



PHONOLOGICAL AWARENESS  
LESSONS RESEARCH STUDY  
KINDERGARTEN,  
2022-2023 SCHOOL YEAR,  
MID YEAR REPORT

Rachel L. Schechter, Ph.D. & Alicia D. Lynch, Ph.D.  
August 2, 2023

LXD RESEARCH  
95 PERCENT GROUP LLC



# Understanding ESSA Evidence

In December 2015, the Every Student Succeed Act (ESSA) was passed encouraging education programs to provide evidence of effectiveness and impact in order to be federally supported. EvidenceforESSA.org provides standards to assess the varying levels of strength of research for education products.

The categories for ESSA Evidence are: strong, moderate, and promising evidence of effectiveness, or demonstrates a rationale to be effective.

- Level/Tier 1: Strong - At least one randomized, well-conducted study showing significant positive effect on at least one outcome measure, analyzed at proper clustering (class/student or school level), with a multi-site sample of at least 350 students.
- Level/Tier 2: Moderate - At least one quasi-experimental (i.e., matched), well-conducted study showing significant positive student outcomes, analyzed at class/student or school level, with a multi-site sample of at least 350 students.
- Level/Tier 3: Promising - Would have qualified for Tier 1/2, but did not account for clustering, but obtained significantly positive outcomes at student level or did not meet sample size required. Post-hoc or retrospective studies may also qualify.
- Level/Tier 4: Demonstrates a Rationale - Well defined logic model based on rigorous research, an effort to study intervention effects is planned or currently underway.

## This study meets the requirements for Level 2: Moderate

- ✓ Study has compared experimental groups to control groups through matching
- ✓ Matching/weighting conducted prior to posttest collection or during the early stage of intervention implementation
- ✓ Studies with fewer than 50 clusters or 350 students need to demonstrate pretest equivalence
- ✓ The dependent variable(s) include a quantitative measure of academic achievement
- ✓ Study duration is at least 12 weeks, from program inception to posttest
- ✓ Study has at least 2 teachers and 30 students per treatment
- ✓ From pretest to posttest, attrition (dropout) is similar between experimental and control groups
- ✓ Study uses a form of a program that could in principle be replicated
- ✓ If subjects were assigned or treated in clusters (classes or schools), statistical significance for clustered designs used HLM, with pretests and other variables as covariates, or other methods accounting for clustering



# 95 Phonological Awareness Lessons Research Study

Kindergarten, 2022-2023 School Year, Mid-Year Report

Prepared by Rachel L. Schechter, LXD Research

Analysis conducted by Alicia D. Lynch, Lynch Research Associates

Abstract	4
Introduction	5
Study Program Description	5
Comparison Programs	6
Table 1. ESSA-Level Evidence on Comparison School programs for Kindergarten	6
Fall 2022 Research Methods	7
Reading Assessments	7
Table 2. Acadience Reading Subtests, Skill Coverage, and Benchmark Goals in Kindergarten	7
Student and School Demographics	8
Fall Implementation	8
95 Percent Group Coaching Summary (Before this Study)	8
Results	8
Sample Descriptions	8
Table 4. Sample sizes at baseline and midyear	9
Table 5. Sample descriptives	9
Table 6a. Reading Composite Score: T-tests were run for Kindergarten Well below/Below students.	9
Table 6b. LNF Score: T-tests were run for Kindergarten Well below/Below students.	9
Table 6c. FSF Score: T-tests were run for Kindergarten Well below/Below students.	10
Analytical Approach	10
BOY-MOY Statistical Results	10
Key Findings	11
Table 7. Results for All Students	12
Conclusion and Next Steps	13
References	13
Appendices	14

## Abstract

This study aims to measure the impact of 95 Percent Group's Phonological Awareness Lessons (PA Lessons), an intervention program designed primarily for Tier 2 or 3. The study is a replication of Schechter & Lynch (2023) research conducted in the same medium-sized school district, Val Verde Unified School District, in California during the 2021-2022 school year. It presents the fall results from September 2022 to January 2023 using Acadience® Reading data. The same school groups (treatment with PA Lessons and comparison without the PA Lessons) continue from the previous study and therefore uses a quasi-experimental design to generate evidence of the program's impact that aligns with evidence standards associated with ESSA Level 2 (Moderate). A total of 527 kindergarten students from 11 schools who were well below/below benchmark in the Fall of 2022 participated in the study. Of these students, 257 were in the treatment group, and 270 were in the comparison group. The demographic characteristics of the comparison group included 56% males, 3% with special education (SPED) status, 44% English Language Learners (ELL), 88% from diverse ethnic/racial backgrounds, and 24% identified as foster/homeless. Similarly, the treatment group demographics comprise 51% males, 2% with SPED status, 36% ELL, 87% from diverse ethnic/racial backgrounds, and 20% identified as foster/homeless. Results of the study found that using PA Lessons substantially impacted student growth for Tiers 2 & 3 over the fall. PA Lesson schools had reduced Well Below Grade Level by half (67% to 28%). Comparison schools only reduced the group by less than a third (67% to 47%). Notably, 41% of PA Lesson students were At/Above grade level mid-year, compared to only 25% in other schools.

*Keywords:* 95 Percent Group, literacy, phonological awareness, Acadience Reading assessment, kindergarten, evaluation

## Introduction

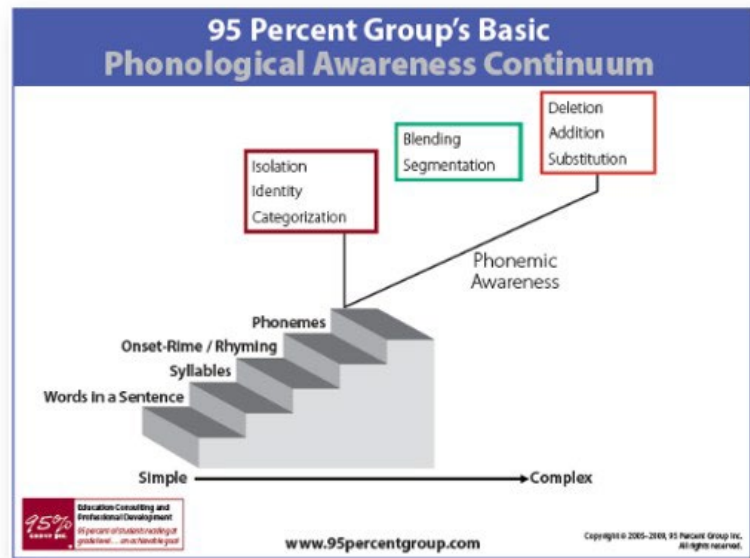
Learning Experience Design Research (LXD Research) is an independent evaluation, research, and consulting division within Charles River Media Group focusing on education. LXD Research designs rigorous research studies, multifaceted data analytic reporting, and dynamic content to disseminate insights. For 95 Percent Group, LLC, the team is conducting a study during the 2022/2023 school year to measure the impact of Phonological Awareness Lessons (PA Lessons) in a medium-sized school district in California. The study is a replication study of Schechter & Lynch (2023), conducted in the same school district (Val Verde Unified School District) during the 2021-2022 school year. This study keeps the same school groups (treatment with PA Lessons and comparison without the PA Lessons), and therefore uses a quasi-experimental design to generate evidence of the program's impact that aligns with evidence standards associated with ESSA Level 2 (Moderate). Half of the schools used the program last year, and are continuing to use PA Lessons with a new group of kindergartners. The other half of the schools are continuing to use their usual programs, as they did last year. This interim report describes the Fall results from September 2022 to January 2023.

To measure the impact of a new program in a classroom, it might be essential to evaluate the program over a few years after initial implementation as teacher experience and comfort with the program might influence student outcomes. A meta-review of 30 studies varying in their target population (grade level, subjects, geography) and methods demonstrate that teacher experience is positively correlated with student achievement gains and on other measures of success (Kini & Podolsky, 2016). Teaching effectiveness rises sharply in the first few years and continues to follow an upward trend well into the teacher's second or third decade of teaching (Kini & Podolsky, 2016). For instance, one study evaluated teachers' perspectives on the implementation of a Kindergarten Readiness Assessment across two years and found that teachers reported the administration of the program to be less burdensome in Year 2 when compared to Year 1 (Schachter et al., 2019). Extending from this, teacher comfort with implementing a new program might be positively correlated with teaching effectiveness and observing an impact on measures of student outcome.

### Study Program Description

The Phonological Awareness (PA) Lessons are designed primarily for Tier 2 or Tier 3 phonological awareness intervention for kindergarten. In the study, the PA Lesson intervention included the use of an initial diagnostic screener, and then the use of the 95 Percent Group's Phonological Awareness Screener for Intervention™ (PASI) to group students into intervention groups based on skill needs every three weeks. Students who were Below Benchmark are identified for intervention through use of a curriculum-based measure or an early literacy screener assessment used by the district, and then placed into lessons along the Phonological Awareness Continuum through the PASI. The PA Lessons support students who are not meeting benchmarks through comprehensive lesson plans that target skills aligned with the Phonological Awareness Continuum, from readiness (understanding concepts

and terms; applying language) through phonological awareness (syllables; onset rimes; phonemes). Students received 20-30 minutes of daily intervention through a push-in model, in small groups of three to four students who are at similar levels. Instructors monitor progress through alternate forms of the PASI and use this data to re-group students every 3 weeks based on the lowest skill on the continuum that needs support. Instruction is grounded in and aligned with evidence-based instructional practices in literacy. Once students reach mastery of skills for their grade level, they have completed the intervention.



### Comparison Programs

A comparison school survey was conducted during the 2021/2022 school year to better understand what “business as usual” looked like in their classrooms. Most teachers (42%) responded that they used their core curriculum, Wonders, to support Tier 2 and Tier 3 reading intervention. A small group of teachers also mentioned using Heggerty Phonemic Awareness resources (17%). At the time of this report, neither of these products had efficacy research for Kindergarten (Table 1). 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 Kindergarten

Product	Evidence for All Students	Evidence Tiers 2 - 3
Wonders	None	None
Heggerty	None	None

## Fall 2022 Research Methods

The goals of the fall activities were to begin to understand the nature and extent of literacy program implementation in comparison schools, as well as to understand mid-year gains. During Fall 2022, district leaders supported data collection to initiate the study and support product coaching services. Those activities included:

- Conducting Acadience® Reading K-6 with all students at the beginning of the year (BOY)
- Conducting PASI at the start of the year and every three weeks for progress monitoring (with treatment schools)
- Conducting and sharing Acadience Reading with all students in the middle of the year (MOY)

This report focuses on the gains from BOY to MOY on Acadience Reading.

## Reading Assessments

Acadience Reading K-6 assessments were administered by a special assessment team (not classroom teachers) at both school districts at the beginning of Fall 2022. As a set of curriculum-based measures Acadience Reading assesses student development as a reader. Designed for universal screening and progress monitoring to determine the appropriate supports for each student, Acadience is administered three times per year in fall, winter, and spring. Assessments take between 3 and 11 minutes per student to complete. Scores include standardized scale scores and on-grade achievement-level placements. Kindergarten Acadience Reading subtests are listed in Table 2, 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 2. Acadience Reading Subtests, Skill Coverage, and Benchmark Goals in Kindergarten*

Subtest	Indicators of These Basic Early Literacy Skills	Benchmark Goals		
		BOY	MOY	EOY
First Sound Fluency	Phonemic Awareness	X	X	
Phoneme Segmentation Fluency	Phonemic Awareness		X	X
Letter Naming Fluency	Indicator of Risk	N/A	N/A	N/A
Nonsense Word Fluency	The Alphabetic Principle and Basic Phonics		X	X
Composite	Overall Estimate of Reading Ability	X	X	X

## **Student and School 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 FRM), homeless status, migrant status, attendance rate, special education status, and whether or not a student was retained in a grade. School characteristics that may affect outcomes were also collected, including percent English language learners, percent students in special education, and total student enrollment.

## **Fall Implementation**

### **95 Percent Group Coaching Summary (Before this Study)**

Training to support kindergarten teachers was provided throughout last year and this year (2022-2023). The special assessment team participated in an initial workshop followed by three coaching visits at their respective schools. Coaches provided guidance on how to use the beginning of year assessments to place students in initial groups. Treatment schools used the PASI to place students into groups for the PA Lessons, used during intervention time. With each cycle, teachers created student groups to focus on specific PA or Phonics skills. Over time, students advance through the 95 Percent Group PA Continuum. Coaches consulted and discussed questions during the Fall and returned for a visit in Winter 2022.

## **Results**

### **Sample Descriptions**

A total of 527 kindergarten students from 11 schools who were well below/below benchmark in the Fall of 2022 participated in a quasi-experimental design examining the effects of the 95 Percent Group's PA Lessons in the Val Verde, California school district between the Beginning of Year (BOY) 2022 and Middle of Year (MOY) 2023. Of these students, 257 were in the treatment group and 270 were in the comparison group (see Table 4). We employed Chi-Square analyses to compare students in the treatment and comparison groups in regard to gender, special education status (SPED), English Language Learner (ELL), Hispanic race/ethnicity and rates of Foster/Homelessness. Results suggested there were no statistically meaningful differences between the treatment and comparison groups in regard to gender, ELL, SPED, Hispanic race/ethnicity and rates of Foster/Homelessness (see Table 5).



Table 4. Sample sizes at baseline and midyear

School Group	BOY	MOY	Sample
	# of Students	# of Students	# of Students
Treatment	276	255	255
Comparison	270	266	266
Total	536	521	521

Table 5. Sample descriptives

Group	Male	SPED	ELL	Ethnicity/ Race	Foster/ Homelessness
Comparison	56%	3%	44%	88%	24%
Treatment	51%	2%	36%	87%	20%

### Baseline Reading Assessment Scores

In Tables 6a-c, we report t-test results comparing baseline assessment scores in the students enrolled in treatment and comparison groups and their effect sizes (Cohen's d). Students enrolled in the treatment and comparison groups were similar in their baseline (BOY) Composite, LNF and FSF scores.

Table 6a. Reading Composite Score: T-tests were run for Kindergarten Well below/ Below students.

Condition	Number of students	BOY Composite Average Score	SD	Significance	Effect Size Cohen's d
Comparison	276	8.63	7.85	p=.61	.04
Treatment	260	8.98	7.95		

Table 6b. LNF Score: T-tests were run for Kindergarten Well below/ Below students.

Condition	Number of students	BOY LNF Average Score	SD	Significance	Effect Size Cohen's d
Comparison	276	4.01	5.01	p=.19	.12
Treatment	260	4.59	5.00		

Table 6c. FSF Score: T-tests were run for Kindergarten Well below/ Below students.

Condition	Number of students	BOY FSF Average Score	SD	Significance	Effect Size Cohen's d
Comparison	276	4.61	6.05	p=.66	.04
Treatment	260	4.39	5.83		

### Analytical 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 (“female”; 1=female, 0=male), SPED status (“sp”; 1=SPED, 0=non-SPED), ELL status (“ELL”; 1=ELL, 0=non-ELL), Hispanic ethnicity (“hispanic”; 1= Hispanic, 0=Not Hispanic), 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. All analyses were conducted separately by grade using the statistical software package R 3.6.2.

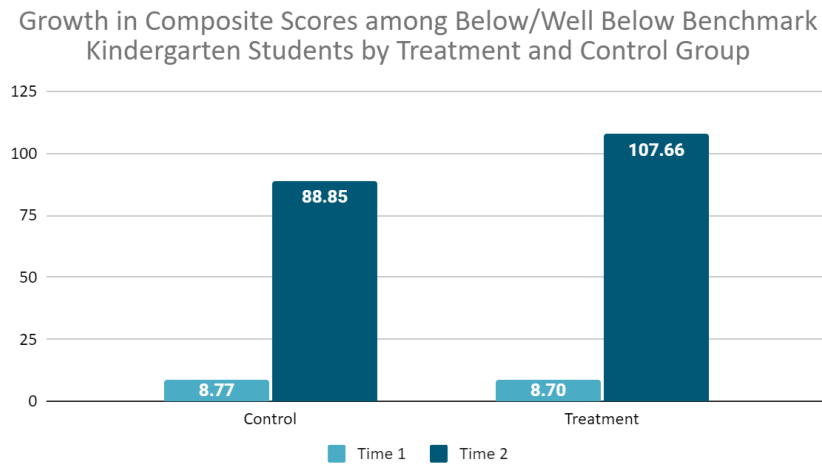
### BOY-MOY Statistical Results

Within the Kindergarten well below/below benchmark sample, we examined growth in Composite, FSF and LNF scores. Because the scores were highly positively skewed counts, we elected to use a poisson distribution to examine changes in scores overtime. There was a significant effect of treatment on composite (IRR=1.22,  $p < .001$ ,  $f^2 = .00$ ) scores (see Figure 1) and FSF (IRR=1.21,  $p < .001$ ,  $f^2 = .00$ ) scores (see Figure 2) with students in the treatment group demonstrating more growth in composite and FSF scores than students in the comparison group. There was not a significant effect of treatment on LNF scores, suggesting that students in the treatment and comparison group demonstrated similar growth in LNF. 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 comparison groups can be found in [Appendix 2](#).

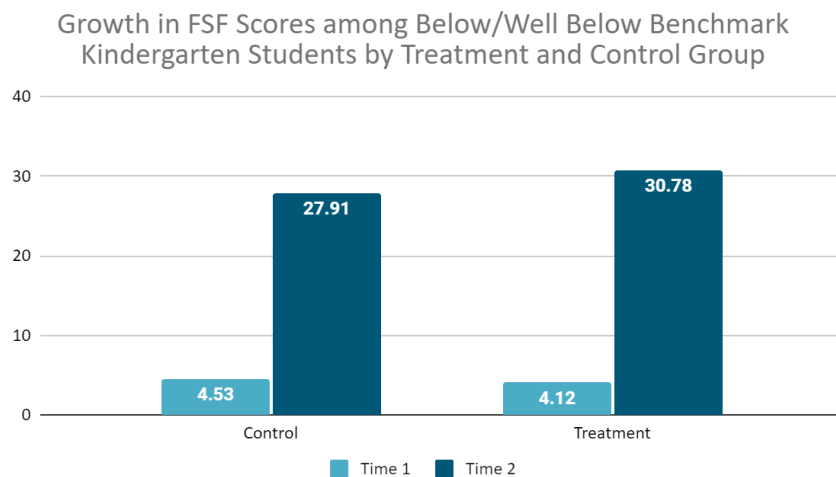
## Key Findings

### Gains on Overall Composite Scores and FSF Scores

Kindergarten students in the treatment group demonstrated significantly more growth in Composite scores than students in the comparison group



Kindergarten students in the treatment group demonstrated significantly more growth in FSF scores than students in the comparison group



### Change of Composite Levels

Using PA Lessons made a substantial impact on student growth for Tiers 2 & 3 over the fall. PA Lesson schools had reduced Well Below Grade Level by half (67% to 28%). Comparison schools only reduced the group by less than a third (67% to 47%) Notably, 41% of PA Lesson students were At/Above grade level mid-year, compared to only 25% in other schools.

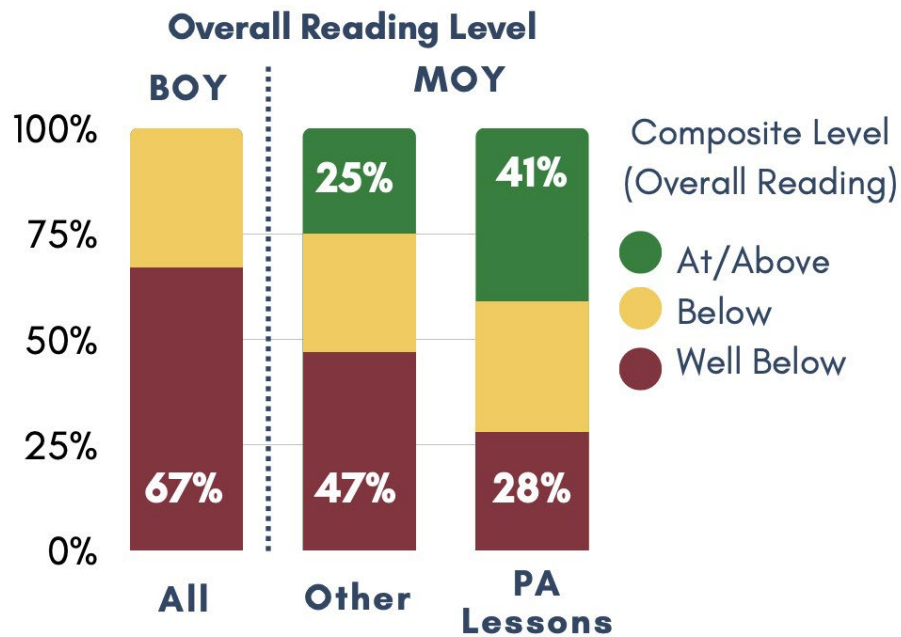


Table 7. Results for All Students

Test	School Group	BOY	MOY	Statistically Different?	Effect Size Cohen's d
Composite Scores	Comparison	8.77	88.85	Yes, the PA Lesson group had higher gains than the comparison group.	.47
	PA Lessons	8.70	107.66		
First Sound Fluency Scores	Comparison	4.53	27.91	Yes, the PA Lesson group had higher gains than the comparison group.	.33
	PA Lessons	4.12	30.78		

## Conclusion and Next Steps

The impact of the PA Lessons during its second year of use by the treatment schools resulted in twice the effect size than seen in the last BOY to MOY study (Schechter & Lynch, 2022). These large effects demonstrate that after some experience using the PA Lessons, teachers are able to use them more effectively to change student's performance levels in just a few short months. Not only did the PA Lessons group outperform the comparison group on the Acadience Reading Composite scores and the First Sound Fluency scores, the PA Lesson students Well Below grade level reduced by half.

Unlike the previous year study, there was less research oversight and data collection. The next steps for this study will be to conduct a teacher survey to learn more about what may have changed in the comparison schools and to hear from the teachers using the PA Lessons about how they are perceiving its impact. Acadience end-of-year testing is scheduled for April and May 2023.

## References

- Kini, T., & Podolsky, A. (2016). Does Teaching Experience Increase Teacher Effectiveness? A Review of the Research. *Learning Policy Institute*.
- Schachter, R. E., Flynn, E. E., Napoli, A. R., & Piasta, S. B. (2020). Teachers' perspectives on year two implementation of a kindergarten readiness assessment. *Early Education and Development, 31*(5), 778-795.
- Schechter, R. L. & Lynch, A. D. (2023). Phonological Awareness Lesson Library with Phonics Chip Kit and Phonological Awareness Screener for Intervention: Fall 2021 - Fall 2022, Grades K-1 efficacy study. Learning Experience Design (LXD) Research. [Full Report](#)

## Appendices

### Appendix 1: Kindergarten Results

For below or well below students:

- Composite score: (IRR=1.22,  $p < .001$ ) - significant differences between treatment and comparison group
- LNF score: (IRR=0.95,  $p = .26$ ) - no significant differences between treatment and comparison group
- FSF score: (IRR=1.21,  $p < .001$ ) - significant differences between treatment and comparison group

Composite Score

<i>Predictors</i>	<b>comp</b>		
	<i>Incidence Rate Ratios</i>	<i>CI</i>	<i>p</i>
(Intercept)	0.86	0.68 – 1.10	0.239
Time	10.14	9.72 – 10.58	< <b>0.001</b>
female	1.13	1.01 – 1.25	<b>0.030</b>
hispanic	0.92	0.78 – 1.08	0.299
ELL	0.83	0.74 – 0.93	<b>0.001</b>
sp	0.20	0.13 – 0.31	< <b>0.001</b>
foshom	0.98	0.86 – 1.11	0.705
intervention	0.81	0.62 – 1.06	0.132
Tigr	1.22	1.15 – 1.30	< <b>0.001</b>
<b>Random Effects</b>			
$\sigma^2$	0.02		
$\tau_{00}$ X.95ID:SchoolName	0.35		
$\tau_{00}$ SchoolName	0.03		
ICC	0.95		
N X.95ID	520		
N SchoolName	11		
Observations	1037		
Marginal R <sup>2</sup> / Conditional R <sup>2</sup>	0.790 / 0.989		

## LNF Scores

<i>Predictors</i>	<b>LNF</b>		
	<i>Incidence Rate Ratios</i>	<i>CI</i>	<i>p</i>
(Intercept)	0.70	0.54 – 0.91	<b>0.007</b>
Time	6.03	5.66 – 6.43	<b>&lt;0.001</b>
female	1.02	0.90 – 1.15	0.770
hisp	0.90	0.75 – 1.09	0.283
ELL	0.81	0.71 – 0.92	<b>0.001</b>
sp	0.38	0.24 – 0.61	<b>&lt;0.001</b>
foshom	0.94	0.82 – 1.09	0.448
intervention	1.15	0.88 – 1.50	0.316
Tigr	0.95	0.87 – 1.04	0.258
<b>Random Effects</b>			
$\sigma^2$	0.08		
$\tau_{00}$ X.95ID:SchoolName	0.43		
$\tau_{00}$ SchoolName	0.02		
ICC	0.85		
N <sub>X.95ID</sub>	520		
N <sub>SchoolName</sub>	11		
Observations	1039		
Marginal R <sup>2</sup> / Conditional R <sup>2</sup>	0.606 / 0.943		



FSF Scores

<i>Predictors</i>	<b>FSF</b>		
	<i>Incidence Rate Ratios</i>	<i>CI</i>	<i>p</i>
(Intercept)	0.74	0.56 – 0.96	<b>0.023</b>
Time	6.16	5.80 – 6.54	<b>&lt;0.001</b>
female	1.19	1.06 – 1.34	<b>0.004</b>
hisp	0.86	0.72 – 1.03	0.107
ELL	0.84	0.75 – 0.96	<b>0.008</b>
sp	0.16	0.10 – 0.26	<b>&lt;0.001</b>
foshom	1.05	0.91 – 1.21	0.516
intervention	0.75	0.56 – 1.01	0.058
Tigr	1.21	1.11 – 1.32	<b>&lt;0.001</b>
<b>Random Effects</b>			
$\sigma^2$	0.07		
$\tau_{00}$ X.95ID:SchoolName	0.41		
$\tau_{00}$ SchoolName	0.03		
ICC	0.87		
N <sub>X.95ID</sub>	520		
N <sub>SchoolName</sub>	11		
Observations	1039		
Marginal R <sup>2</sup> / Conditional R <sup>2</sup>	0.664 / 0.956		

## Appendix 2: Kindergarten t-test Results

In the table below we report effect sizes (Cohen's d) resulting from dependent samples t-tests that compared growth in composite scores and average beginning of year and middle of year composite scores in the treatment and comparison groups.

T-tests were run for Composite Scores

Grade	Condition	Number of students	Average difference in Composite between BOY and MOY	SD	Significance	Effect Size Cohen's d
Kindergarten	Comparison	264	80.59	43.41	p<.001	.47
	Treatment	254	101.77	46.18		

Grade	Condition	Number of students	BOY Composite Average	SD	Significance	Effect Size Cohen's d
Kindergarten	Comparison	266	8.80	7.95	p=.86	.02
	Treatment	255	8.92	7.91		

Grade	Condition	Number of students	MOY Composite Average	SD	Significance	Effect Size Cohen's d
Kindergarten	Comparison	264	89.45	47.05	p<.001	.44
	Treatment	254	110.71	48.63		

T-tests were run for First Sound Fluency (FSF)

Grade	Condition	Number of students	Average difference in FSF between BOY and MOY	SD	Significance	Effect Size Cohen's d
Kindergarten	Comparison	266	23.73	12.44	p<.001	.33
	Treatment	254	28.32	15.41		

Grade	Condition	Number of students	BOY FSF Average	SD	Significance	Effect Size Cohen's d
Kindergarten	Comparison	266	4.66	6.09	p=.58	.05
	Treatment	255	4.37	5.88		

Grade	Condition	Number of students	MOY FSF Average	SD	Significance	Effect Size Cohen's d
Kindergarten	Comparison	266	28.38	14.00	p=.001	.29
	Treatment	254	32.70	16.05		