



**ESSA  
Level 2  
Evidence**

**PHONICS LESSON LIBRARY™ &  
PHONICS CHIP KIT™  
2021-2022, GRADE 1  
EFFICACY STUDY**



**LXD RESEARCH  
95 PERCENT GROUP LLC**

# PHONICS LESSON LIBRARY™ & PHONICS CHIP KIT™



## EFFICACY RESEARCH

### 2021-2022 RESULTS - FIRST GRADE - INTERVENTION

#### PROGRAM DESCRIPTION

The Phonics Lesson Library (PLL) is a Tier 2 and Tier 3 Phonics intervention. The program aligns with the Phonics Screening Inventory (PSI), a diagnostic and progress monitoring tool to create small groups that receive 30 minutes of explicit and systematic phonics instruction using Phonics Chip Kits. Students accelerate skill growth along the Phonics Continuum.

#### SCHOOL DESCRIPTION

LOCATION: California

GRADE: First, Tiers 2 & 3

SIZE: 462 Students

DEMOGRAPHICS: 82% Hispanic | 33% ELL | 4% SPED | 5% Foster/Homeless

#### STUDY DETAILS

Schools in the district were paired by Spring 2021 ELA scores and then assigned to treatment and comparison groups. Schools in the treatment group used Phonics Lesson Library and Phonics Chip Kit (PLL) with first graders to provide targeted, daily, small-group lessons for students who were **Below or Well Below Benchmark** at the beginning of the year.

#### COMPARING RESULTS

Tiers 2 and 3 students using the **95 Percent Group's PLL showed higher gains on the CORE Phonics Survey and Acadience Reading than the comparison group** on multiple measures in first grade.

While Wonders has research studies, none have a comparison group. At the time of this report, Heggerty resources lack any research studies that meet ESSA-level evidence.

##### COMPARISON GROUP

#### LITERACY TOOLKIT

Core: Wonders  
Tiers 2 & 3: Heggerty and a variety of resources

##### TREATMENT GROUP

#### REVISED LITERACY TOOLKIT

Core: Wonders  
Tiers 2 & 3: **Phonics Lesson Library and Phonics Chip Kit**

##### ASSESSMENT

#### CORE PHONICS SURVEY & ACADIENCE READING

CORE Phonics Survey and Acadience Reading K-6 were conducted at the beginning and end of the 2021-2022 school year. Acadience was also conducted mid-year.

Grade	Tiers 2 & 3
Wonders	
Heggerty	
PLL	

No Evidence Limited Evidence Moderate Evidence



For more information about the Phonics Lesson Library, this study or other products, contact [info@95percentgroup.com](mailto:info@95percentgroup.com)



Learning Experience Design  
LXD Research  
a division of Charles River Media Group, LLC

### RESULTS FROM RIGOROUS STATISTICAL MODELS

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

-  Student Demographics: Gender, Ethnicity, & Foster/Homelessness
-  Student Supports: EL Status & SPED
-  Student BOY Benchmark Status
-  Phonics Diagnostic at BOY

Even with statistical controls, schools with PLL had higher gains to comparison schools on multiple measures.

Assessment	Role	Study Outcome
Acadience Reading	Universal Screener (Overall Reading)	
CORE Phonics Survey	Diagnostic	



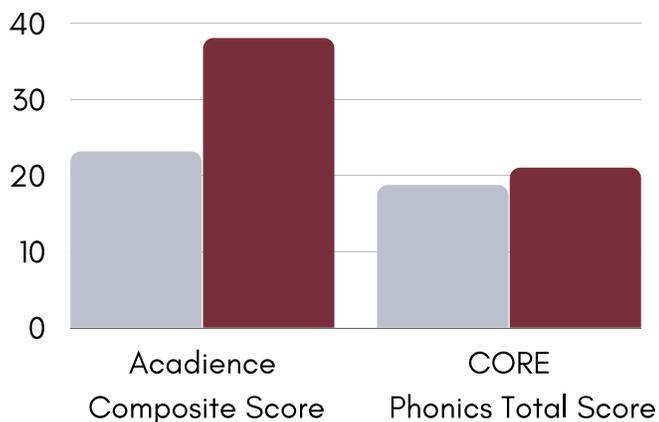
Similar Gains



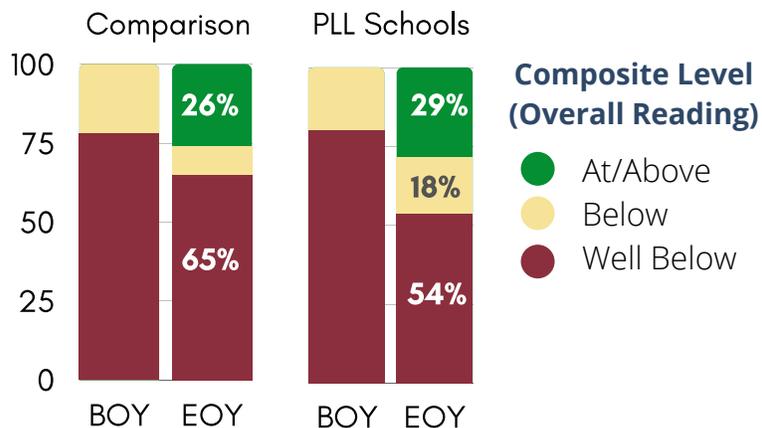
PLL Higher Gains

### BOY TO EOY OUTCOMES FOR PLL SCHOOLS

Students in schools using the Phonics Lesson Library with the Phonics Chip Kits (red) made **higher gains from BOY to EOY** than the comparison group (gray).



Across the year, these gains led to meaningfully different EOY Benchmark for PLL schools.





# 95 Phonics Lesson Library Research Study

## First Grade Results, 2021-2022 School Year

Prepared by Rachel Schechter, LXD Research

Analysis conducted by Lynch Research Associates

Introduction	3
Study Program Description	3
Comparison Programs	4
Table 1. ESSA-Level Evidence on Comparison School Programs for First Grade	4
Research Activities	4
Reading Assessments	5
Table 2a. Acadience Reading Subtests, Skill Coverage, and Benchmark Goals in First Grade	5
Table 2a. CORE Phonics Survey Subtests, Skill Coverage, and Maximum Scores	6
Student and School Demographics	7
PLL Implementation	7
95 Percent Group Coaching Summary	7
Phonics Screening Inventory Implementation Description	7
How many students have received PLL?	7
Table 3. Number of Students in Each Cycle by Study	8
Are students progressing in the program as expected?	8
Figure 1. Number of Students per Skill by Cycle	8
Results for Students Below or Well Below Benchmark at BOY	9
Sample Descriptions	9
Table 4. Sample Descriptives for Treatment and Comparison groups by Study	9
Table 5. Sample of Students with BOY and EOY Composite Scores by group	10

Analytical Approach	10
EOY-BOY Statistical Results	11
Table 7. HLM Results for Students Below or Well Below Benchmark at BOY	11
Figure 2. Students in the treatment group demonstrated significantly more growth in Composite scores than students in the comparison group.	11
Figure 4. Students in the treatment group demonstrated significantly more growth in CORE Phonics Survey total scores than students in the comparison group.	12
Conclusion and Future Research	12
Appendix 1	13
Effect Sizes based on t-tests	13
BOY-EOY	14
Composite Score	14
Phonics Total	15



## Introduction

Learning Experience Design Research (LXD Research) is an independent evaluation, research, and consulting division within Charles River Media Group focusing on education. LXD Research designs rigorous research studies, multifaceted data analytic reporting, and dynamic content to disseminate insights. For 95 Percent Group, LLC, the team is conducting a 2021/22 study of the implementation and impact of the Phonics Lesson Library (PLL) with the Phonics Chip Kit in a medium-sized school district in California with a student population over 80% Hispanic and over 25% English Language Learners. The study used quasi-experimental designs to generate evidence of the program's impact that aligns with evidence standards associated with ESSA.

## Study Program Description

The PLL is designed primarily for Tier 2 or Tier 3 phonics intervention. In the study, the Basic version of the PLL was used with first graders in four intervention schools, which used Wonders as their core curriculum. Teachers used an initial diagnostic screener to place students into intervention groups and then the used of 95 Percent Group's Phonics Screener for Intervention™ (PSI) to progress monitor. Students who were Below or Well Below Benchmark were identified for intervention using Acadience Reading K-6 and then placed into lessons along the Phonics Continuum (see graphic below). The PLL supports students who are not meeting benchmarks through comprehensive lesson plans that target skills aligned with the Phonics Continuum, from learning simple letter-sound correspondences to blending words with more complex and variable letter combinations to using syllabication to decode multisyllabic words. This study combined the PLL with the Phonics Chip Kit.

Students received 30 minutes of daily intervention through a push-in model in small groups of three to four students who were at similar levels. Instructors monitored progress through alternate forms of the PSI and used this data to re-group students every 3 weeks based on their highest-need skill (i.e., the lowest skill on the continuum that needs support). Instruction is grounded in and aligned with evidence-based instructional practices in structured literacy, following the Science of Reading research base. Instruction is systematic, following a developmental progression from simple to complex; explicit, introducing new skills with direct, multisensory instruction and a gradual release of responsibility from teacher to students; and diagnostic, targeted to students' specific skill needs as determined by frequent assessment. Once students reach mastery of skills for their grade level, they have completed the intervention.

---

## Comparison Programs

In the comparison schools survey, most teachers (73%) responded that they used their core curriculum, Wonders, to support Tier 2 and Tier 3 reading intervention. One-third mentioned using Heggerty Phonemic Awareness resources (36%). Other products included but were not limited to: Imagine Learning (44%), Heggerty Phonemic Awareness (44%), Heggerty Bridge the Gap (22%), as well as the core curriculum Journeys (39%). Nearly all of these programs describe their materials as based in the science of reading and represent a relatively high bar as a comparison to the 95 Percent Group programs. Evidence related to these programs' effectiveness for first graders varies, however, with most not having any evidence with first graders (Table 1). In both studies, comparison schools implemented intervention in a variety of ways, usually pulling students out for thirty minutes for Tier 3 and using small-group instruction during the reading block for Tier 2.

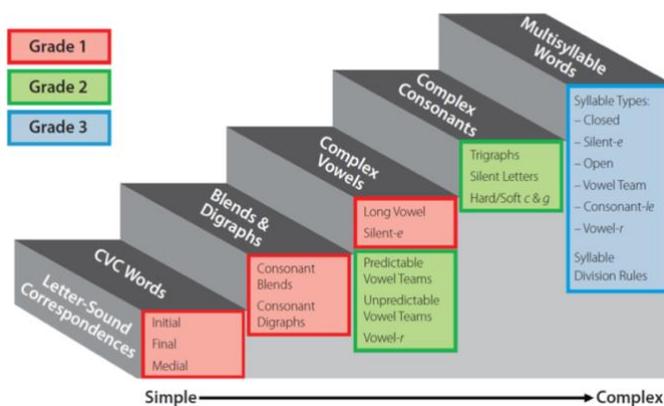


Table 1. ESSA-Level Evidence on Comparison School Programs for First Grade

Product	Evidence for All Students	Evidence for Tiers 2 - 3
Wonders	<a href="#">Limited</a>	None
Heggerty Phonemic Awareness & Bridge the Gap	None	None

## Research Activities

Throughout the school year, multiple research activities took place to understand the nature of the PLL and comparison school phonic implementations and measure the impact of those programs. These activities included:

- Conducting Acadience<sup>®</sup> Reading K-6 with all students at the beginning, middle, and end of the year (BOY, MOY, EOY)
- Conducting segments from the Consortium on Reaching Excellence in Education (CORE) Phonics Survey with all students with Below Benchmark or Well Below Benchmark scores on Acadience

- Teacher surveys, interviews with instructional coaches, and classroom observations

This report focuses on the gains from BOY to EOY on Acadience Reading and the CORE Phonics Survey for the students who performed Below or Well Below Benchmark on Acadience Reading in Fall 2021 (BOY). Results from the qualitative data collected from the surveys, interviews, and observations will be available at the end of Summer 2022.

## Reading Assessments

Acadience Reading K-6 assessments were administered by a special assessment team (not classroom teachers) in both school districts at the beginning of Fall 2021, in January 2022, and May 2022. As a set of curriculum-based measures, Acadience Reading assesses student development as a reader. Designed for universal screening and benchmarking to determine the appropriate supports for each student, Acadience is administered three times per year in the fall, winter, and spring. Assessments are administered observationally in a one-on-one setting and take between 3 and 11 minutes per student to complete. Scores include standardized scale scores and on-grade achievement-level placements. First grade Acadience Reading subtests are listed in Table 2a, along with the skills they assess and the benchmark goals for the times of year they are administered (the measures administered vary by time of year based on expected skill development). Note that the LNF measure does not have benchmark goals because it is an indicator of risk rather than an indicator of a basic early literacy skill. At each administration period, subtest scores are weighted and combined into a Composite Score, which is an overall indicator of reading ability.

*Table 2a. Acadience Reading Subtests, Skill Coverage, and Benchmark Goals in First Grade*

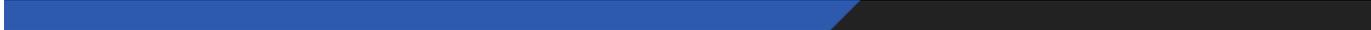
Subtest	Indicators of These Basic Early Literacy Skills	Benchmark Goals		
		BOY	MOY	EOY
Letter Naming Fluency	Indicator of Risk	None	--	--
Phoneme Segmentation Fluency (PSF)	Phonemic Awareness	47 <b>40</b> 25	--	--
Nonsense Word Fluency: Correct Letter Sounds (CLS)	The Alphabetic Principle and Basic Phonics	34 <b>27</b> 18	59 <b>43</b> 33	81 <b>58</b> 47
Nonsense Word Fluency: Whole Words Read (WWR)	The Alphabetic Principle and Basic Phonics	4 <b>1</b> 0	17 <b>8</b> 3	25 <b>13</b> 6

Subtest	Indicators of These Basic Early Literacy Skills	Benchmark Goals		
		BOY	MOY	EOY
Oral Reading Fluency (ORF): Words Correct	Accurate and Fluent Reading of Connected Text	--	34 <b>23</b> 16	67 <b>47</b> 32
Oral Reading Fluency (ORF): Accuracy	Advanced Phonics and Word Attack Skills; Accurate and Fluent Reading of Connected Text	--	86% <b>78%</b> 68%	97% <b>90%</b> 82%
Oral Reading Fluency (ORF): Retell	Reading Comprehension	--	--	17 <b>15</b> 0
Composite	Overall Estimate of Reading Ability	129 <b>113</b> 97	177 <b>130</b> 100	208 <b>155</b> 111

The [CORE Phonics Survey](#) sections C-F (Table 2b) were administered by the special assessment team to both intervention and comparison students at the participating schools in Fall 2021 and Spring 2022. The CORE survey measures phonics and phonics-related skills in beginning reading. Each survey presents several lists of letters and words for students to identify or decode. Like in Acadience, pseudowords are also included to assess decoding skills using words that students have never seen before, requiring them to use their knowledge of letter-sounds and blending without allowing them to rely on context or memorization. The CORE Phonics Survey was used as a diagnostic tool to place students into initial reading groups for this study, and then the PSI was used to track student progress from earlier skills to grade-level mastery (and then exiting intervention as appropriate). The surveys take approximately 10 minutes to administer.

*Table 2b. CORE Phonics Survey Subtests, Skill Coverage, and Maximum Scores*

Section	Description	Maximum Score
C	Letter Sounds, Consonants	21
D	Letter Sounds, Long Vowels	5
D	Letter Sounds, Short Vowels	5



<b>Section</b>	<b>Description</b>	<b>Maximum Score</b>
E	Reading & Decoding, Short Vowels in CVC Words	15
F	Reading & Decoding, Consonant blends with short vowels	15

**Student Demographics**

Student demographics that may be related to outcome measures were collected, including: school, district, gender, grade, race/ethnicity, age, English Language Learner status, economic disadvantage status (the likely proxy is an indicator of whether a student qualifies for free or reduced-price meals [FRM]), homeless status, migrant status, and special education status.

**PLL Implementation**

**95 Percent Group Coaching Summary**

Training to support first-grade teachers in the treatment group was provided in both school districts before school started. Coaches gave guidance on how to use the CORE assessments to place intervention students in initial groups. The use of the PSI began with Cycle 2, and the PLL was used during intervention time. With each cycle, teachers created student groups to focus on specific Phonics skills. If a group of first graders needed phonological awareness lessons before starting the PLL, those lessons were made available to the teachers. Over time, students would advance through the 95 Percent Group Phonics Continuum. Consultants were available to support coaches and discuss questions three times (Fall 2021, Winter 2022, and Spring 2022).

**Phonics Screening Inventory Implementation Description**

Teachers completed the PSI every three weeks as part of the intervention. The results of these screeners informed student groupings and the targeted skill for the cycle’s lessons. This section of the report summarizes the number of students who have been identified and served by literacy intervention. Cycle 1 is not included below because the schools used the CORE phonics survey instead of the PSI to eliminate redundant testing.

How many students have received PLL?

The number of students grouped for intervention during each cycle between August and December 2021 is displayed below (Table 3). All the PSI results were shared through an aggregate report provided to the research team, so it was not possible to follow individual students or connect PSI data



to Acadience data in this study. We learned from the instructional coaches that all schools were nearly fully ramped up by Cycle 3.

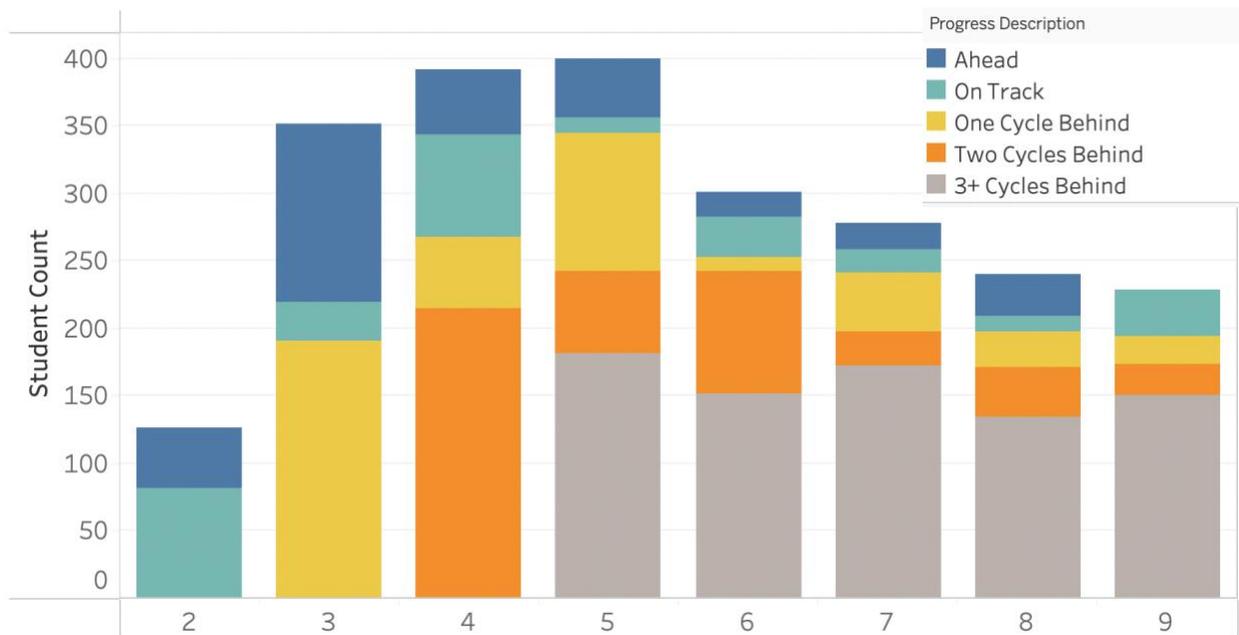
*Table 3. Number of Students in Each Cycle by Study*

Grade	Cycle Number							
	2	3	4	5	6	7	8	9
First Grade	118	242	287	275	190	172	148	176

Are students progressing in the program as expected?

Most students started the school year working on skills at the start of the skills progression for each program. While evidence of students advancing through the program becomes clearer by Cycle 5, many are behind the expected progression (Figure 1). As students show mastery in all the skills, they “place out” and no longer receive phonics intervention. Starting in Cycle 6, the total number of students reduce and less than 25% of students advance categories each week. While it cannot be seen in this data visualization, most students are moving forward in each cycle, even though they are still behind.

*Figure 1. Number of Students per Skill by Cycle*



The PLL is a highly explicit, scripted program that includes multisensory activities for students via the Phonics Chip Kit and requires teachers to provide students with direct feedback as they work. Research with similarly explicit 95 Percent Group materials has shown that it takes some time for teachers to master lesson delivery and that lessons may take longer to provide until they become more familiar with the approach and format. Furthermore, research from the National Council on Teacher Quality ([Drake & Walsh, 2020](#)) has shown that only 68% of teacher preparation programs covered phonics instruction in 2020, while in 2013, only 53% of programs covered phonics. While progress in teacher preparation has been made, many teachers currently in the field are unprepared to provide explicit phonics instruction without additional professional development. Thus, teachers in the treatment group may be learning how to teach the skills in the PLL as they are also adjusting to using a new program within a new intervention model. It is expected that over time, with use of the program and coaching, they may become increasingly familiar with the concepts and comfortable with program delivery, and students will advance through the program more quickly.

## Results for Students Below or Well Below Benchmark at BOY

### Sample Descriptions

Because this program is an intervention program, this report focuses on students who scored Below or Well Below Benchmark in Acadience Reading in Fall 2021. A total of 462 first graders had BOY, MOY, and EOY data. Using a quasi-experimental design to examine the effects of the 95 Percent Group’s PLL, a portion of schools used the walk-to-intervention program (treatment), and another portion did not (comparison). Of these students, 357 students were in the treatment group and 243 students were in the comparison group (see Table 4 for details).

Students in treatment and comparison were similar in regard to gender, Hispanic race/ethnicity, and ELL status. However, 1st grade students in the comparison group were more likely to receive Special education services ( $\chi^2=7.81, p = .005$ ; see Table 4) than students in the treatment group.

*Table 4. Sample Descriptives for Treatment and Comparison groups by Study*

Grade	Group	Male	SPED	ELL	Hispanic	Foster/ Homelessness
First Grade	Comparison	51%	6%	31%	80%	5%
	Treatment	52%	1%	38%	83%	5%

Within the sample of 462 students who had both BOY and EOY data available, we found no statistically significant differences in BOY composite scores in the treatment versus comparison group

( $t=-1.31, p=.19$ ). Table 5 below displays the average BOY scores for students who had BOY and EOY scores.

*Table 5. Sample of Students with BOY and EOY Composite Scores by group*

Grade	Condition	Number of Students	BOY Average	SD	Significance	Effect Size Cohen's d
First Grade	Comparison	214	65.16	34.03	p=.71	0.03
	Treatment	248	66.29	31.24		

### Analytical Approach

Two level hierarchical linear regression models (HLMs) with time (level 1) nested within students (level 2) were employed to examine growth in composite and subscale scores. All models contained a series of covariates including gender (“female”; 1=female, 0=male), Hispanic ethnicity (“hispanic”; 1=Hispanic, 0=Not Hispanic), ELL status (“ELL”; 1=ELL, 0=non-ELL), SPED status (“SPED”; 1=SPED, 0=non-SPED), an indicator of fostering/homelessness (“foshom”; 1= in foster care or homeless, 0=not in foster care or homeless), an indicator of time (“Time”; 1=BOY, 2=MOY), an indicator of whether the student was in the treatment or comparison group (“intervention”; 1=Treatment, 0=comparison), and an interaction between time and group calculated as the product of Time\*group (“Tigr”).

We explored main effects of treatment vs comparison 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 or subscale score is different for the treatment versus comparison groups. We also looked at growth in composite scores separately based on students’ benchmark scores and growth in Phonics Total scores among first graders. All analyses were conducted using the statistical software package R 3.6.2.

In the current sample, some assessments were administered at the Beginning of Year (BOY) and End of Year (EOY), which we refer to in this report as “BOY-EOY assessments”. Other assessments were administered at the Middle of Year (MOY) and again at End of Year (EOY) - we refer to these as MOY-EOY assessments. Table 6 summarizes when the various assessments were administered.

Table 6. List of Assessments by Grade Level by Time Period Sample

Grade Level	BOY-EOY Assessments	MOY-EOY Subtests
1	Composite scores, Nonsense Word Fluency CORE Phonics Survey	Oral Reading Fluency

## EOY-BOY Statistical Results

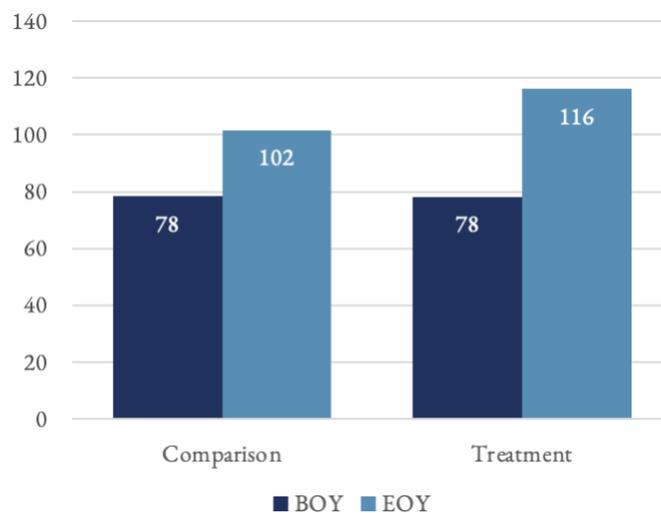
Within the First Grade sample of students who were Below or Well Below Benchmark at BOY, we examined growth in Composite scores, Nonsense Word Fluency (NWF), Oral Reading Fluency (ORF), and CORE Phonics Total scores. For all the Composite and the CORE Phonics Survey, the PLL schools outperformed the schools that did not use PLL (Table 7). In other words, was a significant effect of treatment on Composite scores ( $B=14.90$ ,  $p=.027$ ,  $f^2=.01$ ) and CORE Phonics Survey total scores ( $B=4.31$ ,  $p=.002$ ,  $f^2=.02$ ), suggesting that students in the treatment group demonstrated more growth than students in the comparison group (see Figures 2-4). There were no statistically meaningful differences between treatment and comparison group in regard to NWF or ORF. Complete output for each model can be found in the [Appendix](#).

*Table 7. HLM Results for Students Below or Well Below Benchmark at BOY*

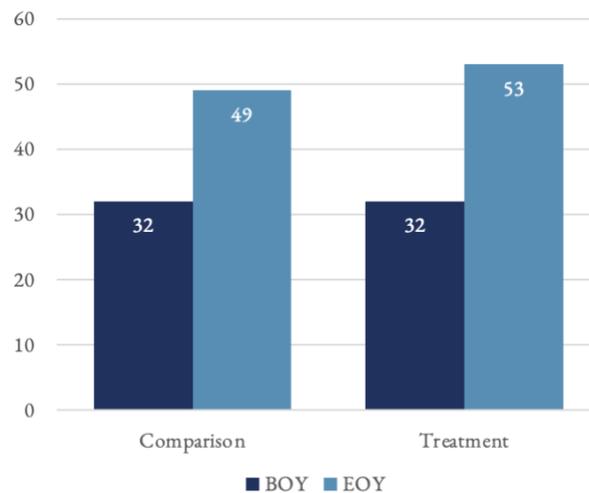
Test	School Group	BOY	EOY	Statistically Different?
Composite Scores	Wonders + Variety	78.39	101.53	Yes, they are different. Treatment group saw significant more growth from BOY to EOY.
	Wonders + PLL	78.11	116.15	
CORE Phonics Survey	Wonders + Variety	32.42	49	Yes, they are different. Treatment group saw significant more growth from BOY to EOY.
	Wonders + PLL	32.06	52.95	

*Acadience Subtests NWF and ORF (from MOY to EOY) showed similar growth for both groups.*

*Figure 2. Students in the treatment group demonstrated significantly more growth in Composite scores than students in the comparison group.*



*Figure 3. Students in the treatment group demonstrated significantly more growth in CORE Phonics Survey total scores than students in the comparison group.*



### **Conclusion and Future Research**

The Phonics Lesson Library and Phonics Chip Kit set is an intensive, highly scripted, and multisensory toolkit. These new routines for teaching phonics explicitly may require an adjustment period for teachers to become comfortable and proficient with them. In addition, changing the model of intervention in a school from a pull-out to a walk-to-intervention model takes many months to adopt and become routine. It is very encouraging to see that despite these challenges, students' overall reading and phonics scores improved across the year and outperformed students in the comparison group that were not using the PLL.

Future research that follows these students into second and third grade could help educators understand the longer-term impact of the walk-to-intervention model and the use of high-quality phonics instructional materials. It would be also helpful to understand if students who received explicit phonological awareness instruction in kindergarten would see increased benefits from an explicit phonics program. Studies to investigate these questions are planned to help both program developers and teachers better understand how to support all students learning to read.

## Appendix

### 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 comparison groups.

T-Tests were run for First Grade Below/Well Below Benchmark Students.

Grade	Condition	Number of students	Composite Difference average	SD	Significance	Effect Size Cohen's d
1st grade	Treatment	248	38.27	68.13	p=.02	0.13
	Comparison	214	22.93	72.74		

### BOY-EOY

- Composite score: (B=14.90, p=.027) - significant differences between treatment and comparison group
- Phonics Total score: (B=4.31, p=.002) - significant differences between treatment and comparison group
- CLS score: (B=3.40, p=.15) - no significant differences between treatment and comparison group
- WWR score: (B=0.43, p=.67) - no significant differences between treatment and comparison group

### MOY-EOY

- ORF WC score: (B=1.95, p=.09) - no significant differences between treatment and comparison group
  - ORF Accuracy score: (B=-0.90, p=.44) - no significant differences between treatment and comparison group
-

## BOY-EOY

### Composite Score

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 comparison versus treatment groups.

<i>Predictors</i>	<i>Estimates</i>	<b>comp</b>	
		<i>CI</i>	<i>p</i>
(Intercept)	55.25	35.56 – 74.94	<0.001
Time	23.14	13.51 – 32.77	<0.001
female	-2.87	-12.54 – 6.81	0.561
hisp	-1.85	-14.69 – 10.99	0.778
ELL	-17.41	-27.93 – -6.90	0.001
sp	-64.46	-91.56 – -37.36	<0.001
foshom	-18.53	-40.34 – 3.27	0.096
R intervention	-15.18	-37.19 – 6.83	0.176
Tigr	14.90	1.73 – 28.06	0.027
<b>Random Effects</b>			
$\sigma^2$	2501.66		
$\tau_{00}$ X.95ID	1416.25		
ICC	0.36		
N X.95ID	447		
Observations	894		
Marginal R <sup>2</sup> / Conditional R <sup>2</sup>	0.114 / 0.434		

## Phonics Total

<i>Predictors</i>	<b>phonics</b>		
	<i>Estimates</i>	<i>CI</i>	<i>p</i>
(Intercept)	15.84	11.75 – 19.94	<0.001
Time	16.58	14.66 – 18.49	<0.001
female	0.80	-1.33 – 2.93	0.463
hisp	-0.49	-3.28 – 2.30	0.731
ELL	-3.20	-5.52 – -0.88	0.007
sp	-12.08	-22.30 – -1.87	0.020
foshom	-6.40	-11.43 – -1.37	0.013
R intervention	-4.67	-9.28 – -0.07	0.047
Tigr	4.31	1.65 – 6.97	0.002
<b>Random Effects</b>			
$\sigma^2$	84.50		
$\tau_{00}$ X.95ID	75.95		
ICC	0.47		
N X.95ID	430		
Observations	778		
Marginal R <sup>2</sup> / Conditional R <sup>2</sup>	0.375 / 0.671		