



# Evaluation of the *World of Wonders* Literacy Program

September 2020

Submitted to:  
McGraw-Hill Education

ICF Contributing Authors:

Samantha Salvador, MA

James Demery, Ph.D

Barbara O'Donnell, Ph.D.

Erica McCoy, M.P.A

**Table of Contents**

**Executive Summary** ..... **i**

**Introduction** ..... **1**

**About World of Wonders**..... **1**

**World of Wonders Evaluation**..... **3**

    Evaluation Questions ..... 3

    Demographics Data Sources ..... 4

        School Characteristics ..... 4

        Kindergarten Student Demographics..... 7

    Literacy Measures Data Sources ..... 8

        Early Literacy Quick Assessment ..... 8

        Developmental Reading Assessment 2<sup>nd</sup> Edition ..... 9

**Findings** ..... **12**

    Pre-K Literacy Achievement..... 12

        Literacy Achievement Pre- versus Post-*World of Wonders* Implementation ..... 12

        To what extent did pre-K students meet target expectations? ..... 16

        Were school characteristics associated with changes in pre-K literacy achievement? ..... 18

    Kindergarten Literacy Achievement ..... 20

        Comparison of Students Based on *World of Wonders* Exposure ..... 20

        Were student demographics associated with DRA2 literacy achievement? ..... 22

        Was full-day versus half-day district pre-K associated with DRA2 literacy achievement? ..... 22

**Conclusions and Recommendations**..... **23**

    Limitations ..... 25

**Appendix A: Multiple Regression Analyses** ..... **28**

## Executive Summary

*World of Wonders (WoW)* is an early literacy program designed to help children ages 3 to 5 years develop the foundational social and academic skills they need to succeed in kindergarten and beyond. This two-year study was conducted to determine the extent to which the program accomplishes that goal during the first two years of implementation in a school district. This study provides evidence for *Every Student Succeeds Act (ESSA)* Tier III promising evidence, with qualifications, that *World of Wonders* supports the early literacy achievement of students. In pre-kindergarten (pre-K) students, there were statistically significant improvements from beginning-of-year (BOY) to end-of-year (EOY) on each of the Early Literacy Quick Assessment (ELQA) subscales within both implementation years of the study. Among students who entered kindergarten in the district, those students who had attended district pre-K implementing *World of Wonders* scored higher on average on the BOY Direct Reading Assessment (DRA2), even after accounting for student demographics.

ESSA, the 2015 revision of the federal law regulating PreK-12 education in the United States, strongly recommends that schools and districts implement interventions with documented evidence of effectiveness (i.e., evidence-based interventions). ESSA classifies research studies in one of four tiers depending on the rigor of design and methods used in the research and the strength of the evidence they produce: Tier 1 (Strong Evidence); Tier 2 (Moderate Evidence); Tier 3 (Promising Evidence); and Tier 4 (Demonstrates a Rationale).

Promising evidence was based primarily from a nonequivalent control group posttest only analysis of BOY literacy assessment data. BOY DRA2 data from the 2018–19 school year for 1,155 kindergarten students provided by a mid-sized school district in central Oklahoma, 742 (64%) of whom attended a district pre-K implementing *World of Wonders* and 413 (36%) whom did not attend district pre-K. Student-level demographics were provided for the kindergarten sample (i.e., race, gender, English language learner, special needs). The two groups of kindergarten students did not differ significantly from one another demographically except for a higher percentage of students identified as special needs who attended district pre-K and experienced *World of Wonders*.

The study also reports on data analyzed from BOY to EOY for two years of pre-K ELQA. The district provided ELQA data for 306 pre-K students from the first year (2017-18) that *World of Wonders* was implemented in the district and from 95 pre-K students in implementation Year 2 (2018-19). In Year 1, 15 of the district's 17 elementary schools participated; in Year 2, 9 schools in the district provided pre-K data for the study. Student-level demographic data were not available for linkage with ELQA. Analyses did include school-level demographics including total school enrollment, percent of Grade 3 students proficient in ELA, percent of minority students, percent of ELL students, percent of students identified as special education, and percent of students eligible for the federal free and reduced-price lunch program.

Qualifications regarding the study findings include the inability to separate out pre-K attendance from *World of Wonders* implementation and the inability to control for factors

beyond provided demographics, including factors that may have influenced families decisions to enroll students in district pre-K or not.

### **Summary of Key Findings**

- There were significantly higher BOY literacy scores on the DRA2 among kindergarten students who entered kindergarten after being exposed to *World of Wonders* in pre-K compared to their counterparts who had not attended district pre-K and presumably were not exposed to the program. The finding held true even after adjusting for student demographics and regardless of whether students attended full-day or half-day district pre-K.
- Within each year of pre-K (i.e., 2017-17 and 2018-19), there were significant increases from BOY to EOY on each of the seven ELQA subscales, suggesting that *World of Wonders* supports expected literacy development. BOY ELQA scores were the best predictor of EOY ELQA scores.
- Regarding meeting target ELQA scores (Letter Sounds do not have target scores and were excluded from these analyses):
  - In year 1 of implementation, significantly more pre-K students met the targets for Rhyming, Uppercase Letter sand Lowercase Letters at EOY, as compared to BOY.
  - In both years, pre-K students on average met targets at EOY for the literacy skills of Receptive Vocabulary and Uppercase Alphabet. In Year 1, they also met the EOY target for Lowercase Alphabet; in Year 2, they also met the EOY target for Expressive Vocabulary.
  - Significantly fewer students at EOY as compared to BOY met targets for Expressive and Receptive Vocabulary in both years, although students on average met EOY targets for Receptive Vocabulary in both years and Expressive Vocabulary in Year 2.
  - Target EOY scores were not met in either year for Print Concepts or Rhyming; however, a more rigorous study would allow for the determination of whether *World of Wonders* could be improved to better emphasize these topics or if teachers needed additional supports to implement activities related to these topics. It is also possible that the ELQA target may be too high for these subscales.

DISCLAIMER: Literacy can improve due to several factors, including innate ability, home experiences and attending preschool in genera. It is unknown what type if any preschool experience kindergarten students had who did not attend the district preschool. Even taking these factors into account, *World of Wonders* may have contributed to improvements in student outcomes including arriving at kindergarten higher in literacy development.

## Introduction

Oklahoma launched its Early Childhood Four-Year-Old Program in 1980 and by 1998 nearly all school districts in Oklahoma (99%) were providing free preschool to four-year-old children.<sup>1</sup> This report provides an account of a study that took place in a mid-sized school district in central Oklahoma that began implementing the *World of Wonders* literacy program in the fall of 2017 in all its elementary schools.<sup>2</sup> At the end of the 2017-18 school year, phase one of the study, McGraw-Hill Education (MHE) contracted with ICF to help them determine the extent to which students in the district who had been exposed to *World of Wonders* in preschool acquired the basic early literacy skills they would need to be successful in kindergarten and beyond.

The first year of the study focused solely on the performance of pre-K students. At the end of the 2018-19 school year the study expanded to phase two which included an examination of the potential effects of *World of Wonders* on literacy skills at the beginning of kindergarten among students in the district who had been exposed to the program as pre-K students in 2017-18 as compared to students who had not attended pre-K in the district. That is, were children who attended the district's preschools and experienced *World of Wonders* better prepared at the start of kindergarten than their non-attending peers regarding literacy skills? Additional pre-K data were also provided for this school year. Two literacy assessments were used to estimate the impact of the *World of Wonders*. In both phases, pre-K teachers administered the Early Literacy Quick Assessment (ELQA) to their students in the beginning-of-year (BOY) and end-of-year (EOY). In phase 2,<sup>3</sup> kindergarten teachers administered the Developmental Reading Assessment, second edition (DRA2) to their students at BOY.<sup>4</sup> The findings and subsequent conclusions presented in this report are based on the results of those assessments.

## About World of Wonders

*World of Wonders* is the early literacy (preschool) component of MHE's *Wonders* comprehensive literacy program. The program's curriculum includes content in science, math, social studies, music, and motion. The program's learning objectives were influenced by those of the Head Start Early Learning Outcomes Framework

---

<sup>1</sup> See [https://sde.ok.gov/sites/default/files/documents/files/Oklahoma\\_YB2018.pdf](https://sde.ok.gov/sites/default/files/documents/files/Oklahoma_YB2018.pdf)

<sup>2</sup> ICF would like to thank the district for their assistance in providing data for this study and for responding to questions regarding the data. See <https://www.mheducation.com/prek-12/program/wonders-2020/MKTSP-BGA07M0.html> for additional information on *World of Wonders* and its follow-on curriculum *Wonders* for grades K-5,

<sup>3</sup> See <https://outreach.ou.edu/educational-services/education/elqa/> for additional information on ELQA.

<sup>4</sup> See [http://assets.pearsonschool.com/asset\\_mgr/current/20139/DRA2\\_Technical\\_Manual\\_2012.pdf](http://assets.pearsonschool.com/asset_mgr/current/20139/DRA2_Technical_Manual_2012.pdf) for additional information on DRA2.

language and literacy component that was developed to help learners, ages 3 to 5, build pre-reading skills.<sup>5</sup>

- Child attends to communication and language from others
- Child understands and responds to increasingly complex communication and language from others
- Child varies the amount of information provided to meet the demands of the situation
- Child understands, follows, and uses appropriate social and conversational rules
- Child expresses self in increasingly long, detailed, and sophisticated ways
- Child understands and uses a wide variety of words for a variety of purposes
- Child shows understanding of word categories and relationships among words
- Child demonstrates awareness that spoken language is composed of smaller segments of sound
- Child demonstrates an understanding of how print is used and the rules that govern how print works
- Child identifies letters of the alphabet and produces correct letter sounds associated with letters
- Child asks and answers questions about a book that was read aloud
- Child writes for a variety of purposes using increasingly sophisticated marks

Table 1 provides an overview of the logic model guiding *World of Wonders* development as well as its intended short-term outcomes.

---

<sup>5</sup> See <https://s3.amazonaws.com/ecommerce-prod.mheducation.com/unitas/school/explore/sites/reading-wonders/wow-head-start-correlations.pdf> and <https://eclkc.ohs.acf.hhs.gov/sites/default/files/pdf/elof-ohs-framework.pdf>. MHE has also correlated across a range of skills, although language and literacy are the focus here.

**Table 1. World of Wonders Logic Model**

Resources	Activities	Outputs	Short-term Outcomes
<p><i>World of Wonders</i> curriculum:</p> <ul style="list-style-type: none"> <li>Balanced literacy</li> <li>Science</li> <li>Math</li> <li>Social studies</li> <li>Music</li> <li>Motion</li> <li>Social-emotional development</li> </ul> <p>Pre-K teachers</p> <p>Early literacy screening and monitoring assessments</p> <p>School/district budget</p> <p>Family/in-home support for literacy</p>	<p>Pre-K instruction in early literacy skills:</p> <ul style="list-style-type: none"> <li>Print concepts</li> <li>Phonological Awareness</li> <li>Vocabulary</li> <li>Alphabet knowledge</li> </ul> <p>Administer routine screening and monitoring of early literacy skills</p> <p>Use assessment results to inform instruction</p>	<p>Better focused literacy instruction</p> <p>Increased use of early literacy assessments results to drive instruction and provide additional support in literacy instruction as needed</p> <p>Increased knowledge of early literacy skills</p>	<p>Teachers: Improved early literacy instruction</p> <p>Pre-K children: Increased level of kindergarten readiness</p>

## World of Wonders Evaluation

This study was conducted primarily to gather evidence on the potential of *World of Wonders* to improve the basic literacy skills of pre-K students and send students to kindergarten better prepared regarding literacy skills. To determine the impact of the program in the first year of implementation the ICF evaluation team used a one-group pretest-posttest evaluation design. This design was selected because there was no comparison group of participants (i.e., students who had not been exposed to *World of Wonders*). The one-group pretest-posttest design is commonly used by researchers seeking to establish Tier 3 Promising Evidence of a program’s efficacy in accordance with the Every Student Succeeds Act (ESSA) standards for evidence-based interventions.<sup>6</sup> To estimate the impact of the program on the BOY performance of kindergarten students who had been exposed to *World of Wonders* in pre-K the team used a nonequivalent control group posttest only design. This design was possible because the district provided literacy achievement data for a comparison group of kindergarten students. As noted earlier, pre-K teachers used the

<sup>6</sup> See <https://www2.ed.gov/policy/elsec/leg/essa/guidanceuseseseinvestment.pdf>

ELQA to assess literacy outcomes achievement among their students and the DRA2 was used to gauge the literacy achievement of kindergarten students.

### **Evaluation Questions**

The evaluation was designed to determine the extent to which early literacy outcomes were associated with exposure to *World of Wonders*, given the available data. Where possible, analyses on subgroups of students (e.g., race, gender, economically disadvantaged status, special education status, and English language learner status) were conducted in order to determine whether any given subgroup of students may have been impacted differentially. Student-level demographic data was not provided for pre-K students. Therefore, school-level demographic characteristics were used to examine whether these were associated with literacy outcomes. Several student-level demographics were provided associated with the BOY kindergarten data. In the end, four primary evaluation questions were posed to address these concerns.

1. How did literacy achievement from BOY to EOY vary when comparing the ELQA results of pre-K students prior to the implementation of *World of Wonders* (i.e., 2016-17), after one year of implementation (i.e., 2017-18), and after two years of implementation (i.e., 2018-19)?
2. To what extent were school-level characteristics such as percent minority related to pre-K literacy achievement?
3. How did the BOY literacy assessment of kindergarten students who were exposed to *World of Wonders* in pre-K compare to that of kindergarten students who had not been exposed to the program in pre-K?
4. To what extent were student-level demographics related to BOY literacy achievement in Kindergarten in students who were exposed to *World of Wonders* in pre-K as compared to that of kindergarten students who had not been exposed to the program in pre-K ?

Table 2 aligns the evaluation questions to the corresponding data methods, sources, and analytic techniques used to answer them.

**Table 2. Evaluation Questions, Data Sources, and Analyses**

Evaluation Question	Data Source	Analysis
How did literacy achievement from BOY to EOY vary when comparing the ELQA results of pre-K students prior to the implementation of <i>World of Wonders</i> (i.e., 2016-17), after one year of implementation (i.e., 2017-18), and after two years of implementation (i.e., 2018-19)?	2016-17, 2017-18, and 2018-19 ELQA results	Paired-sample t-tests; Analysis of Covariance (ANCOVA), McNemar's chi-square test
To what extent were school-level characteristics such as percent minority related to BOY to EOY pre-K literacy achievement?	2017-18 and 2018-19 ELQA results and 2017-18 and 2018-19 School Report Cards	Regression analysis
How did the BOY literacy assessment of kindergarten students who were exposed to <i>World of Wonders</i> in pre-K compare to that of kindergarten students who had not been exposed to the program in pre-K?	2018-19 DRA results	Chi-square test, t-test
To what extent were student-level demographics related to BOY literacy achievement in kindergarten in students who were exposed to <i>World of Wonders</i> in pre-K as compared to that of kindergarten students who had not been exposed to the program in pre-K ?	2018-19 DRA results Student-level demographics	Chi-square test, Logistic Regression, ANCOVA

### Demographics Data Sources

For the pre-K data in both years, student-level demographics were not available; therefore, school-level characteristics were used as a proxy. Student-level demographics were available for kindergarten students.

### School Characteristics

Table 3 displays the demographic profiles of the schools with pre-K students that participated in the study; School P did not participate in Year 1.<sup>7</sup> School-level variables included the average enrollment size, percentage of students rated as proficient on the Grade 3 Oklahoma School Testing Program English Language Arts (OSTP ELA) test, and percentage of students identified as each of the following: racial minority, English language learner (ELL), qualifying for special education, and qualifying for free/reduced lunch.

<sup>7</sup> All students in each district pre-K classroom were exposed to the *Wonders* curriculum. Parent consent was required to receive student achievement data. Therefore, the study is limited to those students in each school/classroom who provided parent consent.

**Table 3. Demographic Profiles of Participating Elementary Schools  
in 2017-18 and 2018-19**

School	School Year	Average Enrollment	% Proficient Gr 3 OSTP ELA	% Racial Minority	% ELL	% Special Educ.	% Free/Reduced Lunch
School A	2017-18	599	48.0%	52.3%	18.0%	13.8%	67.5%
	2018-19	611	63.0%	53.3%	19.7%	14.8%	69.2%
School B	2017-18	585	93.0%	34.0%	4.6%	14.2%	32.8%
	2018-19	592	86.0%	33.0%	6.0%	15.4%	33.8%
School C	2017-18	617	60.0%	38.5%	3.7%	9.3%	52.6%
	2018-19	588	65.0%	38.1%	3.7%	11.7%	46.5%
School D	2017-18	526	43.0%	55.4%	17.7%	13.0%	66.5%
	2018-19	505	37.0%	58.2%	21.6%	13.9%	70.0%
School E	2017-18	413	43.0%	34.5%	2.7%	18.6%	55.8%
	2018-19	421	32.0%	36.5%	3.8%	18.4%	56.1%
School F	2017-18	499	24.0%	54.2%	4.7%	16.2%	88.2%
	2018-19	481	24.0%	54.0%	7.2%	16.9%	88.5%
School G	2017-18	274	50.0%	51.3%	12.2%	22.5%	66.1%
	2018-19	296	32.0%	50.7%	13.4%	22.5%	62.8%
School H	2017-18	407	51.0%	55.3%	6.5%	19.7%	72.1%
	2018-19	400	29.0%	52.6%	7.9%	17.6%	71.0%
School I	2017-18	354	79.0%	33.9%	3.4%	15.5%	62.7%
	2018-19	390	75.0%	36.9%	7.6%	13.2%	28.8%
School J	2017-18	451	49.0%	38.3%	6.8%	13.4%	51.2%
	2018-19	481	56.0%	39.5%	7.9%	13.1%	51.4%
School K	2017-18	526	58.0%	44.2%	4.3%	14.2%	56.3%
	2018-19	557	64.0%	47.5%	5.1%	15.5%	60.3%
School L	2017-18	609	73.0%	35.0%	2.3%	17.3%	20.0%
	2018-19	630	46.0%	36.1%	4.3%	17.7%	20.6%
School M	2017-18	422	70.0%	43.4%	4.0%	16.3%	36.1%
	2018-19	469	67.0%	42.4%	11.1%	11.1%	41.9%
School N	2017-18	558	66.0%	38.5%	2.3%	12.9%	37.1%
	2018-19	564	61.0%	41.1%	4.3%	14.5%	37.5%
School O	2017-18	321	32.0%	44.3%	3.7%	19.1%	62.7%
	2018-19	318	39.0%	43.0%	6.8%	20.1%	73.5%

Table 3. (continued)

School	School Year	Average Enrollment	% Proficient Gr 3 OSTP ELA	% Racial Minority	% ELL	% Special Educ.	% Free/Reduced Lunch
School P <sup>a</sup>	2017-18	NA	NA	NA	NA	NA	NA
	2018-19	196	25.0%	34.7%	3.6%	21.4%	49.5%
District	<b>2017-18</b>	<b>15,942</b>	<b>58.0%</b>	<b>41.5%</b>	<b>4.6%</b>	<b>17.2%</b>	<b>48.8%</b>
	<b>2018-19</b>	<b>15,963</b>	<b>55.0%</b>	<b>42.7%</b>	<b>5.8%</b>	<b>17.6%</b>	<b>49.6%</b>

**Source:** District School Reports (<https://www.edprofiles.info/report-card/district/list>)

**Notes:** Oklahoma School Testing Program English Language Arts (OSTP ELA); English language learner (ELL). To comply with the student privacy protections of the *Family Educational Rights and Privacy Act* (FERPA), the following redaction rules were applied by the district

1. Redact all proficiency categories if the total n size is <10.
2. If the n size is >10, redact proficiency categories with  $\leq 3$  students.
3. If one proficiency level is redacted, an additional proficiency level must be complementary suppressed to not unintentionally reveal a redaction.
4. If one proficiency percentage is 100%, then an additional category will need to be redacted as to not reveal that there are zero in other categories.
5. If the sum of the n size of two categories is equal to the total n, an additional category will need to be redacted as to not reveal that there are zeros in other categories

<sup>a</sup>School P did not participate in the first year of the study.

Although this is not considered a high poverty school district according to the guidelines of the National School Lunch Program,<sup>8</sup> the majority of students at 11 of the 15 schools in 2017-18 qualified for free or reduced price meals; the same was true for 10 of the 16 schools in 2018-19. Average annual enrollment varied widely among the schools from a low of 196 students at School P in 2018-19 to a high of 630 at School L in 2018-19. Half the schools (8) experienced a drop in the percentage of students who demonstrated proficiency on the 2018-19 Grade 3 OSTP ELA, a measure of the extent to which the school may be struggling overall in literacy development. The drop in scores could be related to the fact the state revised the tests in Spring 2017 to make them more rigorous. There were also a few other notable changes in the demography of the participating schools; for example:

- The percentage of students classified as ELL increased at all but one school, School G, where it remained unchanged. At four schools the ELL population increased considerably over the course of the study: School F (4.7% to 7.2%); School I (3.4% to 7.6%); School M (4.0% to 11.1%); and School O (3.7% to 6.8%).
- The percentage of students identified as special education based on having an individualized education program (IEP) or Section 504 plan at each school remained stable; however, at School M the percentage of students identified as

<sup>8</sup> See [https://nces.ed.gov/programs/coe/indicator\\_clb.asp](https://nces.ed.gov/programs/coe/indicator_clb.asp)

special needs dropped considerably from 16% in 2017-18 to 11% in the 2018-19 school year.

- At School I the percentage of students identified as being from economically disadvantaged households based on access to free/reduced lunch fell from nearly two-thirds (63%) of the population in the first year of the study to 29% in 2018-19.

### Kindergarten Student Demographics

Available student-level demographics included race, gender, ELL status, special education status.<sup>9</sup> While some data were provided for 1,317 kindergarten students in 2018-19, data were limited for 162 cases who were excluded from additional analysis. Of the remaining 1,155 students, 742 (64%) attended a district pre-K program in 2017-18 (396 attended full day, 346 attended half day) while 413 (36%) did not attend a district pre-K program.

Based on kindergarten students identified by race, 79 (6.8%) were African American, 199 (17.2%) were Hispanic, 616 (53.3%) were White, and 261 (22.6%) were 'Other' (American Indian, Asian, Multi-racial, or Pacific Islander). When comparing students who attended a district pre-K program to students who did not, we found that the district pre-K sample had a slightly higher percentage of African American students (7.0% vs. 6.5%), a similar percentage of Hispanic students (17.1% vs. 17.4%), a higher percentage of 'Other' races (24.1% vs. 19.9%) and a lower percentage of White students (51.7% vs. 56.2%). However, these slight differences in the distribution of race between the two groups were not statistically significant ( $\chi^2(3) = 3.2, p = .35$ ).

District-wide, approximately half (49.4%) of kindergarten students were female, 14.5% were categorized as ELL, and 8.5% received special education services. The distribution of gender (pre-K: 48.7% female, non-pre-K: 50.6% female) and ELL status (pre-K: 15.0% ELL, non-pre-K: 13.8% ELL) also did not differ significantly between the two groups ( $\chi^2(1) = 0.4, p = .52, \chi^2(1) = 0.3, p = .59$ , respectively). However, the distribution of students receiving special education services did differ significantly ( $\chi^2(1) = 5.9, p < .05$ ), with a higher percentage of students who attended district pre-K receiving special education services in kindergarten (10.0%) than students who did not attend district pre-K (5.8%).

Collectively, the two groups of kindergarten students (i.e., those who attended district pre-K versus those who did not attend district pre-K) were very similar to one another demographically. The only demographic difference between the two groups was the higher percentage of kindergarten students identified as needing special education who attended district pre-K as compared to those who did not attend (10% and 6%,

---

<sup>9</sup> The school identified variable for ELL was categorized as Limited English Proficiency.

respectively). This suggests that families who selected to attend district pre-K were demographically similar to families that did not select district pre-K.

## **Literacy Measures Data Sources**

### **Early Literacy Quick Assessment**

The ELQA is a web-based assessment system that is used to diagnose and monitor four basic early literacy skills: 1) Print Concepts; 2) Vocabulary; 3) Phonological Awareness; and 4) Alphabet Knowledge.<sup>10</sup> During the 2013-14 school year the Educational Training, Evaluation, Assessment, and Measurement (E-TEAM) Department at the University of Oklahoma conducted a validity study of the ELQA. The study was undertaken primarily to establish the construct validity, concurrent validity, and reliability of the assessment. Some 201 Oklahoma pre-K children were used to examine construct validity and reliability, and another 423 pre-K students were used to establish the assessments' concurrent validity.<sup>11</sup>

Regarding the ELQA, construct validity refers to how well its measures assess the early literacy skills it purports to measure. As presented in Table 4, principal components analyses resulted in six separate ELQA early literacy measures that are aligned with the four basic literacy skills; a seventh ELQA measure—Letter Sounds—was not aligned with the four basic literacy skills. The E-TEAM used the Peabody Picture Vocabulary Test (PPVT) and the Phonological Awareness Literacy Screening (PALS) to establish the concurrent validity of the ELQA. Concurrent validity refers to how well an assessment (i.e., the ELQA) assesses a psychological construct in comparison to other well-established measures of that construct in question. Results showed that the ELQA was sufficiently correlated with both the PPVT (.46) and the PALS assessment (.92) to establish concurrent validity as an assessment of early literacy skills. Internal consistency reliability of the ELQA was established using Cronbach's alpha and Guttman split-half reliability. The final Cronbach's alpha reliability coefficients for each ELQA measure were as follows: Print Concepts (.88); Uppercase Alphabet (.97); Lowercase Alphabet (.90); Letter Sounds (.95); Expressive Vocabulary (.88); Receptive Vocabulary (.79); and Rhyming (.89).<sup>12</sup>

---

<sup>10</sup> See <http://www.nifl.gov/nifl/publications/pdf/NELPReport09.pdf>

<sup>11</sup> Early Literacy Assessment (ELQA) Technical Report: Validity and Reliability (2014).  
[https://elqa.ou.edu/media/filer\\_public/b7/f0/b7f0be4e-0ccb-45cf-8765-732a303cbadc/elqa\\_technical\\_report](https://elqa.ou.edu/media/filer_public/b7/f0/b7f0be4e-0ccb-45cf-8765-732a303cbadc/elqa_technical_report).

<sup>12</sup> Ibid

**Table 4. Alignment of ELQA Measures with Basic Early Literacy Skills**

Basic Early Literacy Skills	ELQA Measures
Print Concepts	Print Concepts (10 items)
Vocabulary	Expressive Vocabulary (25 items) Receptive Vocabulary (20 items)
Phonological Awareness	Rhyming (10 items)
Alphabet Knowledge	Uppercase Alphabet (26 items) Lowercase Alphabet (26 items)

**Source:** ELQA Teacher's Guide (2015)

**Note:** The ELQA Letter Sounds measure is not directly aligned with the basic early literacy skills.

The ELQA uses 10 items to assess Print Concepts, some of which ask children to identify specific letters, numbers and words in selected images. They are also asked to identify where on a page they would start reading and in which direction they would read (e.g., left-to-right or right-to-left). The ELQA assesses skills in two dimensions of vocabulary: Expressive Vocabulary and Receptive Vocabulary. To assess Expressive Vocabulary children are shown pictures and asked to name the objects they see; the number of items used to assess Expressive Vocabulary decreased from 25 in 2017-18 to 20 for the 2018-19 school year. For Receptive Vocabulary students are shown 20 pictures and asked to point to the one that represents the concept their teacher presents to them.<sup>13</sup>

The ELQA also assesses phonological awareness, which may be thought of as the ability to detect, manipulate, or analyze the auditory aspects of spoken language independent of meaning. The ELQA uses 10 Rhyming items to assess phonological awareness. Some of those items ask children to determine whether each of five pairs of words rhyme (i.e., rhyme recognition) and to provide a word that rhymes with five different spoken words (i.e., rhyme generation). Finally, Alphabet knowledge--the ability to understand and identify the names and sounds associated with printed letters—is assessed with 52 items including both Uppercase and Lowercase Alphabet letters.<sup>14</sup>

### Developmental Reading Assessment 2<sup>nd</sup> Edition

The DRA2 is a set of Grades K-8 assessments designed to be individually administered by teachers and reading specialists to help determine whether students are reading on, above, or below grade level. DRA2 scores can be used to assist teachers in planning literacy activities and in providing targeted interventions and supplemental support as needed. Pearson Education, the developer of the DRA2, cautions users to not use the

<sup>13</sup> Ibid. Information regarding the reduction in Expressive Vocabulary items was provided in the teacher's guide provided by the school for the study.

<sup>14</sup> Early Literacy Assessment (ELQA) Technical Report: Validity and Reliability (2014).  
[https://elqa.ou.edu/media/filer\\_public/b7/f0/b7f0be4e-0ccb-45cf-8765-732a303cbadc/elqa\\_technical\\_report](https://elqa.ou.edu/media/filer_public/b7/f0/b7f0be4e-0ccb-45cf-8765-732a303cbadc/elqa_technical_report).

DRA2 as the sole measure of reading proficiency when making high-stakes decisions such as those regarding student promotion or retention.

During the assessment, teachers are not supposed to let students skip portions of the test, even if they are reading accurately and effortlessly. The DRA2 assessment begins with the Reading Engagement Survey, which consists of several questions designed to determine the student’s level of engagement with reading. Engaged readers are those who read often, know books and authors, and have goals for themselves as readers. After the engagement survey, teachers ask students to read a fiction or nonfiction selection (or selections) from the DRA2 Benchmark Assessment Books and then retell what they have read. Each Benchmark Assessment Book is identified on a scale from A through 80. As the levels increase, so does the difficulty level for each reading selection.

The time it takes to administer the DRA2 depends on how well the level of the text matches the student’s independent reading level. A student is an independent reader when he or she can successfully engage with the text without assistance from the teacher. Just below the independent reading level (i.e., slightly below grade level) is the instructional reading level. Students scoring at the instructional level may find the text challenging and often require some assistance from the teacher. If a student’s performance is scored as independent, teachers can identify the reading stage as emergent, early, transitional, or extending, depending on the level of the DRA2 assessment administered. For emergent readers who rate at Level A-3, the ELQA administration time is typically 5-10 minutes; for early readers rated at Levels 4-12, the administration time is about 10 minutes; for transitional readers, those rated at Levels 14-24, the administration time is 15-20 minutes; for extending readers (Levels 28-40) the administration time is approximately 45-60 minutes. Table 5 shows the expected level of reading proficiency for kindergarten and Grade 1 students. Note that for kindergarten students achieving at any level above A at BOY indicates the student is above expectations.

**Table 5. Kindergarten and Grade 1 Reading Expectations**

Grade	Time	Independent	Instructional	Intervention
Kindergarten	Fall (BOY)	Pre-A	--	--
	Mid-Year	1	A	Pre-A
	Spring (EOY)	3	2	1
Grade 1	Fall (BOY)	3	2	1
	Mid-Year	8	6	4
	Spring (EOY)	16	14	12

Source: See [http://assets.pearsonschool.com/asset\\_mgr/current/20139/DRA2\\_Technical\\_Manual\\_2012.pdf](http://assets.pearsonschool.com/asset_mgr/current/20139/DRA2_Technical_Manual_2012.pdf)

A student's reading level is determined by his or her scores on the reading engagement survey, oral reading fluency, and reading comprehension. Oral reading fluency is a measure of: 1) how students sound when they read text in terms of mood, pace and tension (i.e., expression); 2) how often they pause when reading a sentence (i.e., phrasing); 3) the number of words read per minute (i.e., rate); 4) the percentage of words read correctly (i.e., accurately); 5) the extent to which students self-correct miscues while reading (i.e., monitoring/self-corrections); and 6) the extent to which students problem-solve unknown words through using cues or teacher assistance (i.e., problem-solving unknown words). Comprehension is the measure of students' ability to retell and understand the text including main ideas, characters, events, or topics.

**Reliability.** Four methods were used to establish the reliability of DRA2 scores: internal consistency reliability, parallel equivalency reliability, test-retest reliability, and inter-rater reliability.<sup>15</sup> Internal consistency reliability depicts the extent to which different items are measuring the same behavioral trait. For oral reading fluency the Cronbach's alpha reliability coefficients of the various DRA2 reading levels ranged from a high of .85 to a low of .54. Regarding comprehension, the coefficients ranged from a high of .85 to a low of .58. Parallel equivalency reliability refers to the extent to which students get similar fluency or comprehension scores on fiction and non-fiction passages of the same level. Multiple analysis of variance (MANOVA) was used to determine parallel equivalency. Results of the MANOVA showed that, except for oral fluency at level 4 and reading comprehension at level 34, the fiction and non-fiction passages were equivalent and can be used interchangeably. Test-retest reliability is a measure of how stable a student's scores are from one administration to the next. Results showed that there were no statistically significant differences between the first administration and the second one (14 days later). Correlation coefficients between the two test periods ranged from .93 to .99. Thirty students were tested by 26 independent raters to examine the inter-rater reliability of the DRA2. Multiple raters rated each student to determine the degree of agreement between raters. Gwet's kappa coefficient was used as a measure of inter-rate reliability. Kappa values from .41 to .60 indicate a moderate level of agreement, and values above .60 indicate substantial agreement. Fluency received a kappa coefficient of .75 and the kappa value for comprehension was .67.

**Validity.** The construct validity of an assessment refers to the extent to which it measures what it is supposed to measure. The DRA2 was designed to measure oral fluency and reading comprehension. The correlations between the fluency and reading comprehension was .41, which indicates the items in each subtest measure unique aspects of reading. The correlations of the two subtests to the entire assessment were .78 and .89, respectively. To examine the DRA2's concurrent validity, DRA2 scores were correlated with scores from other well-known, previously validated tests of reading

---

<sup>15</sup> See DRA2 K-8 Technical Manual (<https://docplayer.net/14668574-K-8-technical-manual.html>)

comprehension and fluency. Specifically, the relationship between the DRA2 and Gray’s Oral Reading Test-4th Edition (GORT-4; Weiderholt & Bryant, 2001), the DIBELS Oral Reading Fluency Test-6th Edition (DORF; Good, Kaminski, & Dill, 2002), was examined. Table 6 shows the results concurrent validity tests:

**Table 6. Correlation Between DRA2 and Other Reading Tests (Grades 1-3)**

DRA2 Measure	GORT-4 (Comp)	GORT-4 (Fluency)	DORF (Median Score)
Comprehension	.60	.65	.70
Fluency	.62	.69	.74

Source: DRA2 K-8 Technical Manual (<https://docplayer.net/14668574-K-8-technical-manual.html>)

## Findings

### Pre-K Literacy Achievement

The ELQA was administered at BOY and EOY in the 2017-18 and 2018-19 school years to determine the degree to which pre-K students possessed basic early literacy skills prior to the implementation of *World of Wonders* and the extent to which those skills had improved at the end of the school year. The average age of students at the time of the BOY assessment was 4.5 years; at the time of the EOY assessment the average age was 5.3 years. The sample size of pre-K students with data in 2018-19 was far smaller than in 2017-18 (95 versus 306 students with BOY and EOY data), limiting potential Year 2 analyses (see Figure 5 later in the report). Given that at least 742 students were identified in kindergarten as having participated in district pre-K in 2017-18, pre-K assessment data were available for only 41% of these pre-K students.<sup>16</sup>

### Literacy Achievement Pre- versus Post-*World of Wonders* Implementation

How did literacy achievement vary when comparing the ELQA results of pre-K students prior to the implementation of *World of Wonders* (i.e., 2016-17), after one year of implementation (i.e., 2017-18), and after two years of implementation (i.e., 2018-19)? In addition to student-level ELQA data for 2017-18 and 2018-19, the district provided researchers with district level mean BOY and EOY ELQA scores for 2016-17.

Figure 1 displays the mean raw scores of pre-K students on each ELQA measure at BOY and EOY from each of the three school years: pre-*World of Wonders* implementation (2016-17), *World of Wonders* Implementation Year 1 (2017-18), and

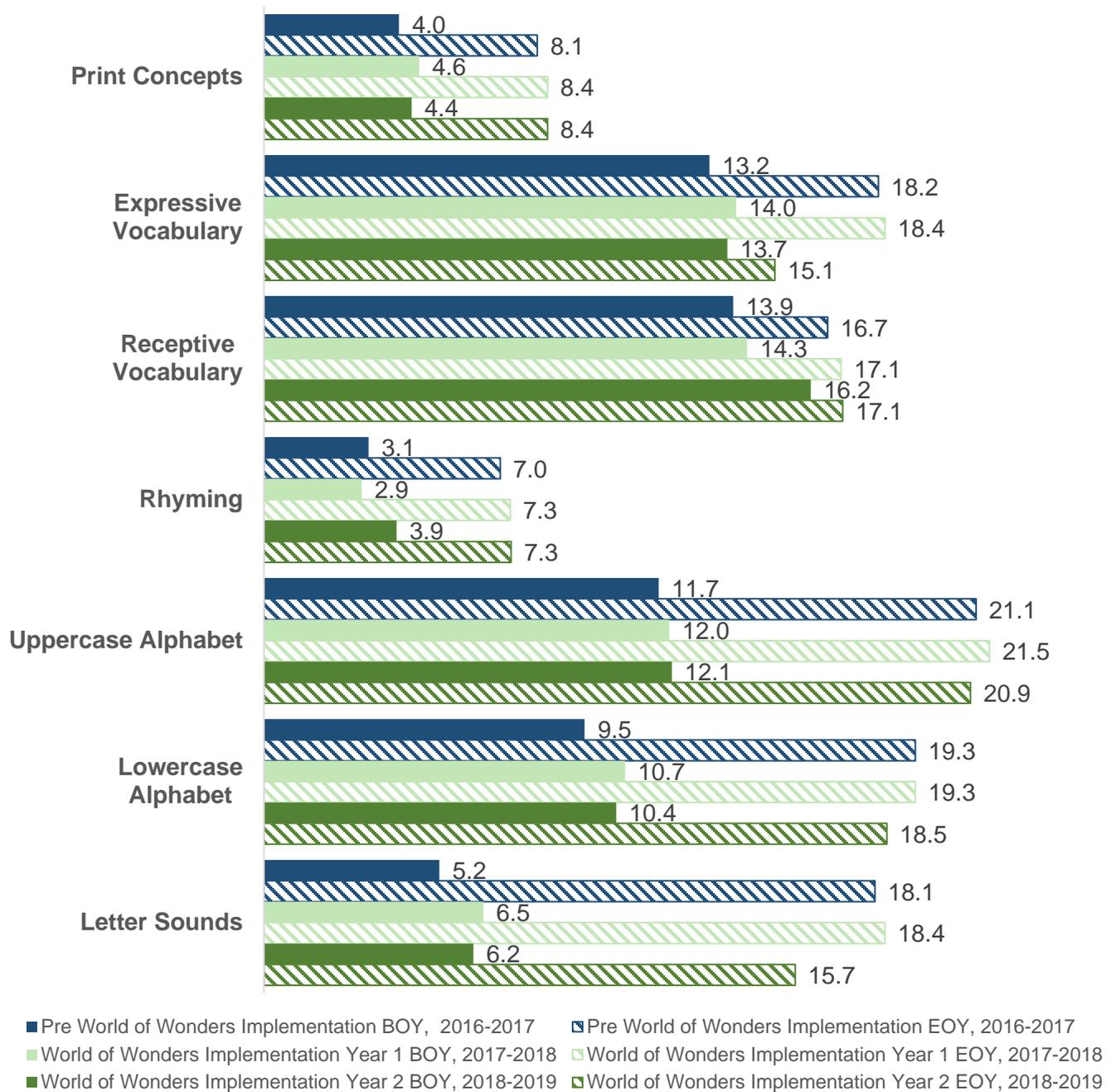
<sup>16</sup> Student demographics were not available for pre-K students, so it was not possible to determine the extent to which the pre-K sample of students with BOY and EOY ELQA data were representative of all preschool students. ELQA data were provided only for those students whose parents consented for the data to be shared. Some pre-K students may have left or entered the district pre-K mid-year leading to missing data. Finally, some students attending district pre-K may have left the district prior to kindergarten. It was not possible to connect pre-K data to Kindergarten data.

*World of Wonders* Implementation Year 2 (2018-19). The BOY scores in each year reflect a student’s starting point prior to experiencing the district pre-K curriculum (*World of Wonders*). Students improved from BOY to EOY on all early literacy measures; however, the nature of the data did not permit researchers to determine whether the levels of improvement from BOY to EOY were statistically significant.<sup>17</sup> In general, the trends from BOY to EOY across the three years can be described as similar. The only exceptions to this were for Expressive Vocabulary and letter sounds in Year 2 of implementation, where EOY scores were lower than EOY scores for the other two samples. Lower Expressive Vocabulary EOY scores in Year 2 are likely related to the fact that only 20 items were used to assess Expressive Vocabulary in this school years, whereas 25 items were used in the first years. It is less clear why letter sounds at EOY were lower in Year 2 but may be associated with the smaller sample being different in some way from the larger sample in Year 1.

---

<sup>17</sup> The data provided for the year prior to *World of Wonders* implementation (2016-17) included only mean scores on each scale at the district level. No student-level or standard deviation data were provided for the study.

**Figure 1. Pre-Kindergarten Mean Beginning- and End-of-Year ELQA Scores Prior to *World of Wonders* Implementation and in *World of Wonders* Implementation Years 1 and 2**



**Source:** *World of Wonders* study data files 2016-17, 2017-18 and 2018-19

**Notes:** BOY = Beginning-of-Year, EOY = End-of-Year. Pre-implementation data based on district level summary data; Implementation Year 1 based on data from 306 students; Implementation Year 2 based on data from 95 students.

Table 7 presents findings from analysis of BOY to EOY ELQA scores in 2017-18 and 2018-19. Paired samples t-tests found that **there were statistically significant improvements from BOY to EOY on each of the ELQA measures within both years of the study.** When **EOY scores were adjusted for BOY score differences, there was no statistically significant change in literacy achievement on any of the ELQA measures from the first year of program implementation to the second.** That is, implementing for a second year did not provide an advantage over having implemented for only 1 year.<sup>18</sup> It was not possible to compare achievement in Expressive Vocabulary across the years of the study because the developer of the ELQA changed the scale from 25 points in the 2017-18 to 20 points 2018-19 on this measure

**Table 7. Average Improvement on ELQA Measures Within and Between School Years**

ELQA Measure	School Year	BOY Mean	EOY Mean	Mean Difference	df	Cohen's d Effect Size
Print Concepts	2017-18	4.6	8.4	3.8*	283	1.7**
	2018-19	4.4	8.4	4.0*	93	1.9**
Expressive Vocabulary	2017-18	14.0	18.4	4.4*	295	1.0
	2018-19	13.7	15.1	1.4*	94	0.4
Receptive Vocabulary	2017-18	14.3	17.1	2.8*	297	1.0
	2018-19	16.2	17.1	0.9*	93	0.4
Rhyming	2017-18	2.9	7.3	4.4*	281	1.5**
	2018-19	3.9	7.3	3.4*	85	1.1
Uppercase Alphabet	2017-18	12.0	21.5	9.5*	298	1.1
	2018-19	12.1	20.9	8.8*	92	1.0
Lowercase Alphabet	2017-18	10.7	19.3	8.6*	224	1.0
	2018-19	10.4	18.5	8.1*	81	1.0
Letter Sounds	2017-18	6.5	18.4	11.9*	204	1.5**
	2018-19	6.2	15.7	9.5*	78	1.1

**Source:** *World of Wonders* study school data files, 2017-18 and 2018-19

**Notes.** df = degrees of freedom; Std. Dev. = Standard Deviation; Cohen's d = mean difference/pooled standard deviation

\* indicates difference is statistically significant at the  $p < 0.001$  level (2-tailed); \*\* indicates effect size of 1.5 or larger

Effect sizes were calculated to express the magnitude of the effect of the *World of Wonders* on early literacy achievement. Effect sizes ranged from 0.4 for Receptive Vocabulary and Expressive Vocabulary in 2018-19, to a high of 1.9 for Print Concepts in 2018-19. A 2007 meta-analysis found that the average annual gain in effect size from

<sup>18</sup> The extent of teacher turnover is unknown. While, it is likely that many of the Year 1 implementation teachers remained in the district in Year 2 some teachers in Year 2 may have been implementing *World of Wonders* for the first time.

nationally normed tests for kindergarten students was 1.5.<sup>19</sup> That is, significant differences at this level or higher are meaningful. District pre-K students matched or exceeded this standard for Print Concepts in in both years; Letter Sounds in Year 1, but not Year 2; and Rhyming in Year 1, but not Year 2.

### To what extent did pre-K students meet target expectations?

Table 8 shows the number of items in each ELQA measure and their corresponding target scores at BOY and EOY. Target scores represent the number of items 4-year old students are expected to answer correctly at BOY and EOY.<sup>20</sup> This table also shows the average number of items participating students answered correctly on each measure. When the average score is equal to or greater than the target score for a given time point, students are considered on average as being on track.

**Table 8. Relative Performance on ELQA Measures (2017-18 and 2018-19)**

ELQA Measure	Number of Items Per Measure	Beginning of Year Performance		End of Year Performance	
		Target Score	Actual Mean Score	Target Score	Actual Mean Score
Print Concepts 2017-18	10	4	4.6*	9	8.4
Print Concepts 2018-19	10	4	4.4*	9	8.4
Expressive Vocabulary 2017-18	25 <sup>a</sup>	14	14.0*	22	18.4
Expressive Vocabulary 2018-19	20 <sup>a</sup>	8	13.7*	14	15.1*
Receptive Vocabulary 2017-18	20	10	14.3*	17	17.1*
Receptive Vocabulary 2018-19	20	10	16.2*	17	17.1*
Rhyming 2017-18	10	3	2.9	8	7.3
Rhyming 2018-19	10	3	3.9*	8	7.3
Uppercase Alphabet 2017-18	26	8	12.0*	20	21.5*
Uppercase Alphabet 2018-19	26	8	12.1*	20	20.9*
Lowercase Alphabet 2017-18	26	6	10.7*	19	19.3*
Lowercase Alphabet 2017-18	26	6	10.4*	19	18.5
Letter Sounds 2017-18	26	N/A	6.5	N/A	18.4
Letter Sounds 2018-19	26	N/A	6.2	N/A	15.7

**Source:** *World of Wonders* study school data files, 2017-18 and 2018-19

**Notes.** There are no target scores for the Letter Sounds Subscale.

\* Indicates met or exceed target score.

<sup>a</sup>Target scores for Expressive Vocabulary in 2017-18 differed from target scores in 2018-19 because the subscale changed from 25 items to 20 items. The most recent target scores are given for all other subscales.

In 2017-18, at BOY, on average students met or exceeded expectations for 5 of the 6 ELQA measures with target scores; the one exception was Rhyming. **At EOY 2017-18, on average students met the target for 3 of the 6 measures: Receptive**

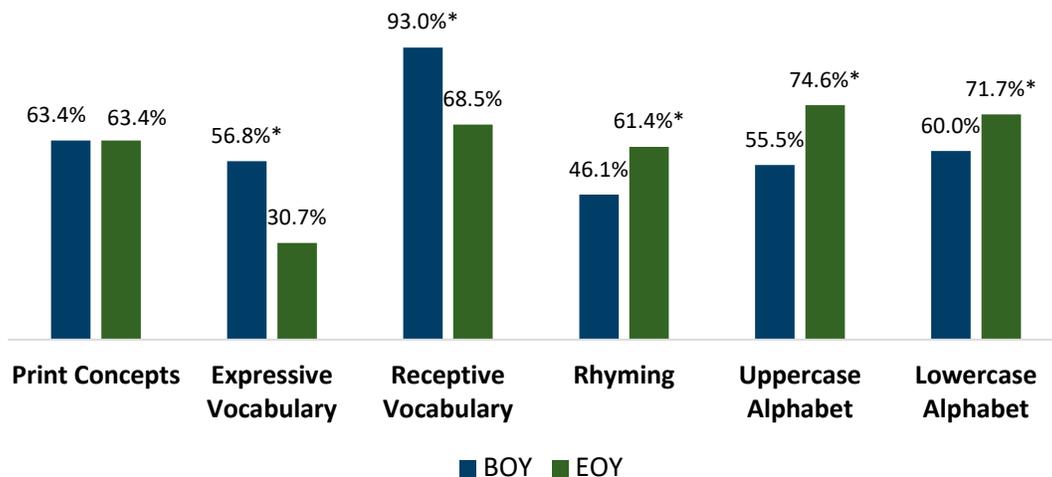
<sup>19</sup> See [https://www.mdrc.org/sites/default/files/full\\_84.pdf](https://www.mdrc.org/sites/default/files/full_84.pdf)

<sup>20</sup> See ELQA Teacher's Guide (2015): <http://eteam.ou.edu/file.axd?file=/ELQA/teachersguide.pdf>

**Vocabulary, Uppercase Alphabet, and Lowercase Alphabet.** In 2018-19, at BOY, on average students were on target for all six of the ELQA measures with targets. **At EOY 2018-19 students met the target for 3 of the 6 measures: Expressive Vocabulary, Receptive Vocabulary, and Uppercase Alphabet.**

Figures 2 and 3 display the percentage of students who met the ELQA early literacy targets in 2017-18 and 2018-19 at BOY and EOY, respectively. Results of McNemar's chi-square test indicate **that there was significant improvement on three of the ELQA measures in 2017-18: Rhyming ( $\chi^2(1) = 16.96, p < .001$ ), Uppercase Letters ( $\chi^2(1) = 47.09, p < .001$ ), and Lowercase Letters ( $\chi^2(1) = 14.70, p < .001$ ), but only Uppercase Letters significantly improved in 2018-19: ( $\chi^2(1) = 8.89, p < .001$ ).** There was no significant change in the percentage of students achieving targets on Print Concepts from BOY to EOY in either year. There were also no significant changes in the percentage of students achieving targets on Rhyming or Lowercase Letters from BOY to EOY in 2018-19. **The percentage of students meeting the targets for Expressive and Receptive Vocabulary *declined* significantly from BOY to EOY in both 2017-18 and 2018-19.** For these two subscales, while there were significant increases from BOY to EOY in both years, this did not translate to meeting target scores at EOY. There are no target scores for the Letter Sounds ELQA measure.

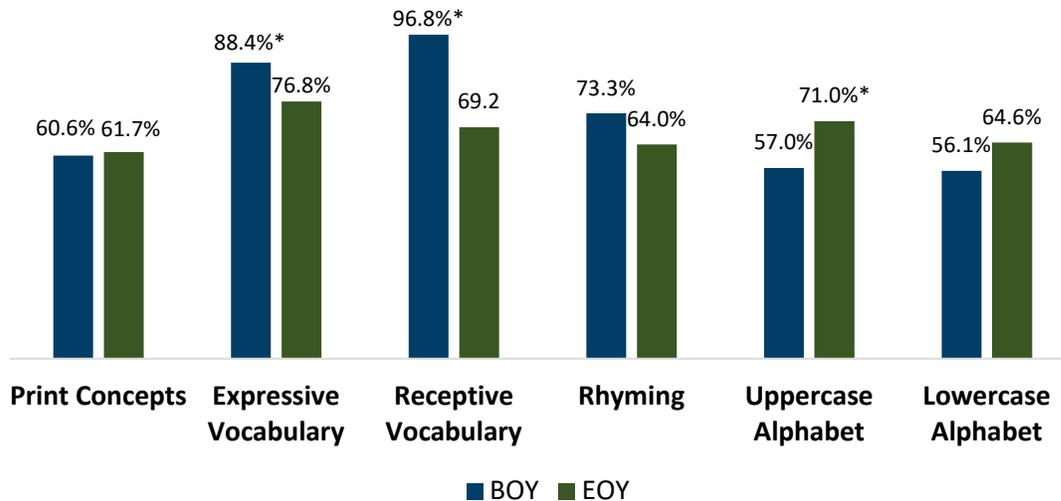
**Figure 2. Chi-Square Analyses Comparing Percentages of Students Meeting ELQA Targets at BOY and EOY 2017-18**



**Source:** World of Wonders study school data files, 2017-18

**Notes:** BOY = Beginning-of-Year, EOY = End-of-Year.

**Figure 3. Chi-Square Analyses Comparing Percentages of Students Meeting ELQA Targets at BOY and EOY 2018-19**



**Source:** *World of Wonders* study school data files, 2018-19

**Notes:** BOY = Beginning-of-Year, EOY = End-of-Year.

### Were school characteristics associated with changes in pre-K literacy achievement?

The evaluation team examined the extent to which each school's prior performance on the Grade 3 OSTP ELA, the percentage of the population that is made up of minority students, and the percentage of students eligible for free or reduced price meals were associated with ELQA scores at the beginning and end of the 2017-18 and 2018-19 school years.<sup>21</sup> Of the 306 students in the 2017-18 sample, 214 (70%) attended a school in which 50% or more of the Grade 3 students demonstrated proficiency on the 2017 OSTP ELA test, 55 students (18%) attended a school in which the majority of students were classified as racial/ethnicity minorities, and 171 students (56%) attended a school in which the majority of students were eligible to receive free or reduced price meals. Regarding the 95 students in the 2018-19 sample, 39 (41%) attended a school in which 50% or more of the Grade 3 students demonstrated proficiency on the 2018 OSTP ELA test, 23 students (24%) attended a school in which the majority of students were classified as racial/ethnicity minorities, and 35 students (37%) attended a school in which the majority of students were eligible to receive free or reduced price meals.

<sup>21</sup> Student-level demographics were not available for the students participating in the study.

Ideally, a multilevel model, in which students are nested within schools, would have been used to examine the impact of school characteristics on student ELQA scores. However, due to the low number of schools in this sample, we were unable to nest students within schools. Instead, we coded school-level variables at the student-level. That is, we indicated whether each student attended a school in which at least half the Grade 3 students demonstrated proficiency on the 2017 Grade 3 OSTP ELA test (i.e., majority proficient); at least half the students identified as a racial/ethnic minority (i.e., majority minority); or at least half the students qualified for free or reduced price meals (i.e., majority FRL) and then ran multiple regression analyses. The Intraclass Correlation Coefficient (ICC) was calculated to determine the proportion of variability in ELQA subscale scores that was accounted for by school-level and student-level variables. Table 9 displays the proportion of variability accounted for at each level for the 2017-18 ELQA results. There were too few students per school to make this determination for the 2018-19 ELQA data.

**Table 9. Percent of EOY ELQA Measure Variance Accounted for at the School and Student-levels (2017-18)**

EOY ELQA Measure	School-level Variance	Student-level Variance
Print Concepts	15.3%	84.7%
Letter Sounds	7.8%	92.2%
Lowercase Alphabet <sup>a</sup>	0.0%	100.0%
Uppercase Alphabet <sup>a</sup>	0.0%	100.0%
Rhyming	21.0%	79.0%
Receptive Vocabulary	6.4%	93.6%
Expressive Vocabulary	9.2%	90.8%

**Source:** School District School Reports, 2017–18

**Note:** <sup>a</sup>Because a majority of students scored at the top of the scale at EOY on Uppercase and Lowercase Alphabet, there was little variability between schools in scores on these measures.

Separate multiple regression analyses were conducted for each of the seven ELQA subscales (See Appendix A). All models except Expressive Vocabulary included the following predictors: BOY ELQA scores for that particular scale, whether the student attended a school in which a majority of students were minorities, whether the student attended a school in which the majority of students were eligible to receive free or reduced price meals, whether the student attended a school in which the majority of Grade 3 students demonstrated proficiency on the 2017 Grade 3 OSTP ELA test, and implementation year: 2017-18 or 2018-19. Because the scale for Expressive Vocabulary changed between 2017-18 and 2018-19, a separate regression was conducted for each year. It should be noted that the strength of these relationships was relatively low and that, given the large number of analyses run with relatively small number of significant

findings, it is possible that some of these findings associated with demographics are random or not practically significant. These findings should be considered exploratory given that the variables are at the school-level. The relevant results of those analyses are as follows:

- Evidence suggests that the school students attended might have influenced pre-K achievement on the ELQA, particularly on the Rhyming and Print Concepts measures as is noted by their high intraclass correlations in 2017-18 (see Table 9). However, **BOY performance was, by far, the strongest predictor of EOY performance on ELQA.**
- When BOY ELQA scores were controlled for, being in a majority Grade 3 OSTP ELA proficient school was a significant predictor of higher Lowercase Alphabet EOY scores, but also a significant predictor of *lower* EOY Expressive Vocabulary and Receptive Vocabulary Scores in 2017-18. Generally, the expectation would be that higher proficiency would have been linked to higher scores making the association with lower scores on Expressive and Receptive Vocabulary unexpected. It is unclear if this is an indication that these two ELQA subscales may be problematic, or an indication that growth in Expressive and Receptive Vocabulary may occur early but be unrelated to literacy achievement as assessed in Grade 3.
- After accounting for BOY ELQA scores, being in a majority minority school was a significant predictor of higher EOY Expressive Vocabulary and Letter Sounds score 2017-18.
- After accounting for BOY ELQA scores, being in a school in which the majority of students qualify for free and reduced meals was a significant predictor of lower EOY Print Concepts and Expressive Vocabulary 2017-18 scores.
- After accounting for BOY ELQA scores, being in implementation year 2 (2018-19) was a significant predictor of lower EOY Receptive Vocabulary scores.

## **Kindergarten Literacy Achievement**

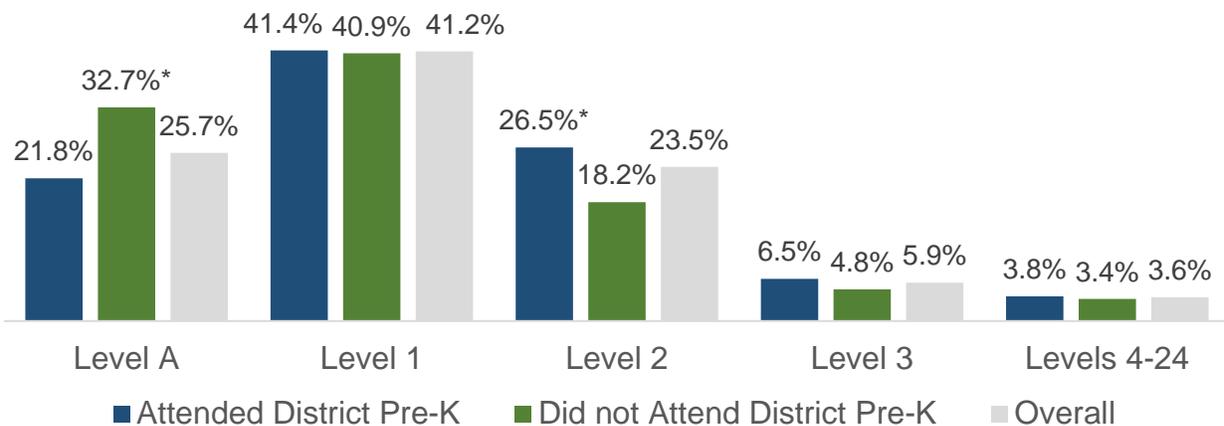
### **Comparison of Students Based on *World of Wonders* Exposure**

The district provided deidentified BOY DRA2 scores for all 1,317 Kindergarten students in 2018-19. This allowed the question of how literacy achievement of kindergarten students who were exposed to *World of Wonders* in pre-K compare to that of kindergarten students who had not been exposed to the program. Of those students, 162 had insufficient data to clearly determine their pre-K status and were excluded from the analysis. In total, **742 attended district pre-K** in the previous school year that used the *World of Wonders* curriculum, with 396 attending a **full day** district pre-K program and 346 students attended a **half day** district pre-K program (morning=167,

afternoon=179. The remaining **413 did not attend a district pre-K program** the previous school year.

Figure 4 provides the BOY DRA2 scale scores (scale ranges from A-80) for students who were exposed to *World of Wonders* curriculum through the district’s pre-K program and those who were not exposed to the program. Overall, about 26% of the sample scored at Level A at BOY while 71% scored between levels 1 and 3 on the kindergarten BOY DRA2 assessment, classifying them as Emerging Readers. The other 4% scored between 4 and 24, classifying them as Early/Transitional Readers. As a reminder, being Level A (or pre-Level A) or higher at BOY and being Level 1 or higher middle-of-year is an independent (if still emerging) reader (see Table 5 earlier in the report). Within district pre-K students 22% scored at Level A and 74% scored at Levels 1 through 3. Regarding students who did not attend district pre-K, 33% scored at Level A and 64% scored at Levels 1 through 3.

**Figure 4. DRA2 Scores for Kindergarten Students who did and did not attend District Pre-K**



**Source:** *World of Wonders* study school data files, 2018-19

**Note:** \* indicates significance at  $p < .05$

**A significantly lower proportion of students who attended the district pre-K program scored at the lowest level, A, in the Emerging Reader category, than students who attended the district pre-K program ( $\chi^2(1) = 16.36, p < .001$ ). A significantly higher proportion of students who attended the district pre-K program scored at Level 2 in the Emerging Reader Category ( $\chi^2(1) = 10.37, p < .01$ ).** These findings suggest that having attended pre-K in the district where *World of Wonders* was the curriculum was associated with fewer students performing at the lowest level of emerging readers, with an increase in the percentage of students at the higher Level 2. While Levels A through 3 are all indicative of emerging readers for Kindergarten BOY, higher levels are associated with higher literacy achievement. That

is, the higher the starting level, the more likely it is that the teacher would have been able to build on this early start.

To allow for further analysis of DRA2 scores, scores of 'A' were converted to zeros so that all the scores would approximate a continuous scale. This allowed for the calculation of mean scores. Outliers that were more than three standard deviations above the sample mean were removed (i.e., above Level 6). This removed four students from the group of students that attended district pre-K and two from the group of students that did not attend district pre-K. **Students who were exposed to *World of Wonders* in pre-K scored significantly higher on the DRA ( $t=3.83$ ,  $p<.001$ ) than students who presumably were not exposed to the program (based on having not attended District pre-K).** The average DRA2 score for the students who were exposed to *World of Wonders* in pre-K was 1.29. The average DRA score for the students who were not exposed to the program was 1.04.<sup>22</sup> **Those results represented a small but positive program effect ( $d=.24$ ).**

### Were student demographics associated with DRA2 literacy achievement?

To examine the impact of the available student-level demographics of race, ELL, and Special Education status, in addition to district pre-K attendance, these variables were included in a logistic regression and predicted the likelihood of pre-K and non-pre-K attending students achieving at each level, while controlling for demographic variables. **Students not attending district pre-K were 1.9 times more likely to be at Level A than students attending district pre-K (adjusted Odds Ratio=1.9,  $p<.001$ ), while students attending district pre-K were 1.7 times more likely to be at Level 2 than students not attending pre-K (adjusted Odds Ratio=1.7,  $p<.001$ ),** after controlling for these demographic variables. That is, the differences found prior to including student demographics were retained after inclusion of these variables. Differences at the other levels remained non-significant while controlling for demographic variables. Similarly, after controlling for demographic characteristics, the adjusted average scores for students attending district pre-K (adj. mean=1.06), remained significantly higher ( $F=17.19$ ,  $p<.001$ ) than students not attending district pre-K (adj. mean=0.8).

### Was full-day versus half-day district pre-K associated with DRA2 literacy achievement?

Comparisons (with outliers removed) were also made between students who attended full-day pre-K ( $n=394$ ), half-day pre-K ( $n=344$ ), and no district pre-K ( $n=409$ ). **Students that attended both full-day ( $M=1.33$ ,  $SD=1.03$ ) and half-day district pre-K ( $M=1.24$ ,**

---

<sup>22</sup> Standard deviations in the two groups were similar: 1.03 among those who attended pre-K and 1.04 among those who did not attend pre-K.

**SD=1.03) both scored significantly higher on the DRA2 ( $F=7.93$ ,  $p<.001$ ) than students who did not attend district pre-K ( $M=1.04$ ,  $SD=1.04$ ).** However, students who attended full-day and half-day district pre-K did not differ significantly from each other on the DRA2 ( $t=1.09$ ,  $p=.27$ ). This pattern held when we controlled for demographic characteristics. Both full and half day pre-K students scored significantly higher than students that did not attend district pre-K ( $F=17.19$ ,  $p<.001$ ), but did not differ significantly from each other after controlling for demographic characteristics.

These findings suggest that there was not an additional advantage with regard to literacy achievement from having attended full-day as compared to half-day district pre-K but that attending district pre-K was advantageous for students with regard to arriving at kindergarten at higher levels of being an emerging reader.

## ***Conclusions and Recommendations***

This study **provides evidence for ESSA Tier III promising evidence, with qualifications**, that *World of Wonders* supports the early literacy achievement of students. **There were statistically significant improvements from BOY to EOY on each of the ELQA literacy subscales within both years of the study. When students entered kindergarten in the district, those students who had attended district pre-K implementing *World of Wonders* scored higher on average on the BOY DRA2, even after account for student demographics.** The finding held true regardless of whether students attended full-day or half-day district pre-K. The kindergarten findings represent a posttest only design given the lack of link to pre-K literacy assessments for all students and included a large enough kindergarten sample size in both groups of students (attending versus not attending district pre-K utilizing *World of Wonders* curriculum), and the two groups were demographically similar with the exception of *more* students identified as special needs having attended district pre-K as compared to not attended district pre-K. This finding suggests that at least during the time when pre-K students were exposed to the *World of Wonders* curriculum, those students were more likely to enter kindergarten at a higher level of becoming a reader than non-attending peers.

The qualifications to identifying as Tier III are primarily associated with two key limitations (discussed further in the limitations section). Notably, it was not possible to differentiate attending pre-K from exposure to *World of Wonders* as all pre-K students in the district were taught using *World of Wonders*. It was also unknown to what extent families who made the decision to enroll a child in district pre-K may have differed from those who did not chose to enroll in district pre-K on variables other than the available demographics. That is, factors outside of *World of Wonders* that cannot be accounted for in this correlational model may have also contributed to the differences between students who did versus did not attend district pre-K.

The two pre-K samples were much smaller in size than the kindergarten sample. The pre-K sample in implementation Year 1 (306 students) was large enough to likely be representative, although student-level demographics were unknown. In implementation Year 2, sample size was much smaller (95 students) decreasing the likelihood that this sample was representative, although findings across the two pre-K implementation years were generally consistent. In both implementation years, students on average met EOY targets for Receptive Vocabulary and Uppercase Alphabet, which suggests *World of Wonders* supports these skills. Findings on 2 of the 6 ELQA measures were mixed with students meeting EOY targets for Expressive Vocabulary in Year 2 but not Year 1 and meeting EOY targets for lowercase alphabet in Year 1 but not year 2. **Targets were not met on average for Print Concepts or Rhyming in either year.** The fact that targets were not met in these areas in either year may indicate that the developers of *World of Wonders* may need to determine if the curriculum adequately supports these skills or if they need to provide additional supports to teachers in implementing the curriculum regarding these skills. It is also possible that the ELQA EOY targets for these skills may not be appropriate. For example, the effect size for change in Print Concepts from BOY to EOY were quite large suggesting meaningful change although the average student did not meet targets at EOY.

While there were significant increases in the percentage of students meeting targets from BOY to EOY (Rhyming, Uppercase Letters, and Lowercase Letters in 2017-18 and Uppercase Letters in 2018-19), there was also a negative finding. Specifically, the percentage of students meeting the EOY, as compared to meeting BOY, targets for Expressive and Receptive Vocabulary declined significantly in both 2017-18 and 2018-19. To be clear, both Expressive and Receptive Vocabulary scores grew from BOY to EOY and on average students were on target for these skill in at least one year if not both; however, some students who were on target for these literacy skills at BOY were no longer on target by EOY. Ideally, to meet Tier III as Promising Evidence there would not be any negative findings in the study. Given that BOY in Kindergarten for students who had attended pre-K were in line with DRA2 targets, concerns around these negative findings are minimal.

Other findings of interest are that:

- The school within the district students attended might have influenced pre-K performance on the ELQA, particularly on the Rhyming and Print Concepts. Also, being in a majority proficient school was a significant predictor of higher Lowercase Alphabet EOY scores, and a significant predictor of lower EOY Expressive Vocabulary in 2017-18, and Receptive Vocabulary Scores.
- Being in a majority minority school was a significant predictor of higher EOY Expressive Vocabulary and Letter Sounds scores in 2017-18.

- Being in a school in which the majority of students qualify for free and reduced meals was a significant predictor of lower EOY Print Concepts and Expressive Vocabulary scores in 2017-18.
- Being in implementation year 2018-19 was a significant predictor of lower EOY Receptive Vocabulary scores.

In light of these findings it is recommend that future studies of *World of Wonders* include a comparison group for all students who were exposed to the program, preferably from the same or similar schools to account for the possible impact of school characteristics. Future studies would also be strengthened by including student-level demographic characteristics (i.e., race, gender, FRL status, ELL status, etc.) for participating students at each grade level. It would also be helpful to include a sufficient number of students and schools so that multilevel analyses can be conducted with students nested within schools. A study of this nature would provide stronger evidence of the program's efficacy.

### **Limitations**

Given the study relied on data provided by the school district, it is worth noting key limitations the readers should keep in mind to best understand report findings. A key limitation of this study is that there was not data available for a comparison group for the pre-K students. Comparison groups are a common feature of evaluation studies because they make it easier to attribute observed changes in the outcome of interest to exposure to the program of interest. Similarly, comparison groups make it easier to rule out the influence of other plausible factors or the possibility that changes in the outcome might have occurred anyway in the absence of the program in question. That is, it is not possible to determine if it was the *Wonders* curriculum per se or the overall preschool experience that were associated with outcomes.

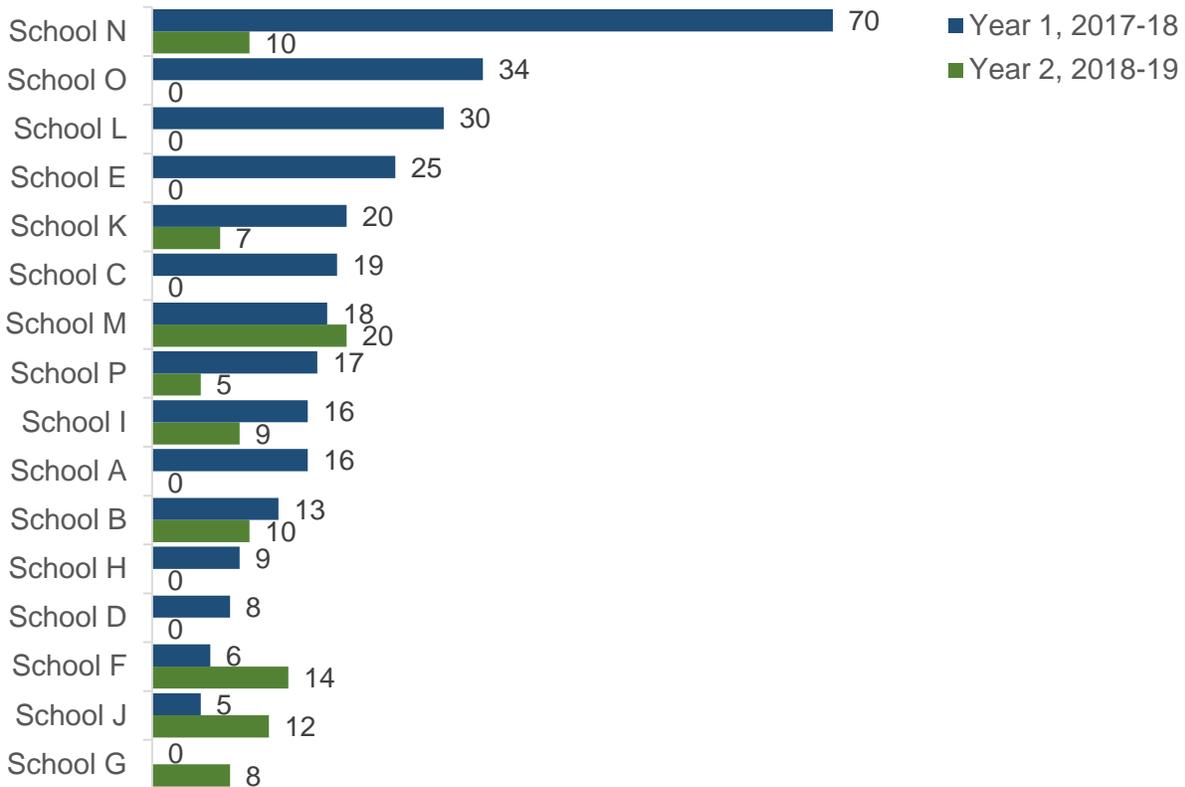
In kindergarten there was a convenience comparison sample of students who had not attended district pre-K. It is unknown if these comparison Kindergarten students may have attended pre-K outside of the district or in another community prior to moving to the district in 2018-19. Demographically, the two groups of kindergarten students were similar to one another suggesting that attendance at district pre-K was not related to child's race, gender, or ELL status. Students attending district pre-K were more likely to have been identified as special needs than those not attending. While demographics are similar, any additional reasons that parents made the decision to not attend district pre-K in 2017-18 are unknown. Parents may have had a child in a day-care setting and not wanted to move the child to a new school, may have perceived that the child was best kept at home prior to kindergarten, or made the decision based on other reasons that are unknown for the study.

In kindergarten, the district was able to provide data for a large sample of students in the district regarding BOY literacy achievement, likely the majority of kindergarten students in the district with BOY DRA2 scores. The sample size was sufficiently large and included student-level demographics, which is a strength of the study. Ideally, EOY kindergarten data would have provided additional evidence (or not) for any advantages associated with attending district pre-K where *World of Wonders* was implemented. Another challenge with the Kindergarten data was that it was not possible to link pre-K performance (based on ELQA) to BOY DRA2 performance. This contributed to the challenge that it was not possible to separate attending preschool effects from exposure to *World of Wonders* effects.

The pre-K samples were more of a challenge with both years of pre-K data much smaller than the number of students attending pre-K. A limitation was that the sample size of pre-K students for whom the district submitted both BOY and EOY ELQA scores in 2017-18 (306) was significantly larger than the 2018-19 sample size of pre-K students (95). Figure 5 shows the number of students at each school in the district for whom the ELQA measures were administered at both time points. In 2017-18, researchers received BOY and EOY ELQA results for 306 students. Nearly one-fourth (23%) of those students attended School N. The school with the next largest group of participants was School O with 34 students. Conversely, four schools—School H (N=9), School D (N=8), School F (N=6), and School I (N=5)—each submitted ELQA results for fewer than 10 students in 2017-18. School G did not submit ELQA scores in 2017-18. For the 2018-19 sample, researchers received BOY and EOY ELQA results for only 95 students, about a third of the number of students from the previous year. Just over a fifth (21%) of the students in 2018-19 attended School M. The school with the next largest group of participants was School F with 14 students, followed by School B (N=10), School N (N=10), School G (N=8), School K (N=7), and School P (N=5). Seven schools did not provide data for 2018-19.

The school district noted that fewer parents had provided permission to share data in 2018-19 than in the prior year. It is unknown if this was because fewer parents were asked to consent or if consent was asked for but not granted. In addition, student-level demographic data were not provided in association with the pre-K ELQA scores. It is unknown if the sample of students in either year is representative of all students in the school. That is, it was also not possible to compare students for whom data were received to those from whom data were not received to determine whether some subgroups of students were particularly likely or not to have data. It is likely that the sample size in 2017-18 was sufficient in size to be representative but given the very small sample size in 2018-19 it may well be that this sample was not representative.

**Figure 5. Number of Students with ELQA Scores at Both Beginning- and End-of-Year by Implementation Year and School**



**Source:** *World of Wonders* study school data files, 2017-18 and 2018-19

The 2018-19 pre-K sample was too small to permit some of the anticipated analyses. The type of analyses conducted on the pre-K sample was also limited by the fact that the evaluation team was not able to collect student-level demographic data, thereby, making it difficult to estimate the extent which each student’s race, gender, special education status, English language proficiency, or family economic student might have influenced their literacy achievement. Finally, because the developer of the ELQA changed the Expressive Vocabulary scale from 25 points in the 2017-18 to 20 points in 2018-19, it was not possible to compare improvement across years on this measure.

## Appendix A: Multiple Regression Analyses

The results of the multiple regression analyses for each ELQA measure are as follows:

- *Print Concepts EOY (n=378)*: The overall model was significant ( $F=22.58$ ,  $p<.001$ ) with an R-square=.23. Significant predictors included:
  - **Print Concepts BOY** ( $F=102.52$ ,  $p<.001$ ) with an estimate of .33. For every 1 point increase of Print Concepts BOY, Print Concepts EOY increased by .33 points, when all other variables in the model are held constant.
  - **Majority FRL** ( $F=12.15$ ,  $p<.05$ ) with an estimate of -.46 indicating that a student in a majority FRL school scores .46 points lower on Print Concepts EOY than a student not in a majority FRL school, when all other variables in the model are held constant.
  
- *Expressive Vocabulary EOY 2017-18 (n=296)*: The overall model was significant ( $F=93.05$ ,  $p<.001$ ) with an R-square=.56. Significant predictors included:
  - **Expressive Vocabulary BOY** ( $F=357.63$ ,  $p<.001$ ) with an estimate of .69. For every 1 point increase of Expressive Vocabulary BOY, Expressive Vocabulary EOY increased by .69, when all other variables in the model are held constant.
  - **Majority Proficient on Grade 3 OSTP ELA** ( $F=9.86$ ,  $p<.01$ ) with an estimate of -1.46 indicating that a student in a majority proficient school scores 1.46 points lower on Expressive Vocabulary EOY than a student not in a majority proficient school, when all other variables in the model are held constant.
  - **Majority Minority** ( $F=7.98$ ,  $p<.01$ ) with an estimate of 1.43 indicating that a student in a majority minority school scores 1.43 points higher on Expressive Vocabulary EOY than a student not in a majority minority school, when all other variables in the model are held constant.
  - **Majority FRL** ( $F=7.83$ ,  $p<.01$ ) with an estimate of -1.27 indicating that a student in a majority FRL school scores 1.27 points lower on Expressive Vocabulary EOY than a student not in a majority FRL school, when all other variables in the model are held constant.
  
- *Expressive Vocabulary EOY 2018-19 (n=95)*: The overall model was significant ( $F=29.14$ ,  $p<.001$ ) with an R-square=.56. Significant predictors included:
  - **Expressive Vocabulary BOY** ( $F=102.97$ ,  $p<.001$ ) with an estimate of .62. For every 1 point increase of Expressive Vocabulary BOY, Expressive

Vocabulary EOY increased by .62, when all other variables in the model are held constant.

- *Receptive Vocabulary EOY (n=392)*: The overall model was significant ( $F=40.35$ ,  $p<.001$ ) with an R-square=.34. Significant predictors included:
  - **Receptive Vocabulary BOY** ( $F=194.31$ ,  $p<.001$ ) with an estimate of .44. For every 1 point increase of Receptive Vocabulary BOY, Receptive Vocabulary EOY increased by .44 points, when all other variables in the model are held constant.
  - **Implementation Year** ( $F=16.43$ ,  $p<.001$ ) with an estimate of -1.04, indicating that a student in 2018-19 scores 1.04 points lower on Receptive Vocabulary EOY than a student in 2017-18, when all other variables in the model are held constant.
  - **Majority Proficient on Grade 3 OSTP ELA** ( $F=7.01$ ,  $p<.01$ ) with an estimate of -.64, indicating that a student in a majority proficient school scores .64 points lower on Receptive Vocabulary EOY than a student not in a majority proficient school, when all other variables in the model are held constant.
  
- *Rhyming EOY (n=368)*: The overall model was significant ( $F=13.08$ ,  $p<.001$ ) with an R-square=.15.
  - **Rhyming BOY** is the only significant predictor ( $F=63.08$ ,  $p<.001$ ) with an estimate of .42. For every 1 point increase of Rhyming BOY, Rhyming EOY increases by .42 points, when all other variables in the model are held constant.
  
- *Uppercase Alphabet EOY (n=392)*: The overall model was significant ( $F=54.07$ ,  $p<.001$ ) with an R-square=.41.
  - **Uppercase Alphabet BOY** is the only significant predictor ( $F=266.96$ ,  $p<.001$ ) with an estimate of .45. For every 1 point increase of Uppercase Alphabet BOY, Uppercase Alphabet EOY increased by .45, when all other variables in the model are held constant.
  
- *Lowercase Alphabet EOY (n=307)*: The overall model was significant ( $F=51.64$ ,  $p<.001$ ) with an R-square=.46. Significant predictors include:
  - **Lowercase Alphabet BOY** ( $F=229.44$ ,  $p<.001$ ) with an estimate of .52 suggesting that for every 1 point increase of Lowercase Alphabet BOY, Lowercase Alphabet EOY increased by .52 points, when all other variables in the model are held constant.

- **Majority Proficient on Grade 3 OSTP ELA** ( $F=4.23$ ,  $p<.05$ ) with an estimate of 1.48 suggests that a student in school in which the majority of Grade 3 students demonstrated proficiency on the ELA test scores 1.38 points higher on Lowercase Alphabet EOY than a student not in a majority proficient school, when all other variables in the model are held constant.
- *Letter Sounds EOY* ( $n=284$ ): The overall model was significant ( $F=27.58$ ,  $p<.001$ ) with an R-square=.33.
  - **Letter Sounds BOY** ( $F=107.41$ ,  $p<.001$ ) with an estimate of .55 indicates that for every 1 point increase of Letter Sounds BOY, Letter Sounds EOY increased by .55 points, when all other variables in the model are held constant.
  - **Majority Minority** ( $F=7.23$ ,  $p<.01$ ) with an estimate of 2.88 indicating that a student in a majority minority school scores 2.88 points higher on Letter Sounds EOY than a student not in a majority minority school, when all other variables in the model are held constant.