Eye on Evaluation

DATA AND ACCOUNTABILITY DEPARTMENT

Letterland Evaluation: 2013-14
Author: Colleen Paeplow, Ph.D.

Abstract

In 2013-14, Letterland had strong implementation, with moderate to high fidelity within approximately 90% of WCPSS K-1 classrooms. The impact of Letterland on students' reading achievement was neutral to positive. A significantly higher percentage of WCPSS kindergarten students were at or above benchmark mid-year on Phoneme Segmentation Fluency (PSF) and Nonsense Word Fluency-Correct Letter Sounds (NWF-CLS) than matched students from a comparison school district. WCPSS students' scores remained significantly higher on the end-of-year NWF-CLS. Participation in Letterland had a significant positive effect for limited English proficient (LEP), Asian, Black, and Hispanic /Latino kindergarten students varying by indicator and benchmark period. The percentage of kindergarten students with an end-of-year PSF score at or above benchmark increased 8.2 percentage points from 2011-12 to 2013-14. Results for NWF-CLS were inconsistent across years. Results suggest Letterland is a promising approach and should be continued.

Executive Summary

Letterland, a phonics-based approach to teaching reading, writing and spelling to students in K-2 classrooms, was implemented in Wake County Public School System's (WCPSS) kindergarten classrooms in 2012-13, grade 1 classrooms in 2013-14, and grade 2 classrooms in 2014-15 (see Figure 1). As part of the WCPSS balanced literacy approach, Letterland explicitly teaches the phonetic patterns based on the six syllable types in the English language. Student participation in phonic story logic and play is integrated with language by using alliteration, rhythm, and rhyme to improve student retention of letter shapes and sounds.

Figure 1
Letterland Implementation Schedule, 2011-12 to 2014-15

2011-12
Training offered to all K teachers

2012-13
• Kindergarten Implementation
• Training offered to new K teachers
• Training offered to all grade 1 teachers

2013-14
• Grade 1 Implementation
• Training offered to new K-1 teachers
• Training offered to all grade 2 teachers

2014-15
Grade 2 Implementation
Training offered to new K-2 teachers

Data Source: Program information provided by WCPSS Academic department staff

Table of Contents

5625 Dillard Drive, Crossroads I, Cary, NC 27518
http://www.wcpss.net

WAKE COUNTY PUBLIC SCHOOL SYSTEM

WCPSS has provided Letterland as a vehicle for teaching the Common Core standards because its systematic, explicit and multi-sensory structure aligns with best practices in reading research. Letterland's design engages students in developing the critical foundational skills and vocabulary needed for reading and writing. Therefore, all kindergarten and grade 1 teachers are expected to:

- Use Letterland as the word work resource within Daily CAFÉ[1].
- Teach the Letterland whole group lessons to all students during word work instruction, ensuring students are being taught the Reading Foundation and Language standards for their respective grade level.
- Differentiate activities from Letterland based on their assessment data within small group and independent work portion of the Daily CAFÉ structure.

Additionally, Special Education Cross Categorical Resource (CCR) teachers, Intervention teachers, and ESL teachers have received training in the Letterland Intervention strand to provide more targeted, strategic interventions for students that they serve.

The initial cost of Letterland was approximately $350,000 per year (2011-12, 2012-13, and 2013-14) during its rollout to WCPSS K-2 classrooms.[2] The vast majority of this cost was instructional materials needed for each teacher within a grade level. A significant portion of the cost also went to providing each teacher a substitute so that they could attend training. Continuing material expenses are expected to cost the district approximately 10% of the initial amount since materials will only need to be purchased for new classrooms.

Results

In 2011-12, WCPSS staff formed a cross-departmental team designed to provide training to teachers and support the implementation of Letterland within K-2 classrooms. The Letterland implementation team set goals related to implementation and outcomes. Short-term goals related to providing training and resources to school staff and to developing an implementation checklist were all met. Some intermediate and long-term goals were met, but others were not. Table 1 summarizes this study's research questions, the implementation team's program goals, and related findings. Overall the results of this study suggest that Letterland was implemented with moderate to high fidelity within WCPSS K-1 classrooms in 2013-14. The impact of Letterland on students' reading achievement as measured by Phoneme Segmentation Fluency (PSF) and Nonsense Word Fluency Correct Letter Sounds (NWF-CLS) were neutral to positive.

- Relative to comparison students in another large school district not implementing Letterland, PSF scores were significantly higher in WCPSS mid-year but not at the end of the year in 2013-14. On the other hand, in 2013-14 a significantly higher percentage of WCPSS kindergarten students were at or above benchmark on the NWF-CLS at the middle-of-year (MOY) and end-of-year (EOY) compared to matched students from the comparison school

---

[1] WCPSS utilizes the Daily CAFÉ model to facilitate the provision of all elements of daily reading instruction.
[2] As of 2013-14, four elementary schools opted not to implement Letterland.

district. In other words, Letterland provided WCPSS kindergarten students an initial advantage in their literacy skills on both measures. By the end of the 2013-14 school year, matched students in the comparison school district caught up with WCPSS kindergarten students' PSF scores, but NWF-CLS scores remained statistically significantly higher for WCPSS students.

- WCPSS trends for kindergarten students' showed that EOY PSF scores increased more than eight percentage points from 2011-12 (prior to Letterland implementation) to 2013-14 (second year of implementation). NWF-CLS trends were less consistent—increasing in 2012-13 and settling back to near baseline in 2013-14.

Recommendations

Given the overall Letterland implementation and initial student outcome data, the results of this study support the continuation of this approach. Observation data from spring 2014 should be examined more closely for possible weak areas of implementation for follow-up. In addition, it is recommended that the implementation team monitor classroom implementation quality more closely to identify differences in teachers with stronger and weaker outcomes. Furthermore, examining literacy outcomes in grades 2 and 3 for the first cohort of students who experienced Letterland is recommended to investigate whether the improvements in foundational skills in kindergarten and grade 1 persist and translate into better reading comprehension performance in later grades.

Table 1
Summary of Research Questions and Program Goals with Study Results

Research Question: Did WCPSS K-1 teachers implement Letterland with fidelity in 2013-14?	89% of kindergarten teachers and 92% of first grade teachers implemented with moderate to high fidelity with approximately 40% of teachers implementing with high fidelity.
Intermediate Goal: 100% of teachers will implement Letterland in 2013-14.	Goal not fully met, but implementation was strong.
Research Question: Was the level of implementation in 2013-14 related to student reading achievement?	A relationship between fidelity of implementation and reading achievement was not demonstrated. This may be because implementation was moderate to high across nearly all teachers on the scale used.
Research Question: Did WCPSS kindergarten students participating in Letterland in 2013-14 have stronger reading outcomes compared to matched students in the comparison school district?	Reading achievement results were neutral to positive. Mid-year scores on PSF and NWF-CLS as well as end-of-year scores on NWF-CLS were statistically significantly higher for WCPSS students as compared to comparison students; however, there was no difference in PSF EOY scores. WCPSS did have fewer students with scores well below grade level than did the comparison school district.
	Participation in Letterland had a significant positive effect for LEP, Asian, Black, and Hispanic/Latino kindergarten students varying by indicator and benchmark period.
Research Question: What were the reading achievement results prior to Letterland implementation compared to year one and year two of implementation?	In kindergarten, the percentage of students who were at or above benchmark on the PSF increased with each year of Letterland implementation, with the percentage who were at or above benchmark on their PSF EOY score increasing 8.2 percentage points from spring 2012 to spring 2014. The percentage of students well below benchmark decreased from 11% to 6%.
Intermediate goal: Higher percentages of kindergarten students will score at or above benchmark by spring of 2013. **Long-term Goal:** Higher percentages of kindergarten students will score at or above benchmark by spring of 2014.	Results for the percentage of kindergarten students who were at or above benchmark on NWF-CLS were less consistent: increasing in 2012-13 and decreasing in 2013-14.
	The intermediate goal was met. However, the long-term goal was not met. The adoption of DIBELS Next may have contributed to a lower percentage of students at or above grade level in the spring of 2014.
Research Question: Did WCPSS K-1 reading achievement results vary by student subgroup and grade?	**Kindergarten:** On PSF EOY, percent on benchmark for all ethnic subgroups considered gradually increased across the first two years of implementation. On NWF-CLS EOY, percent on benchmark for all ethnic subgroups increased after the first year of implementation and decreased during the second year. **Grade 1:** On NWF-CLS, percent on benchmark for all ethnic sub-groups' except Asian students increased following the first year of implementation (2013-14).

Background

North Carolina's Excellent Public Schools Act of 2012, which requires 3rd grade students to be proficient on a state-approved standardized test of reading comprehension (i.e., End-of-Grade reading exam), has amplified the imperative that North Carolina's students are reading on grade level by the end of grade 3 (North Carolina Read to Achieve, 2013). Furthermore, it is an expectation that all students be taught the English Language Arts (ELA) Common Core State Standards for their respective grade levels. WCPSS utilizes the Daily CAFÉ model to facilitate the provision of all elements of reading instruction on a daily basis. In kindergarten and grade 1, the Common Core places a strong emphasis on teaching the critical foundational skills for literacy as outlined in the Reading Foundation and Language strands. Letterland represents WCPSS' efforts to ensure K-2 students gain the early foundational literacy skills needed to be successful readers by grade 3 and throughout their educational careers.

Letterland is a phonics-based early literacy instructional program, which incorporates student interaction through participation in phonic story logic and play with language through alliteration, rhythm, and rhyme. The program is designed to make the task of remembering shapes and sounds of letters easier for students (Ehri & McCormick, 1998). Letterland's phonics-based approach is consistent with the National Reading Panel's findings that the best approach to reading instruction includes explicit instruction in phonemic awareness (PA) and systematic phonics instruction[3]. According to the National Reading Panel's analysis the best approach to reading instruction includes: overt instruction in PA, systematic phonics instruction, efforts to improve fluency, and use of strategies to improve reading comprehension (The US National Reading Panel, 2000). Letterland is designed to improve students' skills in PA, phonics, and executive function. Although reading fluency and comprehension are components of the WCPSS balanced literacy program, these components are not directly targeted with Letterland instruction and are not addressed in this report.

Previous studies have found some evidence to support the use of Letterland in K-2 classrooms to improve students' reading outcomes (Felton & Crawford, 2010; Wendon, 2010). Felton and Crawford (2010) found a significant decrease in the number of students classified as 'at risk' based on Dynamic Indicators of Basic Early Literacy (DIBELS) indicators after students received Letterland instruction for three years. Furthermore, the number of students classified as 'at risk' progressively declined each year that Letterland was implemented.

Phonemic Awareness

Letterland is designed to improve students' PA, which is defined as students' ability to identify and blend phonemes into words. PA refers to the understanding that spoken words can be sub-divided into phonemes or smaller segments of sound. In order to understand PA, teachers typically ask students to isolate phonemes, identify phonemes, or categorize phonemes. For example, a teacher might ask a student, "Which word does not belong - *bus, bun, rug*?" Here,

---

[3] In 1997, a National Reading Panel was assembled to evaluate existing research and evidence to find the best ways of teaching children to read. The Panel reviewed and analyzed several hundred studies from the over 100,000 considered (National Institute of Child Health and Human Development, 2000).

students are required to recognize the word *rug* is the odd sound in a sequence of words (phonemic categorization). Students might also be asked to recognize individual sounds in words. For example, students might be asked to identify the first phoneme or sound in the word "*paste*," in which case students should identify the sound /p/ (phoneme isolation). PA is part of the more encompassing term, phonological awareness instruction (The US National Reading Panel, 2000). Phonological awareness refers to awareness of larger spoken units including syllables and rhyming words, which is also specifically targeted with Letterland instruction. Students might be asked to participate in rhyming exercises and to break sentences into words and words into syllables before they are taught to segment phonemes in words. However, identification and manipulation of sounds at the phonemic level contributes most to helping children learn to read, particularly when phonemes are taught with letters (US National Reading Panel, 2000).

Effective PA instruction helps students break apart and manipulate the sounds in words. Prior research suggests that students' levels of PA in kindergarten predict how well they learn to read during their first two years of school and beyond (Ehri et al., 2001; Gillon, 2004; Stahl & Murray, 1998; The US National Reading Panel, 2000). Based on a meta-analysis of 96 cases that compared the outcomes of treatment and control groups, instructional programs focused on PA, when compared with other types of reading instruction, were associated with increases in:

- students' PA with a large average effect size of $d = .86$ (i.e. Cohen's d[4]), hereafter denoted as d,
- students' overall reading outcomes with a moderate average effect size of .53,
- students' spelling ability with a moderate average effect size of .59,
- students' decoding skills with a moderate effect size of .56, and
- students' reading comprehension with a small effect size of .34 (Ehri, Nunes, Willows, Schuster, Yaghoub-Zadeh, & Shanahan, 2001).

Various subgroups of students increased in levels of PA following PA instruction including: students identified as at-risk of having low reading outcomes, students not identified as at-risk, students with special needs, students with a Limited English Proficient (LEP) status, and students from various socio-economic backgrounds. Overall, the effect of instruction focused on PA on reading outcomes was large in kindergarten ($d = .95$) and small in grade 1 ($d = .48$). Therefore, this research suggests that PA instruction positively contributes to students' early literacy outcomes, especially in kindergarten (Ehri et al., 2001).

---

[4] Cohen's d is a measure of the standardized difference between means. By finding the difference between means and dividing by the standard deviation the difference is standardized (Denis, 2012). Cohen (1988) outlined criteria for gauging effect sizes: small $d = .20$, medium $d = .50$ and large $d = .80$.

Alphabetic Principle (Phonics)

Systematic phonics instruction teaches students to recognize letter-sound correspondences and spelling patterns. Specifically, phonics is the knowledge that letters of the alphabet represent phonemes or sounds that are blended together to form written words. Readers who are skilled in phonics can sound out words they have not seen before, without first having to memorize them (The US National Reading Panel, 2000). Knowledge of the alphabetic system greatly contributes to students' ability to read unfamiliar words in isolation. For example, children may be asked to read pseudowords or nonsense words (e.g., *gan, sig, rav, trusk)* in order to assess their understanding of letter-sound correspondence. Letterland is designed to improve students' ability to blend letter knowledge (e.g., letter name and shape) and PA (sound) to decode written words accurately and automatically using materials that are high-interest and motivational for students. For example, letters are embedded in characters using pictograms and stories, which engages students by providing a visual cue to remember the letter shape and an auditory cue to recall the sound (Letterland International, 2014).

Examining 66 treatment-control comparisons, The US National Reading Panel (2000) found a moderate mean overall effect size produced by phonics instruction ($d = .44$), with a moderate mean effect size for kindergarten students ($d = .56$) and a moderate mean effect size for grade 1 students ($d = .54$), providing support that systematic phonics instruction contributes to students' reading growth when compared with unsystematic or no phonics instruction, particularly in kindergarten and grade 1. Mean effect sizes for phonics instruction were higher for at-risk students in kindergarten ($d = .58$) and grade 1 ($d = .74$).

Executive Function

Executive function skills are inter-related cognitive processes required for goal-directed behavior including memory, attention, and mental flexibility. Emphasis on executive function skills during reading instruction is linked to positive reading outcomes especially for K-1 students (Cartwright, 2012). Because reading is complex and cognitively-demanding, students benefit from support in mentally managing the reading process (Cartwright, 2012). Letter learning requires retaining shapes, names, and sounds in memory and retrieving that information automatically in reading and writing words (The US National Reading Panel, 2000). Letterland is designed to increase automaticity in letter knowledge and PA, which develops students' executive function skills as they read. Brain wave research provides evidence that Letterland promotes the development of executive function skills in preschool-age children. According to Wendon (2010) three and four year olds who were exposed to Letterland activated more of their brain when reading compared to students exposed to more traditional reading programs. Increased brain activity is associated with more developed executive function skills (Cartwright, 2012). The results were found to persist beyond a period of six months, even after instruction had discontinued (Wendon, 2010). These results suggest that students exposed to Letterland are better able to regulate the multiple cognitive processes involved in reading, including attention, memory, language processing, and visual processing (Cartwright, 2012; Wendon, 2010).

Prior research conducted specifically on Letterland was either hosted exclusively on the provider's website (i.e., Letterland's website) or was descriptive in nature. This study provides an independent examination of literacy outcomes for students participating in Letterland using a matched group (quasi-experimental) design. The next section of this report describes the research design and the data used.

Methods

Study Questions

Questions addressed in this evaluation pertained to the level of implementation of Letterland, the attainment of goals set by the literacy staff as benchmarks for success, and the effect of Letterland on reading achievement. The research questions and data sources for this study are outlined in Table 2.

Table 2
Key Evaluation Questions and Data Sources

Research Question	Data Source
Did WCPSS K-1 teachers implement Letterland with fidelity in 2013-14?	2013-14 Student Locator; Letterland fidelity data; and 2013-14 mCLASS® data.
Was the level of implementation in 2013-14 related to student reading achievement as measured by DIBELS data during the 2013-14 academic year?	2013-14 Student Locator; Letterland fidelity data; and 2013-14 mCLASS® data.
Did WCPSS kindergarten students participating in Letterland in 2013-14 have stronger reading outcomes compared to matched students in the comparison school district that was not implementing Letterland?	2013-14 Student Locator, 2013-14 mCLASS® data and demographic and 2013-14 mCLASS® data from comparison school district.
What were the reading achievement results, as measured by DIBELS data, prior to Letterland implementation compared to year one and year two of implementation?	2011-12 through 2013-14 Student Locator files, 2011-12 through 2013-14 mCLASS® and AIMSweb data. DIBELS data were obtained from AIMSweb and mCLASS® in 2011-12 and 2012-13 and from mCLASS® in 2013-14.
Did WCPSS K-1 reading achievement results vary by student subgroup and grade?	2011-12 through 2013-14 Student Locator files, 2011-12 through 2013-14 mCLASS® and AIMSweb data.

Data Sources

This study included DIBELS data retrieved from AIMSweb and mCLASS[®]; student locator data files; and Letterland fidelity data.

- DIBELS assessments were designed to measure early literacy skills such as letter naming fluency, phonemic segmentation fluency, nonsense word fluency, and oral reading fluency[5] (Lenard, 2013). DIBELS data were obtained from AIMSweb and mCLASS[®] in 2011-12 and 2012-13 and from mCLASS[®] in 2013-14.[6] 2013-14 DIBELS data for the population of 12,728 kindergarten students in the large comparison school district were obtained from that district. WCPSS implemented mCLASS[®] system-wide in 2013-14; teachers were expected to record students' DIBELS scores electronically via iPads (new devices).
- Student locator files from 2011-12, 2012-13, and 2013-14 were used to capture WCPSS student demographics and the comparison school district provided student demographic data.
- The Letterland fidelity data were collected during the spring of 2014 via a Google form (see Appendix A). Principals were asked to conduct observations within their kindergarten classrooms utilizing the Letterland fidelity instrument. Observations were conducted by the principal or an assigned school staff member with the goal of observing each kindergarten classroom at least once. Many schools also used the instrument to observe implementation within their grade 1 classrooms; these results are included in Appendix B by school and grade.

Study Design

This study examined both the implementation and impact of Letterland. Literacy outcomes were examined utilizing both descriptive and inferential analyses. To test the impact of Letterland on early reading achievement, WCPSS students participating in Letterland were compared to matched students in a comparison school district that was not using Letterland during the 2013-14 academic year. The district-wide rollout of Letterland in WCPSS prohibited a pilot with a random research design; thus, data from a comparison school district were requested and a student level match of WCPSS kindergarten students to students within the comparison school district was conducted making this study a quasi-experimental design. Descriptively, we assessed whether district goals associated with Letterland were met by comparing DIBELS results prior to Letterland implementation and in year one and year two of implementation in the context of DIBELS Next benchmarks (see Table 3).

---

[5] R. H. Good is credited as the developer of DIBELS measures while the University of Oregon Center on Teaching and Learning contributed to the creation of the DIBELS Data System.
[6] In 2011-12 and 2012-13, 15 schools used mCLASS[®] and the remainder of schools used AIMSweb to enter and access their DIBELS data. In 2013-14, North Carolina's Department of Public Instruction provided mCLASS[®] to all NC schools.

Table 3
Nature of the Study and Valid Conclusions

Research Design	Conclusions that Can be Drawn
☐ Experimental	Program caused the changes identified.
☑ Quasi-Experimental	Program is correlated with changes found (but it is possible that other uncontrolled factors influenced the results).

☐ Descriptive
  ☐ Qualitative
  ☐ Quantitative
	Provides outcome data for the program, but differences cannot be attributed directly to the program because there is no control of other influences. Describes trends which may be actionable and/or lead to changes which can be tested with future research studies.

Fidelity of Implementation

School administrators and/or their appointed staff conducted 1,067 kindergarten and grade 1 classroom observations during the spring of 2014. Some school staff opted to conduct multiple observations within a classroom in order to provide teachers feedback and thereby strengthen the implementation of Letterland. The fidelity checklist used for this evaluation was developed by a school principal outside of WCPSS and shared by Letterland staff (see Appendix A). WCPSS staff created at Google form from the checklist in order to collect observations electronically. A picture cue card was also provided to clarify Letterland concepts and/or activities that were included on the checklist. Members of WCPSS' Letterland Implementation Team tested a more detailed fidelity instrument and determined that in-depth knowledge of Letterland would have been required to administer that instrument. Therefore, in order to enable school administers to collect classroom observations across the district, a simplified fidelity checklist was selected.

The median observation score for each teacher was used to create one fidelity score for each of the 785 kindergarten and grade 1 teachers observed. The teacher fidelity scores were analyzed within a multi-level model (students nested in classrooms nested in schools) to determine the significance of the level of implementation on kindergarten students' reading achievement as measured by PSF and NWF-CLS end-of-year (EOY) scores. The model included student-level variables such as First Sound Fluency (FSF) as a pre-assessment, ethnicity, LEP status and Students with Disabilities (SWD) status. The fidelity score was assigned at the classroom level and the percentage of students receiving free or reduced-price lunch (FRL) was included as a school level variable.

Comparative Analysis

A large school district with a similar student population was contacted and an external data request was made to elicit their kindergarten DIBELS data for 2013-14. Although the comparison school district did not use Letterland district-wide, they used it with their students with disabilities (SWD). We therefore removed SWD students from the matched group analysis. Staff from the comparison school district reported that they had been shifting from a managed

instruction basal program to a more balanced approach to literacy over the last three to four years.[7] They provided schools with a few options to meet their literacy needs including the Columbia Teacher's College Reading and Writing Project Reading and Writing Units of Study; Making Meaning by Developmental Studies Center; Imagine It! (the green band which is the phonics portion of the managed basal program); Month by Month Phonics by Pat Cunningham; Words Their Way by Donald Bear et. al.; and various other school-based programs. After demonstrating acceptable overlap and balance between WCPSS and the comparison school district, outcomes on reading indicators for kindergarten students in WCPSS and the comparison school district were compared. An analysis of variance (ANOVA) significance test was conducted to test the significance of any differences between matched groups.[8] Regression analyses were conducted to determine the size of the impact of program participation on reading indicators.

Propensity Score Matching

WCPSS' kindergarten students participating in Letterland in 2013-14 (n = 11,379) were matched to students enrolled in the large comparison school district (n = 12,059) via a propensity score matched-pair sample using one-to-one greedy matching technique (Dehejia & Wahba, 2002). Overall, 7,739 students from each school district were matched using this matching procedure. A logistic regression model was fit to generate a propensity score for each student in WCPSS and the comparison school district. Students were assigned a probability score based on the matching criteria, and then matched based on this score. The matching criteria used for this study were beginning-of-year (BOY) FSF score, ethnicity, LEP status, and gender. This matching technique attempts to mimic random assignment by creating a control group that is comparable to the study group, and is superior to simple matching procedures that may require weighting of matching criteria (Dehejia & Wahba, 2002). A greedy matching technique was selected as the matching procedure because using this method, a large number of WCPSS and comparison students were retained in the sample, and the samples demonstrated adequate overlap and balance (see Appendix D for more details). Descriptive information for kindergarten students in WCPSS and the comparison school district prior to and after matching are displayed in Table 4. After propensity score matching, these descriptive data suggested that the samples are comparable based on observable covariates. Table 4 includes the entire population of students in each district regardless of whether or not students have a test score at each assessment period. The table provides evidence that the students are well-balanced across districts on the covariates and therefore, comparisons on outcomes between students in the school districts are supported.

---

[7] The basal reading approach refers to teaching basic reading skills utilizing leveled text or reading books.
[8] Analysis of variance (ANOVA) and the associated F-statistic tests the statistical significance of the difference of two or more means (Salkind, 2011).

Table 4

Percentage of Matched Kindergarten Students by Demographic Characteristic, 2013-14

Subgroup	WCPSS Prior to Matching		Comparison District Prior to Matching		T-test Prior to Matching		WCPSS After Matching		Comparison District After Matching		T-test After Matching	
	%	#	%	#	t	Sig	%	#	%	#	t	Sig
Asian	7.4	845	6.7	802	2.4	0.02	6.9	536	6.9	536	0.0	1.00
Black	22.3	2,536	38.4	4,632	-28.2	0.00	28.7	2,223	28.7	2,223	0.0	1.00
Hispanic/Latino	19.1	2,170	22.8	2,750	-6.9	0.00	21.8	1,687	21.8	1,685	0.0	0.97
Multiracial	4.0	452	1.7	206	10.7	0.00	2.0	156	2.0	156	0.0	1.00
Other	0.4	48	0.6	73	-2.2	0.03	0.4	34	0.5	37	-0.4	0.72
White	46.7	5,312	29.8	3,594	27.9	0.00	40.1	3,101	40.1	3,101	0.0	1.00
LEP	13.8	1,574	16.05	1,935	-4.5	0.00	14.5	1,121	14.5	1,121	0.0	1.00
Male	50.2	5,709	50.0	6,032	-0.9	0.39	50.5	3,909	50.5	3,909	0.0	1.00
Female	49.7	5,657	50.0	6,027	-0.9	0.39	49.5	3,830	49.5	3,830	0.0	1.00
All	**100.0**	**11,379**	**100.0**	**12,059**			**100.0**	**7,739**	**100.0**	**7,739**		

Note: 1. Data do not include students with SWD status.
        2. Racial subgroups with small numbers of students are not shown.
Data Sources: WCPSS data: 2013-14 Student Locator; comparison school district data provided by comparison school district

Reading Achievement Trends Across Time

DIBELS results prior to Letterland implementation and in year one and year two of implementation were compared in the context of DIBELS Next benchmarks by grade and student subgroup.

In 2013-14, the DIBELS benchmarks changed following the district's adoption of DIBELS Next. Prior to DIBELS Next, the district implemented DIBELS 6th Edition, which included similar but slightly different measures and benchmark cut scores. Based on our communication with the DIBELS developer, the scores were comparable for the NWF-CLS and PSF across DIBELS 6th edition and DIBELS Next (R. H. Good, personal communication, October 31, 2014). Based on the suggestion of the developer of the DIBELS 6th Edition and DIBELS Next, we used the DIBELS Next benchmark cut scores to describe the scores that were collected using the DIBELS 6th Edition measures. Across years (2011-12 through 2013-14), percentage of students on benchmark and mean scores overall and for subgroups were examined for PSF and NWF-CLS at MOY and EOY for kindergarten and grade 1 students.

Student Demographics

The demographic characteristics of students who participated in Letterland in 2012-13 and 2013-14 are illustrated in Table 5. The population for this study includes all WCPSS kindergarten and grade 1 students enrolled during the 2011-12, 2012-13, and 2013-14 academic years with

DIBELS assessment scores except those attending the four schools not implementing Letterland. Students missing DIBELS data were excluded from the analysis. Generally, over 90% of students had scores each year, with more complete data in 2013-14 than previously. The demographic characteristics closely mirror those of the district since nearly all students are represented.

Table 5
Student Demographic Characteristics, 2011-12, 2012-13, and 2013-14
Kindergarten and Grade 1

| | 2011-12 | | | | 2012-13 | | | | 2013-14 | | | |
| | Kindergarten | | Grade 1 | | Kindergarten | | Grade 1 | | Kindergarten | | Grade 1 | |
Subgroup	%	n	%	n	%	n	%	n	%	n	%	n
Am. Indian	0.3	29	0.2	23	0.9	21	0.3	30	0.3	40	0.2	22
Asian	7.2	766	6.7	721	7.2	840	7.1	835	7.2	862	7.3	928
Black	22.4	2,386	22.2	2,382	22.6	2,636	22.2	2,627	22.6	2,696	23.2	2,932
Hispanic/Latino	18.1	1,929	18.5	1,987	19.2	2,232	18.4	2,172	19.1	2,278	19.0	2,401
Multiracial	3.7	388	3.7	396	3.5	403	3.7	436	3.9	459	3.6	453
Pacific Islander	0.0	2	0.0	0	0.1	12	0.1	17	0.1	11	0.1	15
White	48.3	5,142	48.6	5,210	47.2	5,501	48.2	5,694	46.8	5,571	46.7	5,915
LEP	13.9	1,483	13.6	1,458	21.6	2,518	20.0	2,365	14.4	1,711	13.1	1,654
SWD	4.7	495	6.8	728	6.4	739	8.6	1,013	7.0	832	7.9	1,000
All	**100**	**10,642**	**100**	**10,719**	**100**	**11,645**	**100**	**11,811**	**100**	**11,917**	**100**	**12,666**
Missing	8.8	1,003	7.0	789	7.6	953	8.0	1,026	3.7	441	0.0	0

Data Sources:　2011-14 Student Locator, 2011-14 mCLASS® and AIMSweb data. DIBELS data were obtained from AIMSweb and mCLASS® in 2011-12 and 2012-13 and from mCLASS® in 2013-14.

Results

Figure 2
Pathway to Action
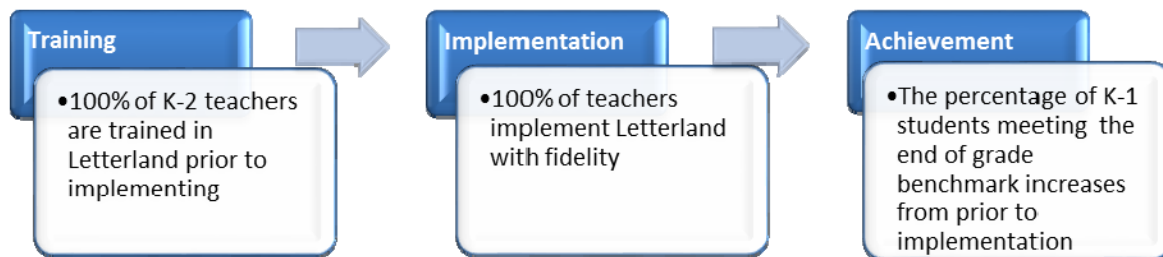
Data Source: Letterland Logic Model, 2011-12 to 2015-16 (see Appendix F)

Teachers were required to attend training prior to receiving Letterland materials; thus, 100% of K-2 teachers were trained prior to implementation. Fidelity of implementation was moderate to

high in the vast majority of WCPSS kindergarten and grade 1 classrooms observed in the spring of 2014. However, a relationship between student reading achievement and level of fidelity to the implementation of Letterland was not demonstrated. Matched student analysis found that by mid-year, a higher percentage of WCPSS kindergarten students were at or above benchmark on PSF and NWF-CLS than in the comparison school district. While WCPSS kindergarten students' NWF-CLS scores at the end of 2013-14 remained significantly higher than in the comparison district, the results for PSF were similar across the two school districts. The percentage of kindergarten students who were at or above benchmark on their PSF EOY increased with each year of Letterland implementation, but the results for NWF-CLS were not consistent across years.

Fidelity of Implementation

Figure 3 displays the distribution of fidelity scores assigned to the 785 kindergarten and grade 1 teachers observed in 2013-14 (see Appendix B for fidelity scores by grade and school). The 412 kindergarten teachers' fidelity scores ranged from zero to 10 with a mean score of 6.0; and the 373 grade 1 teachers' scores were similar ranging from one to 10 with a mean score of 6.2 on a scale of one to 10.

> **More than 89% of teachers implemented Letterland with medium to high fidelity**

Figure 3
*Percentage of Kindergarten and Grade 1 Teachers by Fidelity Score*
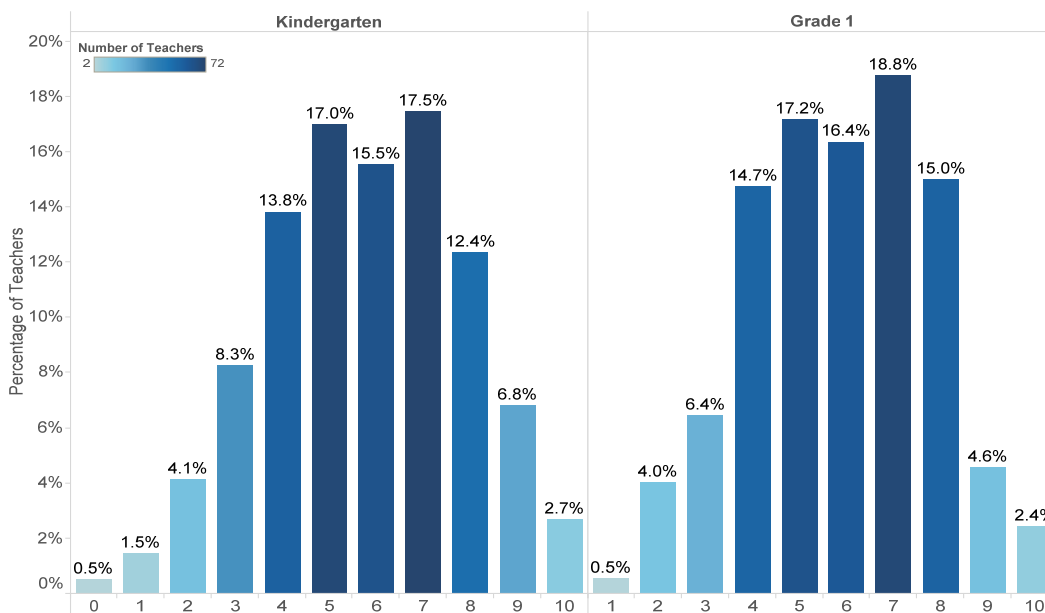
Note: Blue shading indicates the number of teachers who received each score with darker shades indicating more teachers (Number of Teachers 2 — 72)

Data Source: Spring 2014 Fidelity Checklist data

Based on fidelity scores grouped into high, medium, and low (each group representing one third of the scale), 40% to 42% of teachers of each grade received a high fidelity rating, approximately half of kindergarten and grade 1 teachers were implementing Letterland with medium fidelity, while 11 percent of kindergarten teachers and eight percent of grade 1 teachers were implementing at a low level (see Figure 4).

Figure 4
Percentage of Kindergarten and Grade 1 Teachers by Level of Fidelity
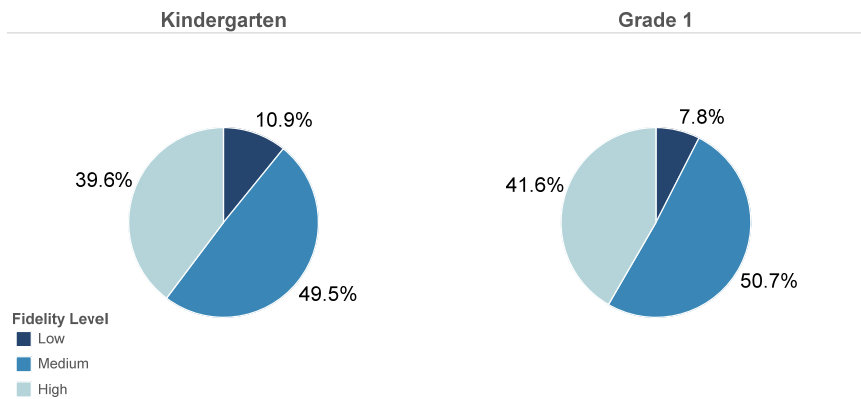
Note: Fidelity level groupings: low fidelity indicates the bottom third of the scoring range (0 to 3.33); medium indicates the middle third (>3.33 to 6.66); and high indicates the top third (>6.66 to 10)
Data Source: Spring 2014 Fidelity Checklist data

Students' PSF and NWF-CLS EOY scores were not related to their teachers' level of fidelity to the implementation of Letterland within the kindergarten and grade 1 classrooms observed during spring 2014 (see Appendix C). In order to examine the degree to which fidelity of implementation explains students' reading outcomes, a multi-level model accounting for the nesting of students in classrooms and classrooms in schools was fit. While the relationship between some student level characteristics (LEP and SWD) and reading PSF EOY and NWF-CLS EOY were statistically significant and varied between classrooms and schools, there was not enough evidence to conclude that the fidelity of implementation impacted reading achievement. It should be noted, when the model was restricted to classrooms with low fidelity scores (< 3.33) and those with high fidelity ($>= 6.66$) the impact of fidelity of implementation on reading achievement was statistically significant at the .10 level (PSF, $p = .06$ and NWF-CLS, $p = .09$). While this did not meet the significance criteria of .05 set for this study, it does provide some evidence that fidelity may impact reading achievement in some classrooms.

Cost of Implementation

The startup costs for providing Letterland materials for all K-2 classrooms and providing all K-2 teachers training were approximately one million dollars for the district. The annual per pupil cost of Letterland's K-2 rollout was approximately $15 per student served in 2012-13, 2013-14,

and 2014-15 (cost included training provided in 2011-12). However, the maintenance cost per student (based on 2014-15 expenditures at the time of this report) is projected to be less than $1 per student. The programmatic cost information regarding Letterland implementation is described in Table 6.

Table 6
Letterland Expenditures, 2011-12 to 2014-15

	Expense Category	2011-12	2012-13	2013-14	2014-15
Instructional Materials	Kindergarten	$334,529	$38,759		
	Grade 1		$280,449	$32,527	
	Grade 2			$295,264	$17,188*
Workshop - Planning & Delivery					
Presenter Costs	Kindergarten	$7,640	$3,400	$1,400	$1,200
	Grade 1		$5,000	$7,800	$1,433
	Grade 2			$9,000	$6,000
Estimated cost for substitutes	Kindergarten	$7,000	$358	$4,095	$1,502
	Grade 1		$6,414	$13,690	$956
	Grade 2			$9,783	$910
		K	**K-1**	**K-2**	**K-2**
Organization Costs		$3,333	$3,333	$5,556	$2,778
Total Expenditure		**$352,502**	**$337,713**	**$346,588**	**$31,967***

Data Source: Program information provided by WCPSS Academic Department staff

Note: *2014-15 cost reflects the mid-year expenditures at the time of this report; however, continuing costs for this program are expected to be considerably lower than those required at start up since material costs only reflect growth (i.e., new classrooms).

For the 23,562 kindergarten students who participated in Letterland in 2012-13 and 2013-14, the total cost of implementing Letterland and the associated training in 2011-12, 2012-13, and 2013-14 was $397,181. To consider the cost effectiveness of Letterland we could divide the cost per kindergarten student by the eight percentage point increase in kindergarten students at or above benchmark on the PSF EOY from 2011-12 to 2013-14.[9] This would mean each percentage point increase in kindergarten students at or above benchmark on the PSF EOY cost $2 per student (for NWF-CLS there was a decrease in the percentage of students at or above benchmark). For grade 1 students there was a small increase in the percentage of students at or above benchmark on the NWF-CLS (approximately one percentage point); thus, the cost effectiveness would be considerably higher (approximately $15 per student). However, this cost would be substantially lower when calculated based on continuing costs rather than start-up costs.

---

[9] Cost-effective analysis is a method of considering costs in terms of dollars and results in terms of an outcome of interest, e.g., student test scores (Levin & McEwan, 2001).

Matched Group Comparison

In order to access the impact of Letterland on kindergarten students' reading achievement, WCPSS students were matched to students within the comparison school district on beginning-of-year (BOY) FSF score, ethnicity, LEP status, and gender. PSF and NWF-CLS MOY and EOY scores were analyzed as outcome measures (kindergarten students are not assessed at the beginning of the year on these measures).

> **A significantly higher percentage of WCPSS kindergarten students were at or above benchmark based on EOY NWF-CLS than matched students.**

Figures 5 and 6 display the percentages of students at each benchmark level by student subgroup for PSF and NWF-CLS EOY, respectively. For each subgroup (with the exception of White students) there was a higher percentage of WCPSS students at or above benchmark than for matched comparison school district students. Overall there was not a significant difference in the percentage of students at or above benchmark on the PSF EOY in WCPSS compared to matched students. Differences in the percentage of matched students at or above benchmark by the EOY ranged by student subgroup from 1.4% lower for WCPSS' White students to 4.3% higher for WCPSS Asian students compared to matched students (see Figure 5).
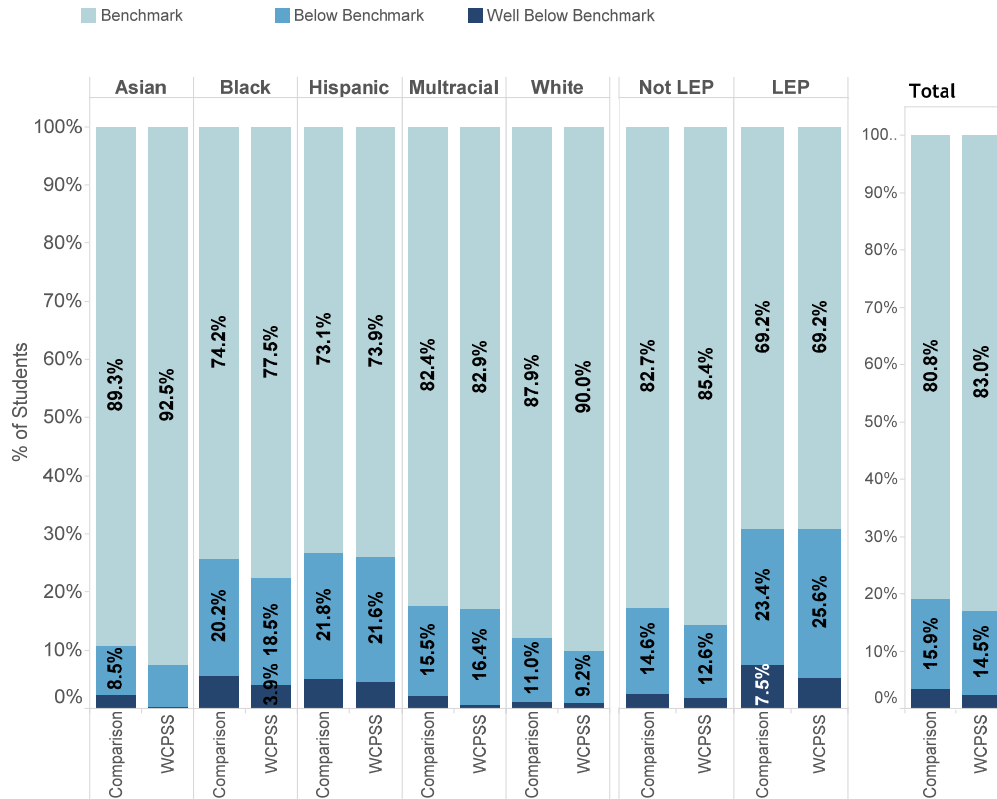
Figure 5
PSF EOY Benchmark Level by Student Subgroup for
WCPSS and Comparison District Kindergarten Students
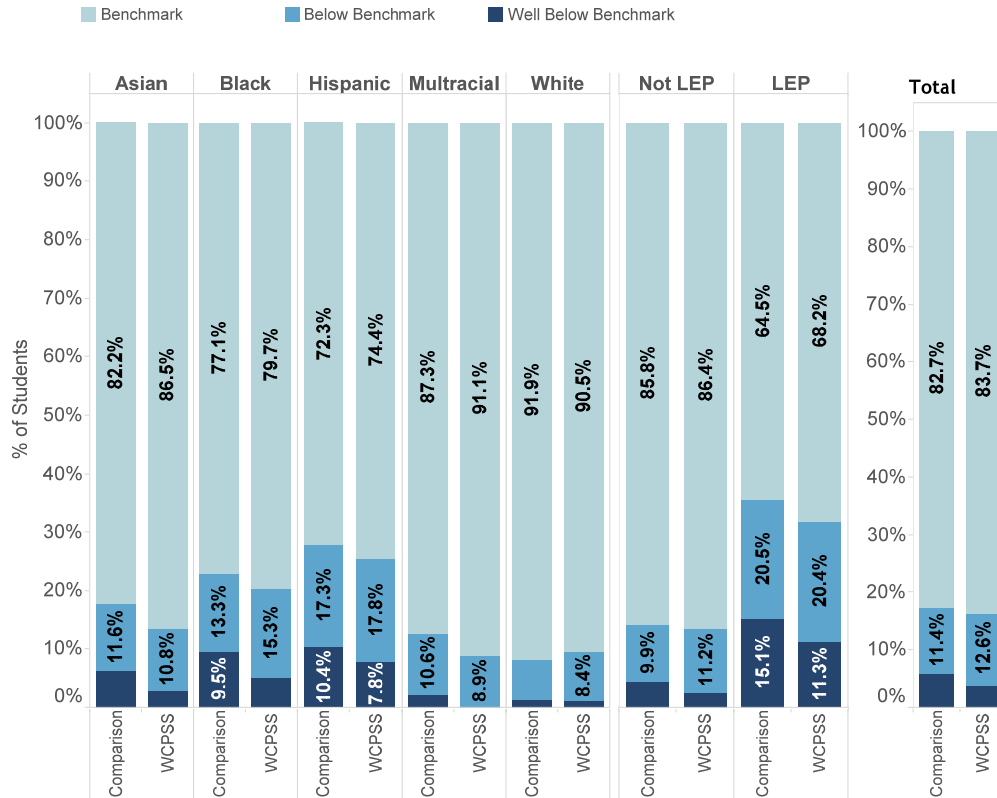
Data Sources: 2013-14 Student Locator, 2013-14 WCPSS and comparison school district mCLASS® data.

A significantly higher percentage of WCPSS kindergarten students were at or above benchmark on NWF-CLS EOY than kindergarten students at the comparison school district. Differences in the percentage of matched students at or above benchmark ranged by subgroup from no difference among LEP students to 3.3% higher for WCPSS' Black students than comparison students.

Figure 6
NWF-CLS EOY Benchmark Level by Student Subgroup for
WCPSS and Comparison District Kindergarten Students

Data Sources: 2013-14 Student Locator, 2013-14 WCPSS and comparison school district mCLASS® data.

An ANOVA was conducted to examine the difference in PSF and NWF-CLS MOY and EOY scores between matched WCPSS kindergarten students and matched comparison students. There was a statistically significant difference between the mean PSF MOY scores of WCPSS and matched comparison students; however, there was no difference in the PSF EOY means. For NWF-CLS there was a statistically significant difference in the mean for both MOY and EOY scores for WCPSS students and matched students (see Table 7). The effect of Letterland on reading outcomes would be considered small for both reading indicators, PSF MOY (d = .25), PSF EOY

(d = -.01), NWF-CLS MOY (d = .33) and NWF-CLS EOY (d = .20), based on Cohen's widely used interpretation of effect sizes.[10] Cohen's scale is not tailored to the effects of intervention studies within education and therefore may be misleading when applied to the generally smaller effect sizes found within these studies (Lipsey et al, 2012). According to Lipsey et al. (2012) effect sizes found across a range of educational interventions were rarely as large as .30.[11] Thus, if we consider the effect size for NWF-CLS with respect to educational interventions measured with a standardized test of narrow scope, the effect sizes may be considered as indicating a positive impact. Table 7 also includes the coefficients and p-values of a doubly-robust OLS regression analysis, which included the matching criteria as covariates (see Appendix D). These OLS coefficients suggest that compared with kindergarten students who did not receive Letterland, students who received Letterland scored 4.4 points higher on PSF MOY, 6.4 points higher on NWF-CLS MOY, and 5.5 points higher on NWF-CLS EOY, all of which are significant at $p < .001$. Analysis examining the interaction of participation in Letterland and student subgroup indicated a significant effect for LEP, Asian, Black, and Hispanic/Latino students varying by indicator and benchmark period (see Appendix D).

Table 7
PSF and NWF EOY Scores WCPSS and Comparison District, Kindergarten

Benchmark Period	WCPSS Students			Comparison Students			Difference in Means	F Value	P Value	Effect Size	OLS Coef.
	n	Mean Score	Standard Deviation	n	Mean Score	Standard Deviation					
PSF MOY	7,540	36.6	17.1	7,543	32.2	18.0	**4.4**	30.3	<.001	.25	4.43 p<.001
NWF-CLS MOY	7,540	34.7	20.7	7,543	28.4	17.9	**6.3**	41.4	<.001	.33	6.39 p<.001
PSF EOY	7,391	49.9	12.8	7,208	50.0	14.2	-0.1	0.4	.543	-.01	-0.04 p=.843
NWF-CLS EOY	7,391	51.7	29.3	7,208	46.4	25.5	**5.3**	72.4	<.001	.20	5.46 p<.001

Note: Bolded blue/shaded font indicates statistically significant differences between WCPSS and comparison mean scores.

Data Sources: 2013-14 Student Locator, 2013-14 WCPSS and comparison school district mCLASS® data.

---

[10] Cohen (1988) outlined criteria for gauging effect sizes: small *d* = .20, medium *d* = .50 and large *d* = .80.
[11] Lipsey et. al., (2012) reported that for standardized tests of narrow scope given at the elementary level the median effect size was .17 and the mean effect size was .25 for the 374 studies considered.

Figures 7 and 8 utilize boxplots to depict the mean, median, and range of the middle and end-of-year PSF and NWF-CLS scores for matched WCPSS and comparison school district students enrolled in 2013-14.

- The box represents the majority of student scores (25th to 75th percentile).
- The "whiskers," or vertical lines, extending from the box represent the range of scores, with the most extreme scores denoted by the dark line of small circles.
- Within each box, the mean is signified by a diamond and the median by a horizontal line in the middle of the box.

For WCPSS students the mean PSF MOY score was significantly higher (more than four points higher) than for matched comparison students. While both WCPSS and comparison school district students' MOY scores had a similar range in scores, WCPSS scores between the 25th and 75th percentile (the box) were higher than students in the comparison school district resulting in a higher median. The mean for the EOY PSF score was approximately the same (differences were not statistically significant).[12] Although the average scores were similar, WCPSS students had a smaller range of scores between the 25th and 75th percentile at the EOY (see Figure 7).

Figure 7
PSF by Benchmark Period for WCPSS and Comparison District
Kindergarten
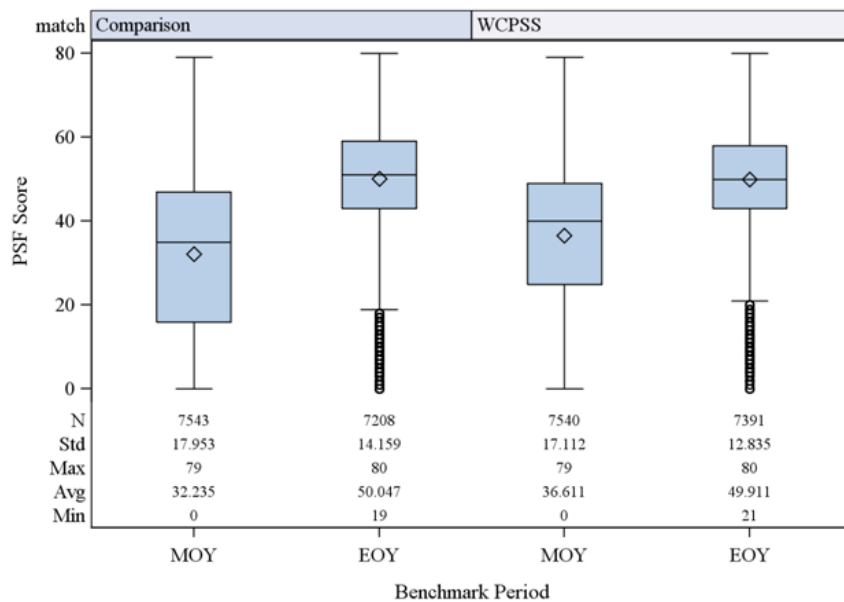
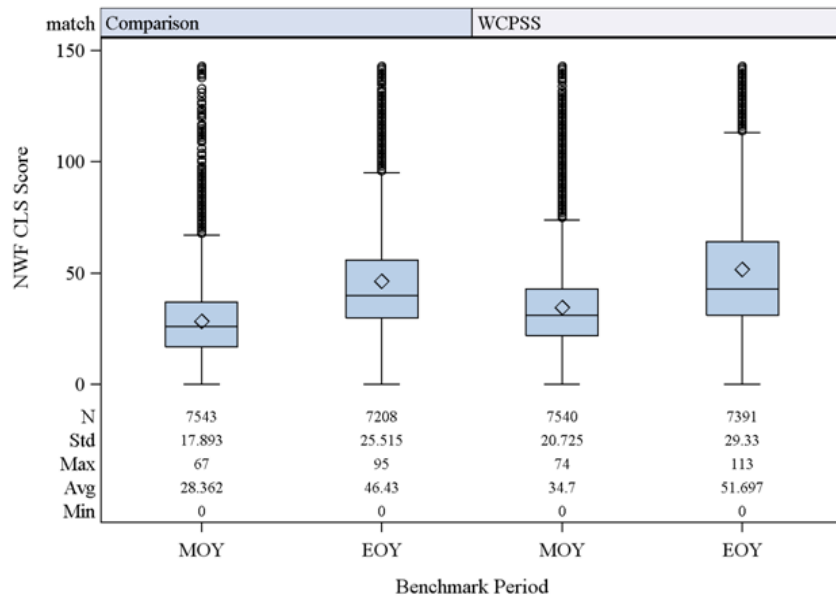Data Sources: 2013-14 Student Locator, 2013-14 WCPSS and comparison school district mCLASS® data.

The mean NWF-CLS MOY score was more than six points higher for WCPSS students than for matched comparison students and five points higher at the EOY (differences were statistically

---

[12] Statistical significance and effect size on MOY and EOY scores were based on ANOVA and regression analyses.

significant[13]). WCPSS NWF-CLS EOY scores between the 25th and 75th percentile (the box) were higher than students in the comparison school district resulting in higher median score (see Figure 8).

Figure 8
NWF-CLS by Benchmark Period for WCPSS and Comparison District
Kindergarten

Data Sources: 2013-14 Student Locator, 2013-14 WCPSS and comparison school district mCLASS® data.

The means, sample sizes, and standard deviations on the outcome variables overall and by subgroup are displayed in Table 8. The results of a two sample t-test are also included. These results indicate that on the PSF MOY, overall and students in all subgroups performed statistically significantly better in WCPSS than the comparison school district. On the other hand, White students in WCPSS performed significantly lower than the comparison school district on the PSF EOY in kindergarten.

On the NWF-CLS MOY, students overall and in all subgroups performed statistically significantly better in WCPSS than the comparison school district. Finally, on the NWF-CLS EOY, overall and students in all subgroups except for students who identify as Other performed statistically significantly better in WCPSS than the comparison school district.

Thus, results indicate, similar to overall results that among the majority of student subgroups for both PSF and NWF-CLS students' scores at the middle of the year were significantly higher for WCPSS students than students in the comparison school district. Students in the comparison school district were able to catch up to WCPSS students by the end-of-year on the PSF but remained statistically significantly lower on the NWF-CLS.

---

[13] Statistical significance and effect size on MOY and EOY scores were based on ANOVA and regression analyses.

Table 8

WCPSS and Comparison Means and Standard Deviations for Outcome Variables after Matching, Kindergarten

Outcome Variables		WCPSS After Matching			Comparison After Matching			T-Test on Difference in Means		
		n	Mean	SD	n	Mean	SD	Diff.	t	Sig.
PSF	**Overall**	**7,540**	**36.61**	**17.1**	**7,543**	**32.23**	**18.0**	**4.38**	**15.3**	**0.000***
(MOY)	LEP	1,120	23.84	16.8	1,092	19.88	16.0	3.96	5.7	0.000***
	Female	3,734	37.75	16.9	3,745	33.59	17.9	4.16	10.3	0.000***
	Male	3,806	35.50	17.3	3,798	30.90	17.8	4.60	11.4	0.000***
	Asian	520	39.54	16.2	518	34.07	17.9	5.47	5.16	0.000***
	Black	2,132	31.65	17.8	2,153	26.12	17.6	5.54	10.3	0.000***
	Hispanic/Latino	1,643	29.19	17.8	1,643	24.61	16.8	4.58	7.6	0.000***
	Other	33	34.33	17.0	70	30.24	17.9	7.74	1.9	0.058
	Multiracial	150	40.20	14.6	146	36.14	16.1	4.06	2.3	0.024*
	White	3,060	43.41	13.4	3,046	40.24	15.3	3.16	8.6	0.000***
PSF	**Overall**	**7,391**	**49.91**	**12.8**	**7,208**	**50.05**	**14.2**	**-0.14**	**-0.6**	**0.543**
(EOY)	LEP	1,114	43.20	14.9	1,044	42.39	16.8	0.81	1.2	0.233
	Female	3,659	51.28	12.4	3,586	51.63	13.4	-0.36	-1.2	0.240
	Male	3,732	48.57	13.1	3,622	48.47	14.7	0.09	0.3	0.771
	Asian	510	50.66	12.0	484	49.83	14.6	0.84	1.0	0.324
	Black	2,080	48.37	13.3	1,991	48.12	15.8	0.25	0.5	0.584
	Hispanic/Latino	1,607	45.86	14.1	1,592	45.59	15.6	0.26	0.5	0.619
	Other	32	50.00	14.6	63	50.00	17.1	0	0.0	1.000
	Multiracial	146	53.12	10.5	142	52.65	12.2	0.48	0.4	0.719
	White	3,014	52.85	11.1	2,968	53.64	10.8	-0.79	-2.8	0.005**
NWF-	**Overall**	**7,540**	**34.70**	**20.7**	**7,543**	**28.36**	**17.9**	**6.34**	**20.1**	**0.000***
CLS	LEP	1,120	24.83	16.3	1,092	20.19	15.4	4.63	6.9	0.000***
(MOY)	Female	3,734	34.75	20.1	3,745	29.05	17.2	5.69	13.2	0.000***
	Male	3,806	34.65	21.3	3,798	27.68	18.5	6.97	15.2	0.000***
	Asian	520	54.13	32.4	518	40.65	28.2	13.48	7.1	0.000***
	Black	2,132	29.81	16.5	2,153	24.78	14.8	5.03	10.5	0.000***
	Hispanic/Latino	1,643	26.93	15.6	1,643	21.89	14.4	5.04	9.63	0.000***
	Other	33	43.55	27.9	70	34.46	28.0	17.19	2.7	0.000***
	Multiracial	150	37.12	20.2	146	28.68	15.2	8.44	4.1	0.000***
	White	3,060	38.77	20.1	3,046	32.30	17.3	6.47	13.5	0.000***
NWF-	**Overall**	**7,391**	**51.70**	**29.3**	**7,208**	**46.43**	**25.5**	**5.27**	**11.6**	**0.000***
CLS	LEP	1,114	38.88	22.5	1,044	36.75	19.5	2.12	2.3	0.019*
(EOY)	Female	3,659	51.39	28.2	3,586	46.53	24.8	4.87	7.8	0.000***
	Male	3,732	52.00	30.38	3,622	46.33	26.3	5.66	8.5	0.000***
	Asian	510	77.70	38.3	484	64.85	35.3	12.84	5.5	0.000***
	Black	2,080	43.71	23.3	1,991	39.56	20.3	4.15	6.0	0.000***
	Hispanic/Latino	1,607	40.76	21.8	1,592	38.66	19.3	2.10	2.9	0.004**
	Other	32	59.50	34.2	63	53.5	33.3	12.11	1.5	0.151
	Multiracial	146	53.30	29.7	142	46.40	22.0	6.90	2.2	0.026*
	White	3,014	58.49	30.2	2,968	52.19	26.8	6.30	8.5	0.000***

Notes: 1. Students with disabilities were not included in analysis. Data only includes Kindergarten students.

2. Differences between means that are statistically significant at the .05 level are indicated with *, at the .01 level with **, and at the .001 level with ***.

3. Students may appear in more than one student subgroup.

Data Sources: WCPSS data: 2013-2014 Student Locator; comparison school district data: provided by comparison school district

Reading Achievement Trends Across Time

To assess whether WCPSS district goals for Letterland were met, results for DIBELS measures including PSF and NWF-CLS were examined for students in kindergarten and grade 1 across three years (2011-12 through 2013-14).[14] Letterland had not been implemented during 2011-12, which offers a baseline year. Letterland implementation began in kindergarten in 2012-13 and expanded to include kindergarten and grade 1 in 2013-14. The percent of students at each benchmark is depicted in Table 9 by grade and by assessment across years. Bar graphs depicting the percent of students at benchmark by student subgroup by assessment across years are available in Appendix E.

> **The percentage of kindergarten students at or above benchmark based on EOY PSF increased by 8.2 percentage points from 2011-12 to 2013-14.**

We first examined all students taking the PSF and NWF-CLS assessments across time, which meant different cohorts of students each year. As measured by PSF in kindergarten, the percentage of students who were at or above benchmark increased with each year of Letterland implementation (8.2 percentage points from 2011-12 to 2013-14) and the percentage of WCPSS students who were well below benchmark based on their PSF EOY score decreased between 2011-12 and 2013-14 from 11.2% to 6.2%. For NWF-CLS, trends were not consistent in direction across years in kindergarten. Across time as measured by the NWF-CLS in kindergarten, the percentage of students who were at or above benchmark increased (5.6 percentage points) after one year of implementation and decreased (7 percentage points) from year one to year two, which resulted in an overall 1.4 percentage points decrease from 2011-12 to 2013-14. The percentage of WCPSS kindergarten students who were well below benchmark based on their NWF-CLS EOY score decreased from 2011-12 to 2012-13, but increased in 2013-14.

The 2012-13 kindergarten students continued Letterland in grade 1 during the 2013-14 academic year (see cells shaded dark blue in Table 9). Since Letterland was implemented for kindergarten in 2012-13, the 2011-12 kindergarten students did not receive Letterland and continued to not receive Letterland in grade 1 during the 2012-13 academic year (see cells shaded light blue in Table 9). Due to student mobility, these groups are not exact cohorts. However, these groups are approximately the same cohorts of students across time. Looking at the two cohorts of students who were followed from kindergarten into grade 1, kindergarten students who received Letterland decreased more on measures of NWF-CLS and PSF at the beginning of grade 1 compared to students' scores at the beginning of grade 1 who did not receive Letterland. However this finding may be due to the change in assessments and benchmarks across years which impacted only the 2012-13 cohort.

---

[14] DIBELS Next benchmark cut scores were used to describe the scores that were collected using the DIBELS 6th edition measures.

Table 9
Percentage of K-1 Students Meeting Benchmark by Benchmark Period
2011-12 to 2013-14

DIBELS Assessments		2011-12 DIBELS Next Benchmarks Applied			2012-13 DIBELS Next Benchmarks Applied			2013-14 DIBELS Next		
Grade/Test	Benchmark	BOY %	MOY %	EOY %	BOY %	MOY %	EOY %	BOY %	MOY %	EOY %
K PSF	At or above	---	69.2	**72.6**	---	73.7	**78.6**	---	76.9	80.8
	Below	---	13.6	**16.2**	---	12.1	**12.9**	---	11.7	13.0
	Well below	---	17.2	**11.2**	---	14.3	**8.5**	---	11.4	6.2
	Total N		*10,642*	*10,642*		*11,645*	*11,645*		*11,683*	*11,570*
NWF-CLS	At or above	---	76.4	**81.2**	---	83.5	**86.8**	---	83.2	79.8
	Below	---	12.9	**14.1**	---	10.1	**10.4**	---	10.8	16.2
	Well below	---	10.7	**4.8**	---	6.4	**2.8**	---	6.0	4.0
	Total N		*10,642*	*10,642*		*11,645*	*11,645*		*11,682*	*11,570*
1 PSF	At or above	53.7	---	---	**62.4**	---	---	**52.1**	---	---
	Below	30.1	---	---	**24.6**	---	---	**34.6**	---	---
	Well below	16.2	---	---	**13.0**	---	---	**13.4**	---	---
	Total N	*10,719*			*11,811*			*12,110*		
NWF-CLS	At or above	72.0	74.1	70.4	**73.6**	74.1	71.2	**69.8**	73.7	72.1
	Below	14.6	13.9	13.8	**14.4**	13.3	13.1	**18.1**	13.1	11.9
	Well below	13.4	12.0	15.8	**12.0**	12.6	15.7	**12.2**	13.1	16.0
	Total N	*10,719*	*10,719*	*10,719*	*11,811*	*11,811*	*11,811*	*12,120*	*12,161*	*11,991*

Notes: 1. Light blue shading indicates baseline year prior to Letterland implementation and dark blue shading indicates first year of implementation.

    2. DIBELS Next began in 2012-13. Cut-offs for DIBELS Next were applied to 2011-12 and 2010-11 for comparability.

    3. Counts only include students at schools that participated in Letterland during 2012-13 and 2013-14 with scores for PSF or NWF-CLS in mCLASS® or AIMSweb data. Dash lines indicate benchmark periods that do not include DIBELS Next benchmark goals or cut points.

Data Sources: 2011-12 through 2013-14 Student Locator files, 2011-12 through 2013-14 mCLASS® and AIMSweb data. DIBELS data were obtained from AIMSweb and mCLASS® in 2011-12 and 2012-13 and from mCLASS® in 2013-14.

In order to further investigate student achievement prior to Letterland and during year one and year two of Letterland implementation, differences in achievement by sub-groups (e.g., ethnicity, LEP, SWD, and gender) across years were examined for PSF and NWF-CLS. Figure 9 through 12 illustrate the mean PSF and NWF-CLS scores at BOY or EOY by ethnicity, LEP, and SWD for kindergarten and grade 1 students. Appendix E includes bar graphs depicting percentages of students at or above benchmark across time for PSF and NWF-CLS at BOY and EOY by ethnicity, LEP, and SWD.

Figure 9 shows the mean PSF and NWF-CLS EOY score for kindergarten students. This figure suggests that following the implementation of Letterland, the average PSF EOY scores (left) for Asian, Hispanic/Latino, Multiracial, and White students increased across the first year of

implementation but decreased following the second year (2013-14). Black students' average score on the PSF gradually increased each year of implementation.

Across the ethnic subgroups depicted in Figure 9, kindergarten students' NWF-CLS EOY scores (right) increased after the first year of implementation and decreased during the second year. Because these data are cross-sectional and include different cohorts of students, graphs should be interpreted with care. The raw mean scores for students, which are illustrated in Figures 9 through 12, are displayed in Appendix E.

Figure 9
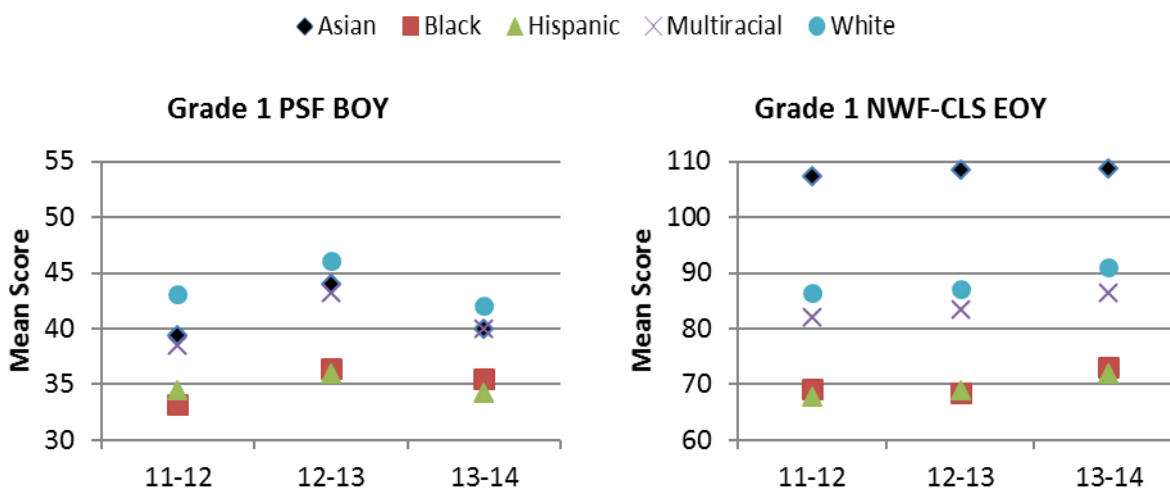*Kindergarten Students' Mean Scores on PSF and NWF-CLS (EOY Only)*
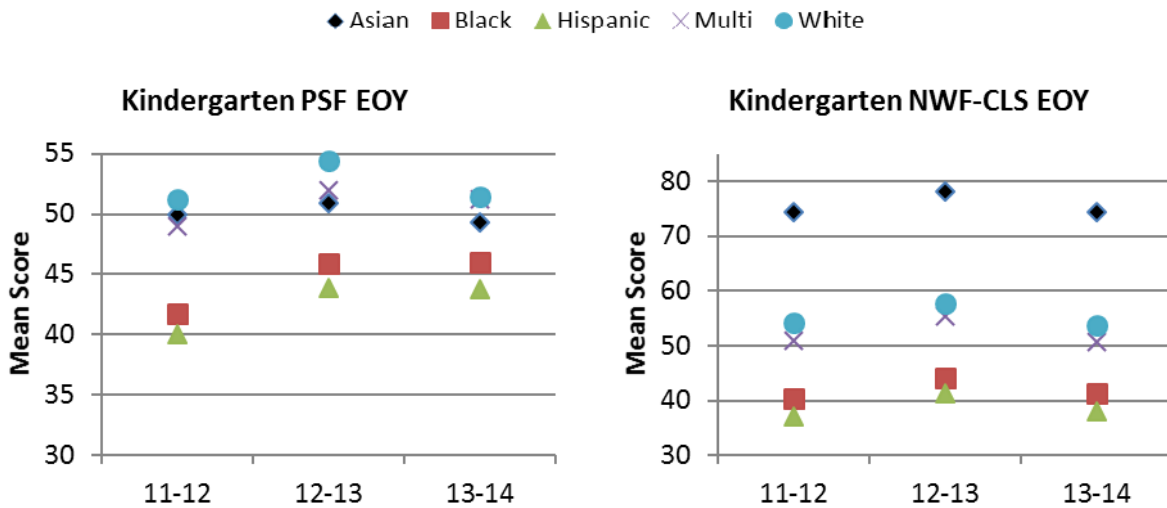
Note: 1. Counts only include students at schools that participated in Letterland during 2012-13 and 2013-14 with scores for PSF or NWF-CLS in mCLASS® or AIMSweb data.
2. Differences in y-axes reflect different score ranges on the PSF and NWF-CLS.
Data Sources: 2011-12 through 2013-14 Student Locator files, 2011-12 through 2013-14 mCLASS® and AIMSweb data. DIBELS data were obtained from AIMSweb and mCLASS® in 2011-12 and 2012-13 and from mCLASS® in 2013-14.

Figure 10 displays grade 1 students' average PSF BOY and NWF-CLS EOY scores for students. Letterland was not implemented in grade 1 until the 2013-14. Therefore the average PSF BOY scores by subgroup (left) are prior to Letterland implementation. The average NWF-CLS EOY scores by subgroup (right) illustrate a slight increase following the first year of Letterland implementation (2013-14).

Figure 10
Grade 1 Students' Mean Scores on PSF (BOY only) and NWF-CLS (EOY only)

Note: 1. Counts only include students at schools that participated in Letterland during 2012-13 and 2013-14.
2. Counts only include students with scores for PSF or NWF-CLS in mCLASS® or AIMSweb data.
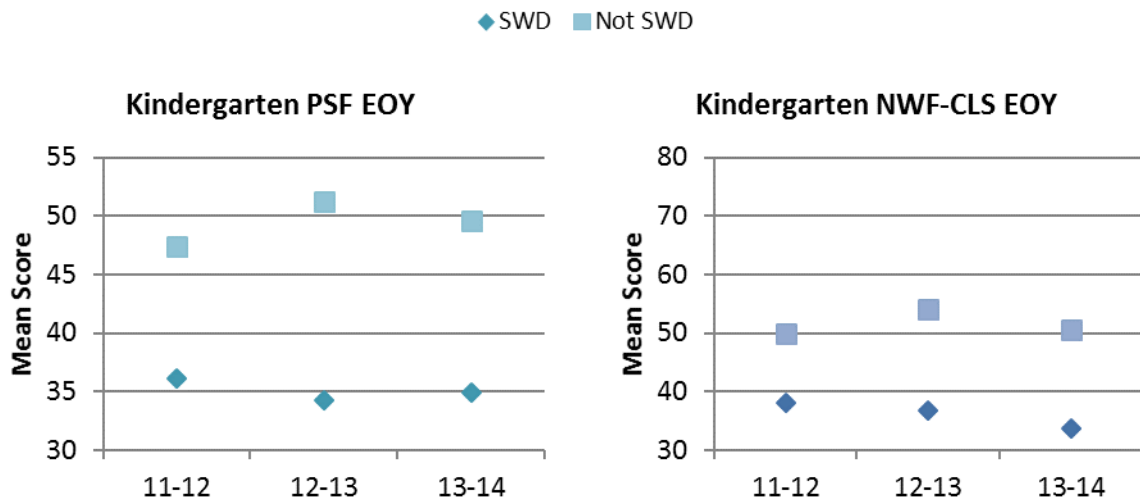3. Differences in y-axes reflect different score ranges on the PSF and NWF-CLS.

Data Sources: 2011-12 through 2013-14 Student Locator files, 2011-12 through 2013-14 mCLASS® and AIMSweb data. DIBELS data were obtained from AIMSweb and mCLASS® in 2011-12 and 2012-13 and from mCLASS® in 2013-14.

Kindergarten mean PSF and NWF-CLS scores disaggregated by LEP and SWD status were not consistent in directionality from 2011-12 through 2013-14. Figures 11 and 12 depict the mean scores for kindergarten students on PSF and NWF-CLS EOY for students based on SWD and LEP respectively. See Appendix E for raw mean scores and bar graphs depicting percentage of students at or above benchmark based on SWD and LEP by grade.

Figure 11 illustrates that for kindergarten students based on SWD, the average PSF EOY score (left) decreased following the first year of Letterland implementation (36.2 to 34.2) and increased slightly following the second year of Letterland implementation (34.9). The average NWF-CLS EOY score based on SWD in kindergarten (right) decreased following the first and second year of implementation.

Figure 12 shows that for kindergarten students identified as LEP, average PSF EOY score for students identified as LEP (left) increased following the first year of implementation (37.6 to 44.5) and then decreased following the second year of Letterland implementation (41.2). Similarly, average NWF-CLS EOY score (right) increased following the first year of Letterland implementation (36.1 to 48.1) and then decreased following the second year of Letterland implementation (36.9). This cross-sectional description of trends in scores across three years is limited in that the DIBELS Next benchmarks were used with scores collected with the DIBELS 6th Edition assessments. Additionally, the description is cross-sectional and describes scores across different groups of students**.**

Figure 11
Average Score on PSF and NWF-CLS for Kindergarten Students based on SWD
(EOY Only)

◆ SWD   ■ Not SWD

Note: 1. Counts only include students at schools that participated in Letterland during 12-13 and 13-14.
    2. Counts only include students with scores for PSF or NWF-CLS in mCLASS® or AIMSweb data.
    3. Differences in y-axes reflect different score ranges on the PSF and NWF-CLS.

Data Sources: 2011-12 through 2013-14 Student Locator files, 2011-12 through 2013-14 mCLASS® and
    AIMSweb data. DIBELS data were obtained from AIMSweb and mCLASS® in 2011-12 and 2012-13 and
    from mCLASS® in 2013-14.

Figure 12
Average Score on PSF and NWF-CLS for Kindergarten Students Based on LEP
(EOY Only)
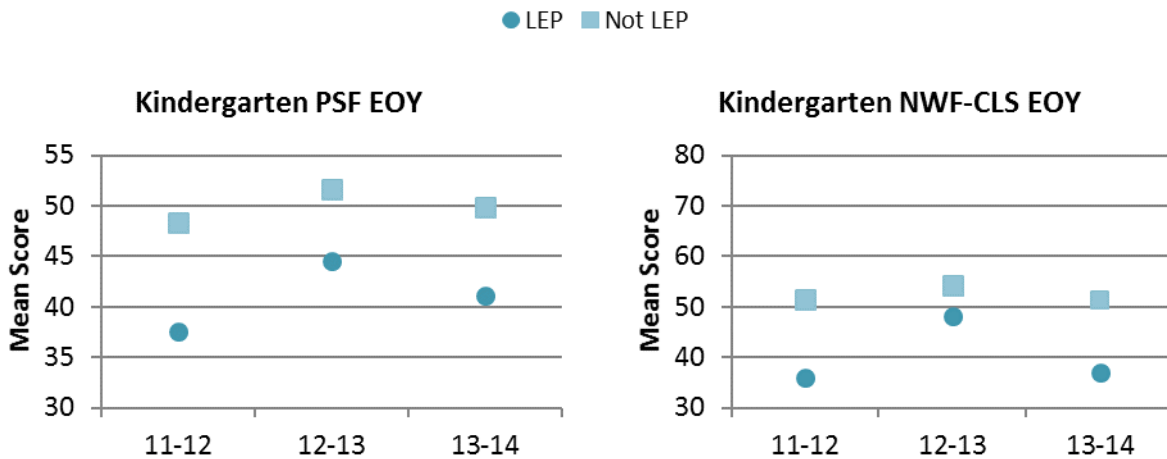
● LEP   ■ Not LEP

Note: 1. Counts only include students at schools that participated in Letterland during 12-13 and 13-14
    2. Counts only include students with scores for PSF or NWF-CLS in mCLASS® or AIMSweb data.
    3. Differences in y-axes reflect different score ranges on the PSF and NWF-CLS.

Data Sources: 2011-12 through 2013-14 Student Locator files, 2011-12 through 2013-14 mCLASS® and
    AIMSweb data. DIBELS data were obtained from AIMSweb and mCLASS® in 2011-12 and 2012-13 and
    from mCLASS® in 2013-14.

Conclusion/Discussion

The implementation of Letterland within WCPSS' K-1 classrooms was strong. The cross-departmental Letterland implementation team met monthly to discuss and plan for training as well as review data in an effort to strengthen implementation and provide support where needed (see Appendix F for Letterland Logic Model). Moreover, teachers were required to attend training prior to obtaining Letterland classroom materials, which is partially attributable to 100% of WCPSS K-1 teachers receiving training. In the spring of 2013-14, the second year of Letterland implementation within WCPSS' kindergarten classrooms, nearly 90% of the kindergarten teachers implemented Letterland with moderate to high fidelity. Implementation was equally strong within the grade 1 classrooms in their first year of implementation. Fidelity of implementation is a fundamental precondition for the success of programs/initiatives (Durlak, 2013). Thus, the quality of the implementation of Letterland in the classroom was promising. Indeed, all of the short-term goals related to Letterland's implementation that were set by the implementation team were met at the time of writing this report. The establishment and continuous work of the implementation team that oversaw the rollout of this district-wide effort appeared to contribute greatly to its successful implementation and therefore should be considered a model.

Although the intermediate goal of 100% of kindergarten teachers implementing with fidelity was not met, the fact that the vast majority of classrooms observed were implementing with moderate to high fidelity exceeds other program implementation efforts within WCPSS at a similar point in implementation such as the Daily CAFÉ (Rhea, 2014). The intermediate goals around higher student outcomes by spring of 2013 were met; although the adoption of DIBELS Next made it difficult to assess the success of the higher achievement goals set for 2014 (see Appendix F).

A relationship between fidelity of implementation and reading achievement was not demonstrated. However, the fidelity checklist, which was used by principals and principal designees to assess implementation, may not have been sensitive to the qualitative differences in Letterland implementation that influenced student outcomes. In other words, the checklist offered a simplified measure of program implementation but was likely not sensitive to the nuances in the quality of teacher implementation that mattered for student learning. The bluntness of the measure did not lend itself to understanding the extent to which various components of Letterland may be more or less critical to improving student outcomes. Therefore the critical core program components (Blase & Fixen, 2013) or the essential activities that are necessary to achieve outcomes are not clear following this study.

WCPSS' kindergarten students' PSF and NWF-CLS MOY scores were statistically significantly higher than matched students in the comparison school district. By the end of the year, WCPSS students' NWF-CLS scores remained statistically significantly higher than matched students in the comparison school district while PSF scores were similar for students in both districts. Although PSF EOY results were similar for WCPSS kindergarten students and matched students, this study's findings were that significantly fewer WCPSS students' scores were well below benchmark on the DIBELS Next indicator than matched students' scores in the comparison school district; and the percentage of WCPSS students who were well below benchmark based

on their PSF EOY score decreased from before implementation (2011-12 to 2013-14) and were also consistent with prior research findings (Felton & Crawford, 2010).[15] Although year one increases in kindergarten students' NWF-CLS EOY scores were reversed in year two (2013-14), 2013-14 NWF-CLS EOY scores were statistically significantly higher for WCPSS students compared to their counterparts in the comparison school district.

It should be noted that although the effect sizes found would be considered small using the interpretations suggested by Cohen (1988), when compared to effect sizes found in other educational interventions they could be considered moderate to high. This study represents an independent evaluation of Letterland using a quasi-experimental design thereby adding valuable evidence of neutral to positive student outcomes following Letterland participation. While the quasi-experimental design enabled us to conclude that participation in Letterland was correlated with the changes found, it precluded us from ruling out the possibility that other uncontrolled factors influenced the results. Furthermore, the US National Reading Panel (2000) study found that stronger research designs resulted in larger effect sizes. Indeed, they suggested maximizing the rigor of a study's research design by limiting as many external and internal threats to validity as possible. They found studies with an experimental design not only provided more conclusive findings which allow for greater confidence in results, but also increased the chance of finding treatment effects where they exist. In light of this study's findings, there are questions that remain to be answered.

While the change to DIBELS Next prohibited a longitudinal cohort study, within the current study DIBELS Next benchmarks were applied across years. Among kindergarten students, the percentage of students who were at or above benchmark increased and the percentage of students below benchmark decreased on the PSF with each year of Letterland implementation. Results for the percentage of kindergarten students on NWF-CLS were not consistent showing positive results in year one and then reversing in year two of implementation.

The change in assessments and benchmarks, following the district adoption of DIBELS Next during the 2013-14 academic year may have complicated the examination of reading achievement across years for several reasons. For example, despite the fact that DIBELS Next benchmark cut scores were applied across years, scores are not independent of benchmarks. Raters' knowledge of the current benchmark cut scores sometimes influences score assignment (e.g., a rater or teacher may encourage students to aim for a score and set goals based on the current benchmark cut scores). When those benchmarks change, explicit goals may change. Another reason the change in benchmarks complicates analysis is that students who were close to being on benchmark may have been allowed to retake the assessments in an effort to achieve an at or above benchmark score. Benchmark cut scores for the DIBELS Next were higher than the DIBELS 6th edition. Although the changes in the measures themselves were minor, these changes also complicated comparisons in score inferences across years.

---

[15] Well below benchmark replaced 'at risk' as the lowest benchmark category in the DIBELS Next edition.

Questions to Consider

Letterland is still in its early years of implementation in WCPSS. Questions to consider as WCPSS moves forward include:

1. Are teachers stronger at implementation of some components of Letterland than others? If so, what could improve implementation?
2. Are there core components within Letterland that K-2 teachers should emphasize in order to strengthen student outcomes, particularly for specific subgroups of students (LEP and/or SWD)?
3. For students struggling with early literacy skills, what supports are classroom teachers providing to students?
4. What is the level of fidelity implementation for intervention teachers supporting students using Letterland within and outside of the classroom?
5. How is the overall literacy program being implemented K-3? How much time is spent on various components during literacy and to what extent are those components implemented with fidelity? To what extent to variations in implementation affect student learning?
6. How are teachers holding students accountable for transferring the foundational skills taught in Letterland throughout the day?
7. Does a systematic program teaching foundational literacy skills via Letterland impact reading proficiency at the end of grade 3?

Recommendations

Overall, Letterland implementation and initial student outcome data suggest continuation of this approach with consideration of ways to make the model even stronger and shore up implementation in weak to moderate implementation classrooms. The following recommendations are made with this goal in mind.

Continue monitoring classroom implementation. Offer support for those with lower levels of implementation. (Could be done by school if isolated teachers.) The fact that the vast majority of teachers (89%) observed implemented Letterland with moderate to high fidelity in kindergarten classrooms and reading achievement was not significantly related to fidelity may either indicate that moderate implementation was not sufficient to impact student achievement or that the simplified fidelity checklist used to measure implementation did not adequately distinguish subtle differences in the fidelity of implementation of Letterland. Furthermore, while across all observed kindergarten classrooms student reading achievement was not related to the level of fidelity within the classroom, there was some evidence to suggest that high fidelity may be related to higher student EOY scores within classrooms (analysis focused on classrooms with either very high or very low fidelity scores), thereby indicating that there may be an impact of implementation that is not evident in classrooms that implemented with moderate fidelity.

Moreover, the fidelity instrument used was not sensitive enough to isolate whether there were essential Letterland components (Blase & Fixen, 2013) that were necessary to improve student

achievement. Thus, classroom observations should be conducted with a more detailed implementation instrument in order to identify these critical components. This information should then be shared with teachers in order to further strengthen Letterland's implementation.

The implementation team should review data to select a sample of kindergarten, grade 1, and possibly grade 2 classrooms to observe. The 2014 fidelity data could be used to select schools and classrooms with the highest and lowest implementation for further observation.

Sustain the program. Given student outcomes were neutral to positive, offering promising albeit inconclusive evidence that Letterland had a positive impact on student reading achievement, and the annual cost of continuing the program is approximately 10% of the initial annual investment, WCPSS should continue the program. Moreover, the calculated cost effectiveness of Letterland demonstrated that the cost per student for continuing the program is substantially lower than start-up costs. Based on continuing costs, the cost effectiveness analysis indicted that: (a) the program cost is $2 per kindergarten student for each percentage point increase in the overall percentage of students at or above benchmark on the PSF EOY (for NWF-CLS there was a decrease in the percentage of students at or above benchmark); and (b) the program cost is $15 per Grade 1 student for each percentage point increase in the overall percentage of students at or above benchmark on NWF-CLS EOY.

Examine DIBELS results in grade 2 and reading comprehension in grade 3. An examination of grade 2 students' DIBELS Next patterns in 2014-15 and student reading comprehension outcomes in 2015-16 is advised. 2014-15 results will allow a two year comparison of DIBELS Next scores. In 2015-16, the 2012-13 kindergarten cohort of WCPSS students, the first cohort of Letterland kindergarten students, will be in grade 3 and their reading beginning-of-grade and end-of-grade exams should be examined to determine if reading comprehension has been impacted by participation in Letterland.

The program implementation offers a highly scalable and sustainable model for those interested in implementing similar programs districtwide. The model included grade level roll out and offered continuous support and training for all K-1 teachers, which resulted in high levels of program implementation within classrooms. Impacts of the program on student learning were neutral to slightly positive; thus, continuation of the program over time may contribute to larger impacts on the student outcomes.

References

Blase, K. & Fixen, D. (2013). Core intervention components: Identifying and operationalizing what makes programs work. *Office of the Assistant Secretary for Planning and Evaluation Research Brief.* Retrieved from http://aspe.hhs.gov/hsp/13/KeyIssuesforChildrenYouth/ CoreIntervention/rb_CoreIntervention.cfm

Cartwright, K.B. (2012). Insights from cognitive neuroscience: The importance of executive function for early reading development and education. *Early Education and Development*, 23(1), pp. 24-36, DOI: 10.1080/10409289.2011.615025

Cohen, J. (1988), *Statistical Power Analysis for the Behavioral Sciences, 2nd Edition.* Hillsdale: Lawrence Erlbaum.

Dehejia, R. H. & Wahba, S. (2002). Propensity score-matching methods for nonexperimental causal studies. *The Review of Economics and Statistics*, *84*(1): 151–161.

Denis, D. J. (2012). *Understanding Cohen's d.* Retrieved from http://www.statpt.com/applied_gen/cohen_d.pdf

Durlak, J. (2013). The importance of quality implementation for research, practice, and policy. *Office of the Assistant Secretary for Planning and Evaluation Research Brief.* Retrieved from http://aspe.hhs.gov/hsp/13/KeyIssuesforChildrenYouth/ImportanceofQuality/rb_QualityImp.cfm

Ehri, L. C., & McCormick, S. (1998). Phases of word learning: Implications for instruction with delayed and disabled readers. *Reading and Writing Quarterly,* 14, 135-163.

Ehri, L. C., Nunes, S. R., Willows, D. M., Schuster, B. V., Yaghoub-Zadeh, Z., & Shanahan, T. (2001). Phonemic awareness instruction helps children learn to read: Evidence from the National Reading Panel's meta-analysis. *Reading Research Quarterly*, *36*(3), pp. 250-287.

Felton R., & Crawford, L. (2010). *Letterland in a Florida Public School Kindergarten.* Retrieved from http://www.letterland.com/documents/pdf/Letterland%20Boosts%20Dibels%20Targets%20(July%202010).pdf

Gillon, G. T. (2004). *Phonological awareness: From research to practice.* New York: The Guilford Press.

Letterland International (2014). *Letterland: Child-friendly phonics.* Retrieved from www.letterland.org

Lenard, M. (2013). *Student performance and achievement gaps on measures of early literacy: 2011-12.* Raleigh, NC: Wake County Public School System. Data & Accountability department.

Levin, H. M. & McEwan P. J. (2001). *Cost-effectiveness analysis, 2nd Edition: Methods and application.* Thousand Oaks, CA: Sage Publications, Inc.

Lipsey, M. W., Puzio, K., Yun, C., Hebert, M. A., Steinka-Fry, K., Cole, M. W., …Busick, M. (2012). *Translating the statistical representation of the effects of education interventions into more readily interpretable form* (NCSER 2013-3000). National Center for Special Education Research. Retrieved from http://ies.ed.gov/ncser/pubs/20133000/pdf/20133000.pdf

National Institute of Child Health and Human Development. (2000). Report of the national reading panel: Teaching children to read. Retrieved from http://www.nichd.nih.gov/publications/pubs/nrp/Pages/smallbook.aspx

North Carolina Department of Public Instruction. (2013, October). *North Carolina read to achieve: A guide to implementing House Bill 950/S.L. 2012-142 Section 7A*. Retrieved from http://www.dpi.state.nc.us/docs/k-3literacy/resources/guidebook.pdf

Rhea, A. (2014). *Daily CAFÉ survey, 2014.* Raleigh, NC: Wake County Public School System. Data & Accountability department. Unpublished manuscript.

Salkind, N. J. (2011). *Statistics for people who hate statistics* (4 ed.). Los Angeles, CA: Sage.

Stahl, S. A. & Murray, B. A. (1998). Issues involved in defining phonological awareness and its relation to early reading. In J. Metsala and L. C. Ehri (Eds.), *Word recognition in beginning literacy* (pp. 65-87). Mahwah, NJ: Erlbaum.

Tabachnick, B. G., & Fidell, L. S. (2007). *Using multivariate statistics*. Boston, MA: Pearson Education, Inc.

The US National Reading Panel. (2000). The US national reading panel report, chapter 2, part II: Phonics instructions. pp. 2-125. Retrieved from http://www.nichd.nih.gov/publications/pubs/nrp/Documents/report.pdf

Wendon, L. (2010, July 9). Letterland increases brain activity [Web blog post]. Retrieved from http://www.letterland.com/news/2010-07-09-letterland-increases-brain-activity

Letterland Lesson Checklist

Refer to Letterland Illustrated Lesson Checklist

Grade Level

○ Kindergarten

○ Grade 1

School Name

[   ]

Teacher Name

[   ]

Letterland Story Logic

1 or more of the activities below is in each lesson

☐ Teacher talks

☐ Letterland software

☐ ABC; Beyond ABC; or Far Beyond ABC books

☐ Letterland songs

☐ Children tell the story logic

☐ Role play

☐ Students' picture code

Date

[   ]

Observer

[   ]

Lesson

[   ]

Letterland Sound Review

1 of the activities below is included in each lesson

[   ]

Blending activities

The two (2) essential items below are a part of all blending activities

- ☐ Essential steps
- ☐ Essential multi-sensory step
- ☐ Live reading
- ☐ Pocket chart or software
- ☐ Teacher writes the word
- ☐ Letter sets for each child

Segmenting activities

The 2 essential items below are a part of all segmenting activities

- ☐ Essential steps
- ☐ Essential multi-sensory step
- ☐ Live Spelling
- ☐ Pocket chart or software
- ☐ Individual letter sets
- ☐ Children write the words

Other word level activities

- ☐ Reading word cards on the pocket chart
- ☐ Sorting word cards on pocket chart
- ☐ Written word sort
- ☐ Practicing spelling and writing Tricky Words
- ☐ Writing words to dictation or writing words independently
- ☐ Playing games with individual words cards
- ☐ Reading word lists

Sentence/text level

One (1) activity below is included in each lesson

- ☐ Word Detectives
- ☐ Reading Review Sentences
- ☐ Reading decodable stories; booklets/plays
- ☐ Discussing stories
- ☐ Role playing stories

Environment

All 4 items below are displayed in Letterland classrooms

☐ Aa-Zz frieze or Alphabet poster

☐ Other Letterland posters (Vowels; Magic e; Syllables Types; etc.)

☐ Students' picture-coded letters or words

☐ Reading direction sign

Comments

Appendix B

Table B1

2013-14 Letterland Average Fidelity of Implementation Scores by School

School Name	Kindergarten		Grade 1	
	Average Fidelity Score	Number of Teachers	Average Fidelity Score	Number of Teachers
Adams Elementary	8.3	3	7.0	1
Apex Elementary	7.3	2	0.0	0
Aversboro Elementary	6.8	4	7.0	5
Ballentine Elementary	4.9	6	5.8	8
Banks Road Elementary	5.1	5	6.1	6
Barwell Road Elementary	6.9	9	8.0	6
Brassfield Elementary	0.0	0	10	2
Brentwood Elementary	5.5	7	6.3	6
Briarcliff Elementary	7.8	2	6.1	6
Brier Creek Elementary	4.8	7	6.4	7
Brooks Elementary	5.5	6	5.2	4
Bugg Elementary	4.1	3	4.1	5
Carpenter Elementary	5.4	6	6.9	5
Carver Elementary	3.5	3	7.8	2
Cary Elementary	9.4	4	8.4	6
Cedar Fork Elementary	5.2	5	4.4	5
Combs Elementary	7.1	5	7.1	4
Conn Elementary	6.1	4	5.0	4
Davis Drive Elementary	7.0	4	6.4	4
Dillard Drive Elementary	8.9	6	6.4	6
Douglas Elementary	3.4	8	6.6	5
Durant Elementary	5.9	8	4.6	7
East Garner Elementary	5.0	1	6.5	2
Farmington Woods Elementary	7.0	6	7.7	3
Forestville Road Elementary	8.3	9	7.4	11
Fox Road Elementary	5.0	5	5.7	7
Fuller Elementary	4.8	4	4.8	3
Fuquay-Varina Elementary	4.3	5	7.3	4
Green Elementary	6.4	6	5.8	4
Green Hope Elementary	8.6	6	7.3	5
Harris Creek Elementary	4.4	6	4.2	3
Herbert Akins Elementary	6.8	8	6.8	8
Heritage Elementary	4.5	2	5.6	5
Highcroft Drive Elementary	5.8	8	5.7	7
Hilburn Drive Elementary	6.5	2	4.5	3
Hodge Road Elementary	4.2	3	3.7	3
Holly Grove Elementary	6.4	7	4.9	12
Holly Ridge Elementary	5.8	6	7.4	7
Holly Springs Elementary	6.8	10	7.9	5
Hunter Elementary	9.0	4	8.0	6
Jeffreys Grove Elementary	10	2	0.0	0
Jones Dairy Elementary	6.6	5	7.5	6
Kingswood Elementary	6.8	3	8.3	3
Lacy Elementary	8.0	6	8.3	6

School Name	Kindergarten		Grade 1	
	Average Fidelity Score	Number of Teachers	Average Fidelity Score	Number of Teachers
Lake Myra Elementary	8.1	4	7.9	5
Lead Mine Elementary	4.3	6	4.8	4
Leesville Elementary	5.7	5	4.3	5
Lincoln Heights Elementary	6.0	2	0.0	0
Lockhart Elementary	5.0	6	5.3	6
Lynn Road Elementary	9.5	1	8.5	3
Millbrook Elementary	7.0	6	9.8	6
Mills Park Elementary	9.0	2	0.0	0
Morrisville Elementary	4.0	6	4.5	7
North Forest Pines Elementary	5.4	8	4.1	7
Northwoods Elementary	5.6	4	6.7	5
Oak Grove Elementary	8.5	4	0.0	0
Olds Elementary	2.5	3	2.0	2
Olive Chapel Elementary	6.6	5	6.7	6
Penny Road Elementary	6.7	5	6.4	5
Pleasant Union Elementary	5.0	2	5.0	2
Poe Elementary	5.9	4	4.3	4
Rand Road Elementary	8.3	2	0.0	0
Reedy Creek Elementary	5.1	6	5.8	6
Richland Creek Elementary	4.0	2	3.0	1
River Bend Elementary	4.2	6	5.6	4
Rolesville Elementary	6.9	4	7.7	3
Salem Elementary	3.4	5	0.0	0
Sanford Creek Elementary	7.9	5	8.5	5
Smith Elementary	3.5	2	0.0	0
Stough Elementary	9.3	3	0.0	0
Swift Creek Elementary	7.5	4	0.0	0
Sycamore Creek Elementary	5.1	9	4.9	9
Timber Drive Elementary	2.9	4	5.0	3
Underwood Elementary	6.8	3	5.2	4
Vance Elementary	6.8	4	7.3	3
Vandora Springs Elementary	7.1	2	0.0	0
Wake Forest Elementary	5.2	4	4.8	5
Wakefield Elementary	8.7	3	10.0	1
Wakelon Elementary	2.9	5	6.0	5
Walnut Creek Elementary	8.5	1	9.0	2
Washington Elementary	7.2	3	5.5	4
Weatherstone Elementary	3.1	6	3.8	7
Wendell Elementary	4.2	6	4.1	4
West Lake Elementary	5.0	6	6.9	6
Wilburn Elementary	9.5	5	4.0	5
Wildwood Forest Elementary	5.4	4	7.3	4
Willow Springs Elementary	8.6	4	9.0	1
Yates Mill Elementary	5.6	6	6.8	5
Zebulon Elementary	7.0	4	7.0	2
WCPSS	6.0	412	6.2	373

Data Source: Spring 2014 Letterland fidelity data
Note: Only schools that provided Letterland fidelity data are listed.

Appendix C

Kindergarten students' FSF scores at the beginning of the year were not correlated to the level of fidelity within the classroom (see first plot in Figure C1). Furthermore, reading achievement outcomes (PSF and NWF-CLS scores) at the end of the year were also not correlated to their teacher's fidelity score (see plots 2 and 3 in Figure C1).

Figure C1
Average Students' Score on Key Kindergarten Literacy Indicators
by Average Fidelity Score

Data Sources: WCPSS data: 2013-14 Student Locator; spring 2014 Letterland fidelity data; and 2013-14 mCLASS® data.
Notes: 1. The correlation coefficient for average fidelity score and FSF BOY r = .01, PSF EOY r = .04, and NWF-CLS r = .02.
          2. Differences in y-axes reflect different score ranges on the PSF and NWF-CLS.
Interpretation Example: The lack of clustering in the scattered plot on the far right shows that NWF-CLS scores at the end of the year were not correlated to teacher's fidelity score.

Appendix D

Propensity Score Matching Technique

The propensity score-match was run utilizing a greedy match procedure that matches students' propensity scores initially using five decimal places—approximately two-thirds of students matched on the first iteration (Dehejia & Wahba, 2002). In other words, WCPSS students were matched to students in the comparison school district with the exact same propensity score to the fifth decimal point. Next, WCPSS students who did not have exact matches to the fifth decimal point were matched with comparison students with the exact same propensity score to the fourth decimal point, which did not result in any matches (see Table D1). The procedure continues to attempt matches using the exact same propensity score to the third, second, and finally first decimal point. Given the differences in the mean FSF (pre-assessment score) for these matched groups and the small percentage of students that would have been added to the study based on subsequent matches, only students matching based on five decimal places were used in this study's matched group analysis.

Table D1
2013-14 Number and Percentage of Students Matched Based on Greedy-Match Procedure within WCPSS and Comparison School Districts

Propensity Match	District	Number of Students Matched	% of Study Population	Mean on FSF (pre-assessment)
Five decimal places	Comparison	7,739	64%	13.79
	WCPSS	7,739	68%	13.79
Four decimal places	Comparison	0	N/A	N/A
	WCPSS	0	N/A	N/A
Three decimal places	Comparison	123	1%	15.82
	WCPSS	123	1%	21.59
Two decimal places	Comparison	206	2%	10.58
	WCPSS	206	2%	10.46
One decimal places	Comparison	112	1%	13.39
	WCPSS	112	1%	12.41

Data Sources: 2013-14 Student Locator, 2013-14 WCPSS and comparison school district mCLASS® data.

Propensity scores were predicted using a logistic regression model, which included the nine covariates (see Table D2). The WCPSS group had an average propensity score of .438 a standard deviation of .155, a minimum score of .063, and a maximum score of .825. The comparison group had an average propensity score of .536, a standard deviation of .141, a minimum score of .122, and a maximum score of .825. Figure D1 illustrates the overlap in propensity scores for the WCPSS and the comparison groups. Figure D1 illustrates that the propensity scores for students in the comparison school district were similar to the propensity scores for the students in WCPSS. This graph supports that comparisons between students in the comparison school district and WCPSS can be made because students were similar across the two school districts as indicated by the adequate overlap in distribution of propensity scores in each group. In order to

further establish that the students in WCPSS and the comparison school district were comparable, balance was investigated by examining standardized bias.

Figure D1
Overlap in Propensity Scores between WCPSS and Comparison school district
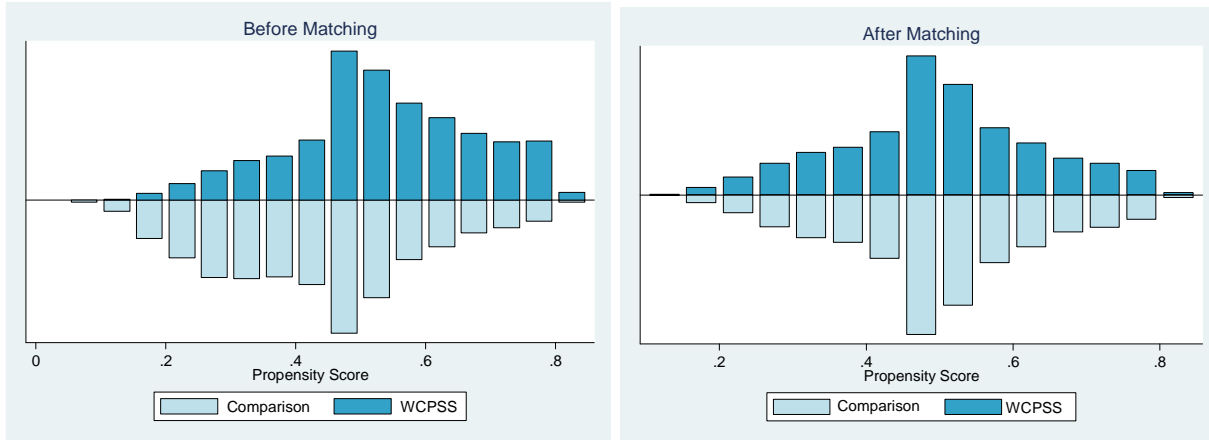
Note: Data do not include SWD status. Before matching, populations include 11,366 WCPSS students and 12,059 comparison district students. After matching, samples include 7,739 WCPSS students and 7,739 comparison district students.
Data Sources: WCPSS data: 2013-14 Student Locator file; comparison school district data provided by that district.

Table D2 displays the percentage of standardized bias for covariates prior to and after matching. The percentage of standardized bias between differences on the covariates was decreased to below one for all covariates after matching. Overall, the two groups were almost perfectly balanced on the eight covariates. The standardized bias estimates suggest that the bias between the two groups on each of the covariates were reduced to the extent that the differences were no longer significant.

Table D2
WCPSS and Comparison Means on Variables Prior to and After Matching

Variable	WCPSS Prior to Matching (n=11,379)	Comparison Prior to Matching (n=12,059)	% Bias Prior to Matching	WCPSS After Matching (n=7,739)	Comparison After Matching (n=7,739)	% Bias After Matching
	%	%		%	%	
LEP status	13.8	16.0	-6.2***	14.5	14.5	0.0
Female	49.8	50.0	-0.4	49.5	49.5	0.0
Asian	7.4	6.7	3.0**	6.9	6.9	0.0
Black	22.3	38.4	-35.6***	28.7	28.7	0.0
Hispanic/Latino	19.1	22.8	-9.2***	21.8	21.8	0.1
Multiracial	4.0	1.7	35.3***	2.0	2.0	0.0
White	46.7	29.8	35.3***	40.1	40.1	0.0
Other	0.4	0.6	-2.6	0.4	0.5	-0.6
	Mean Score	Mean Score		Mean Score	Mean Score	
FSF (BOY)	12.4	16.5	-35.2***	13.8	13.8	0.0

Note: Students with disabilities were not included in analysis. Analysis only includes kindergarten students.
    Differences between groups that are statistically significant at a .05 level are indicated with *, at a .01 level
    with **, and a .001 level with ***.
Data Sources: WCPSS data: 2013-14 Student Locator file; comparison school district data provided by that district

Linear regression analyses were conducted to estimate the average treatment effect of Letterland on student outcomes including PSF MOY, PSF EOY, NWF-CLS MOY, and NWF-CLS EOY. Matching covariates were included in the models as control variables, making the models doubly-robust. Correlations across covariates were low, and the variance inflation factor was below five in each model, indicating that the models did not include unacceptable multicollinearity.[16] The following model was estimated for PSF and NWF-CLS:

$$PSF \text{ or } NWF\ CLS = B_0 + B_1 letterland + B_2 FSF + B_3 gender + B_4 asian + B_5 black + B_6 hispanic + B_7 multi + B_8 other + B_9 LEP + B_{10} letterland * LEP + B_{11} letterland * asian + B_{12} letterland * black + B_{13} letterland * hispanic + B_{14} letterland * multi + B_{15} letterland * other + B_{16} letterland * gender + u$$

Where *PSF* or *NWF CLS* are students' scores at the middle or end of the year (models were fit for both middle and end of the year for each measure), *letterland* is a dichotomous variable indicating whether the student received Letterland, *FSF* is the student's score on the FSF assessment at the beginning of kindergarten, *gender* is a dichotomous variable indicating that the student is female, *Asian, Black, Hispanic/Latino, Multi,* and *Other* are dichotomous variables indicating ethnicity, and *LEP* is a dichotomous variable indicating that the student is identified as limited English proficient. Interaction terms were included for students who received Letterland and their ethnicity, LEP status, or gender. However, in order to estimate the effect of Letterland participation on outcomes controlling for prior achievement (FSF) and student characteristics, four models were first fit without interaction terms (see Table D3). Second, interaction terms were included to estimate the effect of Letterland on outcomes for subgroups (see Table D4).

---

[16] Multicollinearity occurs when two or more predictor variables are highly correlated (Tabachnick & Fidell, 2007).

four models were first fit without interaction terms (see Table D3). Second, interaction terms were included to estimate the effect of Letterland on outcomes for subgroups (see Table D4).

Table D3
Effects of Letterland on PSF and NWF-CLS

Variables	MOY PSF		EOY PSF		MOY NWF-CLS		EOY NWF-CLS	
	Coef.	*S.E.*	*Coef.*	*S.E.*	*Coef.*	*S.E.*	*Coef.*	*S.E.*
Letterland	4.43***	0.23	-0.04	0.21	6.39***	0.27	5.46***	0.40
FSF	0.65***	0.01	0.23***	0.01	0.69***	0.01	0.91***	0.02
Gender (Female)	1.23***	0.24	2.40***	0.21	-0.55*	0.27	-1.87***	0.40
Asian	-2.80***	0.46	-1.59***	0.43	13.27***	0.89	17.67***	1.11
Black	-7.39***	0.32	-3.01***	0.27	-2.33***	0.33	-5.91***	0.50
Hispanic/Latino	-3.88***	0.40	-2.14***	0.34	-1.53***	0.44	-3.44***	0.65
Multi	-2.12**	0.82	0.11	0.68	-0.94	0.98	-3.36**	1.48
Other	-5.79**	1.76	-0.85	2.00	4.50	2.90	5.04	3.61
LEP	-7.72***	0.45	-5.26***	0.43	-4.31***	0.48	-4.78***	0.69
Constant	26.92***	0.34	47.73***	0.29	19.74**	0.36	36.47***	0.53
Number of observations	14,083		14,599		15,083		14,599	
R-squared	.33		.11		.27		.24	

Note: Statistical significance at the p<.05 level is indicated with *, at the p<.01 level with **, and at the p<.001 level with ***.

Data Sources: WCPSS data: 2013-14 Student Locator file; 2013-14 mCLASS® data

Table D4
Effects of Letterland on PSF and NWF-CLS with Interactions by Subgroup

Variables	MOY PSF		EOY PSF		MOY NWF-CLS		EOY NWF-CLS	
	Coef.	*S.E.*	*Coef.*	*S.E.*	*Coef.*	*S.E.*	*Coef.*	*S.E.*
Letterland	3.45***	0.41	-0.54	0.35	7.20***	0.53	6.90***	0.81
FSF	0.65***	0.01	0.23***	0.01	0.69***	0.01	0.91***	0.02
Gender (Female)	1.44***	0.33	2.61***	0.31	0.09	0.36	-1.44**	0.53
Asian	-4.21***	0.65	-2.44***	0.64	9.34***	1.14	13.67***	1.50
Black	-8.57***	0.44	-3.55***	0.40	-1.61***	0.42	-4.84***	0.65
Hispanic/Latino	-5.12***	0.55	-2.49***	0.49	-1.72**	0.55	-2.33**	0.85
Multi	-2.66**	1.21	-0.63	1.04	-2.02	1.23	-3.97**	1.85
Other	-6.71**	2.50	-0.33	3.14	0.37	3.81	3.95	4.83
LEP	-6.85***	0.62	-5.64***	0.64	-2.67**	0.63	-2.92**	0.90
Letterland and LEP	-1.77**	0.88	0.71	0.85	-3.24**	0.96	-3.62**	1.34
Letterland and Asian	2.81**	0.93	1.66*	0.85	7.84***	1.77	7.80***	2.21
Letterland and Black	2.36***	0.59	1.06**	0.53	-1.44*	0.63	-2.10**	0.95
Letterland and Hispanic/Latino	2.50**	0.78	0.73	0.67	0.38	0.86	-2.17*	1.28
Letterland and Multi	1.07	1.65	1.47	1.37	2.10	1.96	1.19	2.95
Letterland and Other	1.65	3.50	-0.96	4.01	8.31	5.76	1.82	7.21
Letterland and Gender (Female)	-0.42	0.47	-0.41	0.42	-1.25**	0.55	-0.83	0.80
Constant	27.41***	0.39	47.98***	0.33	19.33***	0.41	35.73***	0.61
Number of observations	15,083		14,599		15,083		14,599	
R-squared	.34		.11		.28		.24	

Note: Statistical significance at the p<.05 level is indicated with *, at the p<.01 level with **, and at the p<.001 level with ***. Interactions between Letterland and subgroups are shaded light blue.

Data Sources: WCPSS data: 2013-14 Student Locator file; 2013-14 mCLASS® data

Appendix E

Figure E1
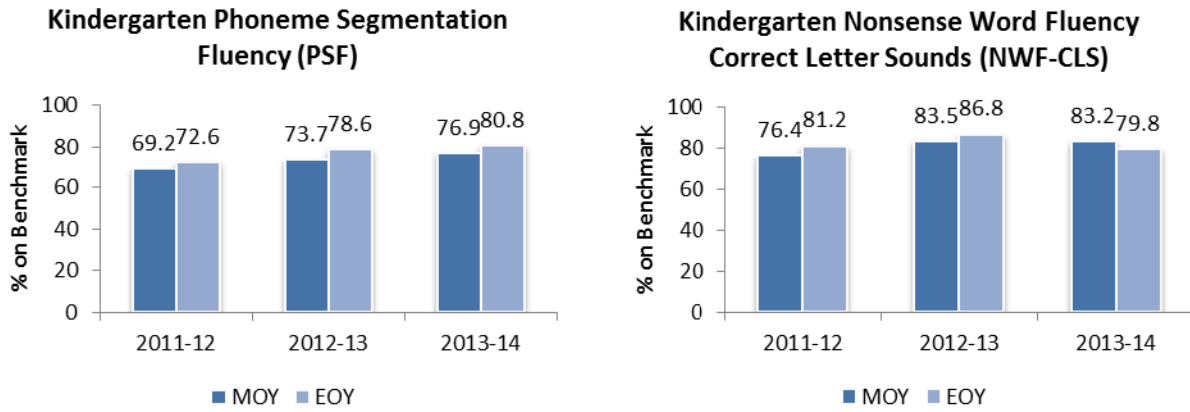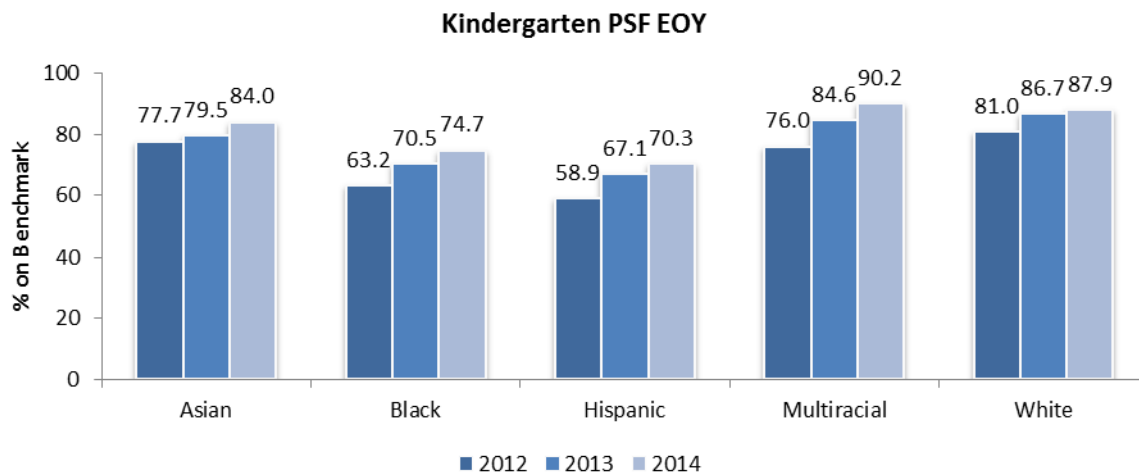*Percentage of Kindergarten Students Meeting Benchmark*

Kindergarten Phoneme Segmentation Fluency (PSF)

	MOY	EOY
2011-12	69.2	72.6
2012-13	73.7	78.6
2013-14	76.9	80.8

Kindergarten Nonsense Word Fluency Correct Letter Sounds (NWF-CLS)

	MOY	EOY
2011-12	76.4	81.2
2012-13	83.5	86.8
2013-14	83.2	79.8

Figure E2
Percentage of Kindergarten Students Meeting Benchmark by Ethnicity

Kindergarten PSF EOY

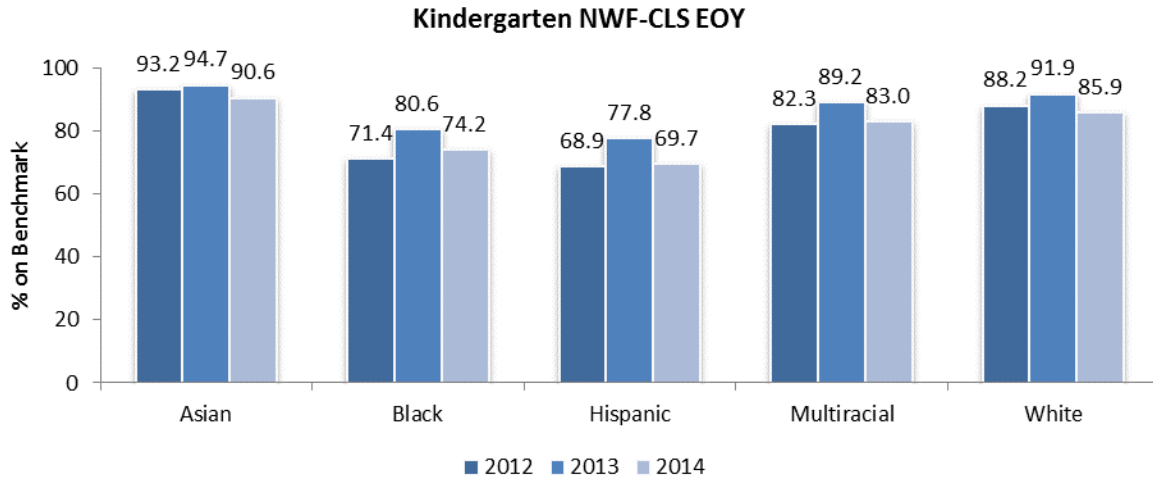|	2012	2013	2014
Asian	77.7	79.5	84.0
Black	63.2	70.5	74.7
Hispanic	58.9	67.1	70.3
Multiracial	76.0	84.6	90.2
White	81.0	86.7	87.9

Kindergarten NWF-CLS EOY

Figure E3
Percentage of Kindergarten Students Meeting Benchmark based on Students with Disabilities (SWD)
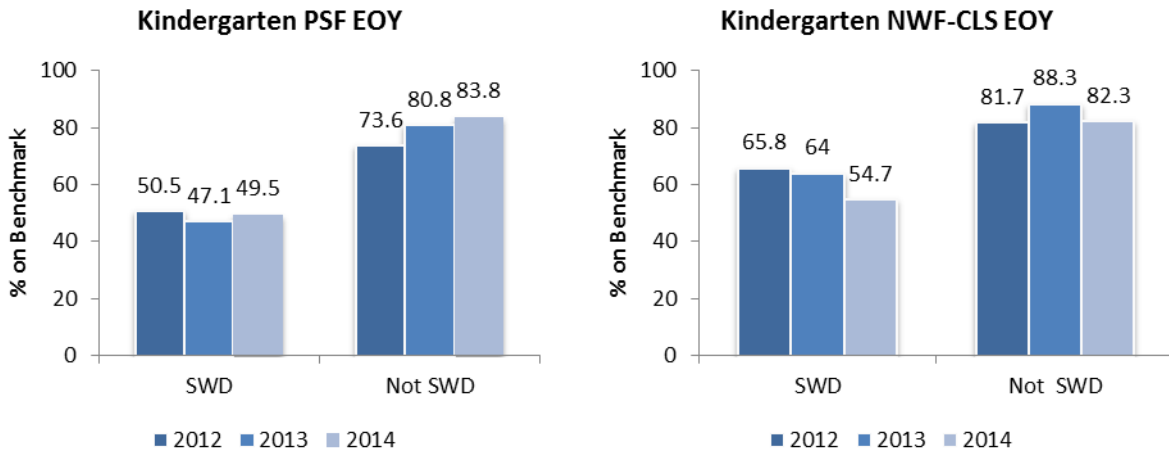

Figure E4
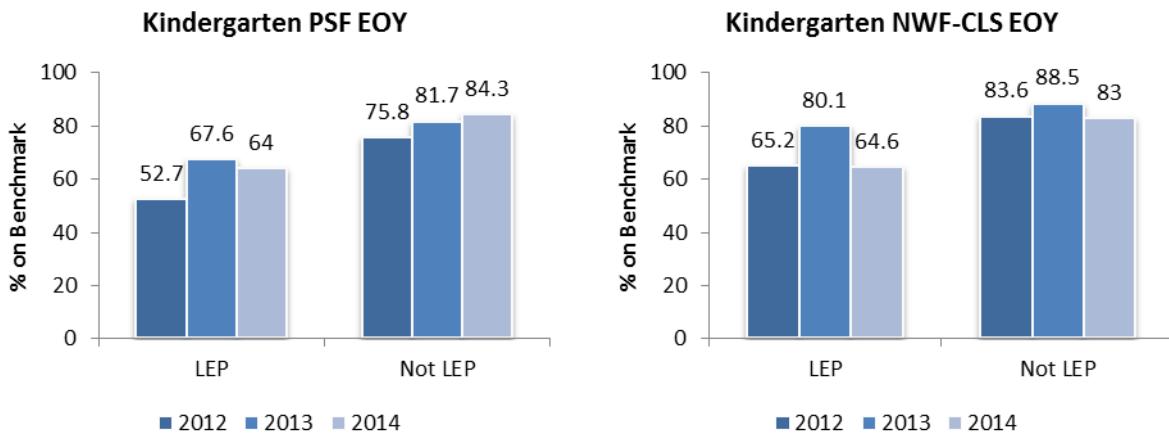*Percentage of Kindergarten Students Meeting Benchmark, Limited English Proficiency (LEP)*

Figure E5
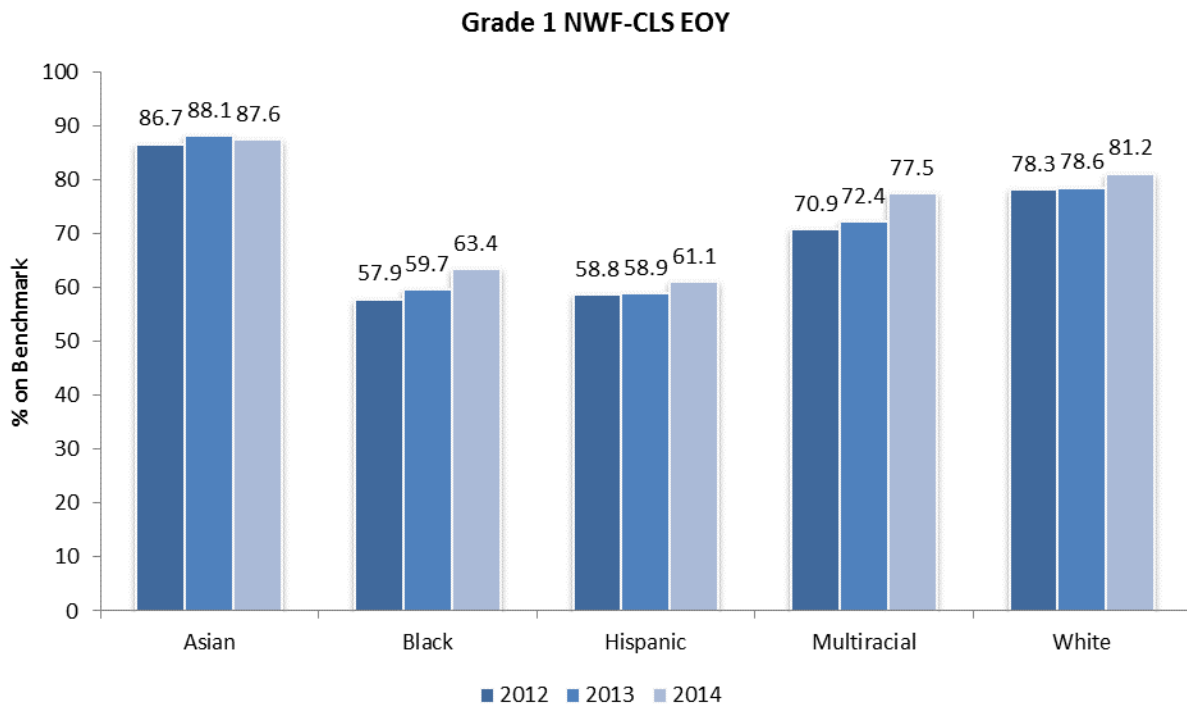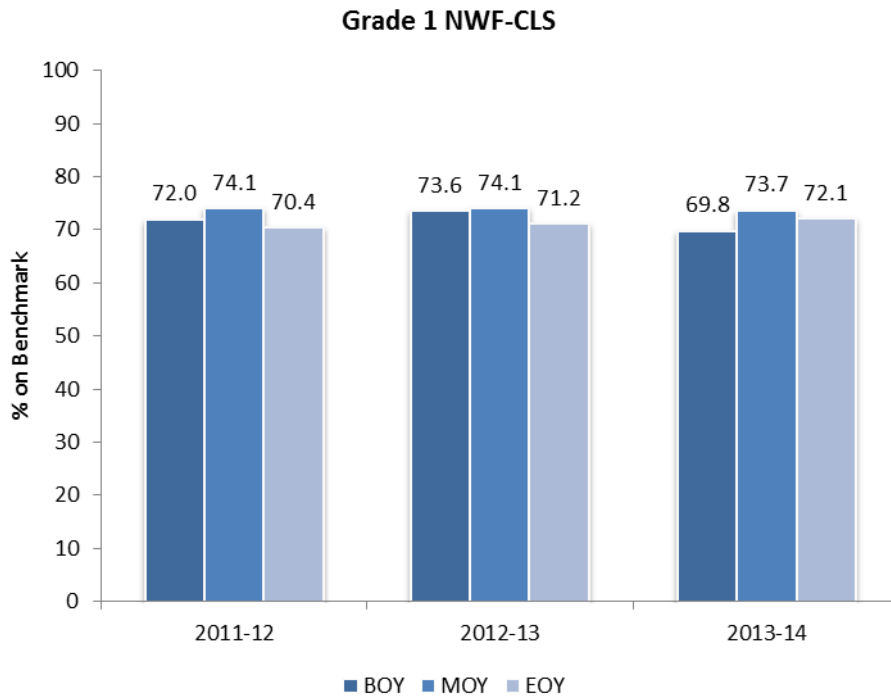*Percentage of Grade 1 Students Meeting Benchmark*

Grade 1 NWF-CLS

Grade 1 NWF-CLS EOY

Table E1
Raw Mean Scores on Indicators Across Years

PSF Kindergarten EOY Only

Variable	2011-12	2012-13	2013-14
Asian	49.95	50.87	49.36
Black	41.75	45.86	46.03
Hispanic/Latino	40.01	43.88	43.73
Multiracial	48.98	51.93	51.22
White	51.22	54.38	51.42
SWD	36.15	34.21	34.91
Not SWD	47.42	51.18	49.62
LEP	37.62	44.49	41.15
Not LEP	48.41	51.71	49.86
All	46.90	50.12	48.58

PSF First Grade BOY Only

Variable	2011-12	2012-13	2013-14
Asian	39.40	44.04	40.00
Black	33.19	36.43	35.54
Hispanic/Latino	34.49	35.94	34.24
Multiracial	38.55	43.17	39.96
White	43.12	46.04	42.01
SWD	29.21	29.96	27.20
Not SWD	39.64	42.98	39.85
LEP	30.87	36.34	30.67
Not LEP	40.16	43.26	40.05
All	38.92	41.88	38.84

NWF-CLS Kindergarten EOY Only

Variable	2011-12	2012-13	2013-14
Asian	74.31	78.08	74.44
Black	40.28	44.21	41.31
Hispanic/Latino	37.02	41.32	38.07
Multiracial	50.83	55.43	50.80
White	54.29	57.83	53.67
SWD	38.00	36.66	33.77
Not SWD	49.90	54.04	50.45
LEP	36.09	48.06	36.90
Not LEP	51.51	54.35	51.39
All	49.36	52.95	49.27

NWF-CLS First Grade EOY Only

Variable	2011-12	2012-13	2013-14
Asian	107.31	108.51	108.68
Black	69.20	68.37	72.98
Hispanic/Latino	67.89	68.89	71.90
Multiracial	82.01	83.48	86.50
White	86.50	87.20	91.03
SWD	64.52	57.57	57.75
Not SWD	81.81	83.14	86.78
LEP	65.29	74.69	67.84
Not LEP	83.11	82.55	87.06
All	80.66	80.90	84.45

Appendix F
Letterland Logic Model, 2011-12 to 2015-16

Need: Ongoing systematic K-2 research-based foundational literacy strand (word work). In 2012-13, 48% of students not meeting 3rd grade reading expectations based on 3rd grade EOG. Prior to implementation 80% or fewer students were meeting benchmark goals for early critical literacy skills on Phoneme Segmentation Fluency, Nonsense Word Fluency, and Oral Reading Fluency.

INPUTS	STRATEGIES	OUTCOMES – IMPACT		
		Short-Term	**Intermediate**	**Long-Term**
Implement Letterland, a phonics-based approach to teaching reading, writing and spelling to 3-8 year olds, in K-2 classrooms. • Find funding for program supplies • Letterland implementation team	**1. Professional Development** (webinar & face to face) • Kindergarten—spring & summer 2012 • Intervention teachers – fall 2012 & winter 2013 • K-2 Literacy teachers—spring 2012 • 1st Grade—spring, summer, & fall 2013 • **Ongoing PD** – new hires Kindergarten – 2012-13 • **Establish cohort of teacher trainers**—kindergarten—winter 2012 • **Presentation at principal meeting** to inform school administrators of Letterland—April 2012, March 2013 • **Educating SLP/OT and IRTs** for support during district wide meetings 2012-13, 2013-14 **2. Provide materials and resources** • Distribution and management of program materials—2012-13, 2013-14 • Blackboard site for program resources (videos/pacing guides/implementation tips)—2012-13, 2013-14 **3. Support** • On-site support through K-2 literacy teachers—March 2013 • Consulting with school-based PLTs to problem-solve implementation—2012-13, 2013-14	100% of teachers are trained in Letterland prior to implementing • K- spring and summer 2012✔ • 1st- spring, summer, fall 2013✔ • Intervention fall 2012 and winter 2013✔ • 2nd-spring/summer 2014✔ 80% of teachers are implementing on track (checking instructional status using survey instrument) • K—March 2013**X** (fall 2014 ✔) • 1st—spring 2014✔ • 2nd –spring 2015 Implementation fidelity tool created • K—March 2013 **X** (fall 2014 ✔) • 1st—fall 2013✔ • 2nd –fall 2014 School support personnel trained & inter-rater reliability established in order to monitor fidelity • K—March 2013 **X** (fall 2014 ✔) • 1st—fall 2013✔ • 2nd –fall 2014 The expectations for on-site support to teachers and district level support to schools communicated to all stakeholders. • July 2013✔ Critical instructional video(s) on Blackboard • July 2013✔ Improved resources available on Blackboard • **webinar access**✔ • **Implementation tips** —July 2013✔	100% of teachers implementing Letterland with fidelity • K—spring 2014**X** (89% high or moderate fidelity) • 1st—spring 2015 (Spring 2014 92% high or moderate fidelity) • 2nd –spring 2016 By spring 2013, the percentage of students exiting kindergarten meeting the end of grade benchmark will increase from 71% to 78% for PSF ✔ and from 81% to 87% for NWF ✔ By spring 2014, the percentage of students exiting 1st grade meeting the end of grade benchmark will increase from 81% to 87% for oral reading fluency **X** and from 80% to 87% for NWF **X (NWF split into NWF-CLS and NWF-WWR)**	By fall 2014, the percentage of students entering 2nd grade meeting the benchmark will increase from 63% to 70% for NWF-WWR **X** By spring 2015, the percentage of students exiting 2nd grade meeting the end of grade benchmark will increase from 73% to 80% for oral reading fluency By spring 2016, the percentage of 3rd grade students reading on grade level as measured by the reading EOG will increase from 52% to 60%