0.563		

Second Grade: Retell Quality of Response (RETELLQR)

Predictors	retellqr 2		
	Estimates	CI	p
(Intercept)	1.39	0.97 – 1.80	<0.001
Time	0.49	0.27 – 0.71	<0.001
Gender	0.16	0.06 – 0.26	0.002
SPED	-0.47	-0.65 – -0.28	<0.001
ELL	0.07	-0.12 – 0.26	0.466
group	0.13	-0.12 – 0.38	0.320
Tigr	-0.07	-0.21 – 0.06	0.288

Random Effects

σ^2	0.57
$\tau_{00 \text{ student\_id:class\_name}}$	0.35
$\tau_{00 \text{ class\_name}}$	0.04
ICC	0.41
$N_{\text{student\_id}}$	963
$N_{\text{class\_name}}$	60
Observations	1926
Marginal R^2 / Conditional R^2	0.061 / 0.443