Literacy Footprints Guided Reading System

Effects on Grade 3 English Language Arts Assessment Performance

Prepared for Pioneer Valley Books by McREL

November 2022

McREL INTERNATIONAL

curiosity for better learning

McREL project team

Study Authors:
Bradley Rentz, Ph.D.

Robert Merth

Darienne Dey

Project Co-Leads:
Lisa M. Jones, Ph.D.

Sheila A. Arens, Ph.D.

Contributors:
Faith Connelly

Paul Burkander

McREL
INTERNATIONAL

© McREL International

This study evaluates the efficacy of a commercial product, Literacy Footprints, developed by Pioneer Valley Books. The authors of this paper are employed by McREL International, a private 501(c)(3) nonprofit corporation specializing in research and evaluation services, which was contracted by Pioneer Valley Books to design and carry out the study. None of the researchers receive commission on sales of the products.

McREL International is a nonprofit, nonpartisan research, development, and service organization that helps schools, districts, and education agencies improve outcomes for all students. We provide:

- analysis and alignment of curriculum, instruction, assessment, and standards.

- professional development, learning, and coaching for teachers, principals, and system leaders.

- research, data analysis, logic models, and program evaluations.

- school improvement recommendations, action planning, implementation support, and technical assistance.

Executive Summary

Pioneer Valley Books (PVB) contracted McREL International to investigate the effects of the third-grade Literacy Footprints (LFP) guided reading system on English language arts (ELA) performance on state assessments. LFP is a research-based, comprehensive, small-group literacy tool that facilitates explicit and systematic guided reading instruction to students in grades K–6.

McREL researchers conducted a matched pair cluster randomized control trial to investigate LFP's possible effects on student achievement. The trial, conducted during the 2021/22 school year, included a sample of 3,071 students in 29 public elementary schools in a large suburban Florida school district. The district purchased LFP kits for grade 3 teachers in 14 elementary schools (SY2021/22) that were randomly assigned to the treatment group. Treatment teachers also obtained access to the *Digital Reader* and received virtual training and follow-up support to implement materials with students. Business-as-usual (BAU) instruction occurred at 15 schools (the control group). After the trial was completed, control group schools received the kits in the 2022/23 school year with access to the *Digital Reader*.[1]

> The study found a statistically significant difference in the average English language arts standardized test scores between treatment and control groups.

Using a design-based estimator that accounts for both school-level clustering and school-matched pairs, the study found a statistically significant difference in the average English language arts standardized tests scores between treatment and control groups. By the end of the study, the treatment group students on average earned a 4.31-point higher scale score than the control group students. Results showed that, despite implementation challenges and limitations, students who received the treatment outperformed their control group peers on the Florida Statewide Assessment in English language arts. However, although statistically significant, the effect size (Hedges' g = 0.19) is small. This study showed promising evidence for the LFP guided reading system. Further research is needed to examine the generalizability of these results across different school districts and contexts and under more favorable conditions (e.g., without frequent COVID disruptions).

Background

Children who struggle with reading during elementary grades are more likely to struggle in school during later grades (Mather et al., 2001). Moreover, students from low-income backgrounds have consistently scored lower on achievement measures, including reading, than high- or middle-income children (Lacour & Tissington, 2011). In response to this achievement gap, the federal government passed and implemented the No Child Left Behind (NCLB) Act in 2001 and then updated the legislation with the Every Student Succeeds Act (ESSA) in 2015. These landmark education policies have significantly altered instruction in schools, particularly in ELA and mathematics. Specifically, ESSA largely restricts the use of federal monies to those school practices, curricula, and materials that meet one of four evidence tiers (strong evidence, moderate evidence, promising evidence, and strong rationale). As a result, school district leaders need to be much more circumspect in selecting

---

[1] Ongoing challenges due to the COVID-19 pandemic (e.g., frequent teacher and student absences, limited professional development time) prevented consistent, systematic implementation of the kits across treatment teachers and schools.

school practices, curricula, and materials—prioritizing those supported by robust research evidence. Developers are, therefore, seeking to provide school district leaders with robust evidence of the efficacy of their tools. The Literacy Footprints Guided Reading System—a reader that provides students with access to readings that are appropriately tailored to their reading level and supported by high-quality instruction in a small group format—is grounded in research literature, but which lacked empirical evidence of effectiveness. The product is informed by an extant body of literature, which is summarized below.

High-Quality Literacy Instruction

Research consistently shows a direct relationship between teacher's knowledge and skills, effective literacy instruction, and intended student outcomes (Darling-Hammond, 2000; McCombes-Tolis & Feinn, 2008; Piasta et al., 2009). Effective literacy instruction techniques occur before, during, and after students read a text (Shanahan, 2010). The LFP guided reading system incorporates seven literacy instruction techniques: comprehension monitoring, guided reading instruction, writing instruction, leveled reading, systematic phonics instruction, appropriately challenging texts, and small and flexible groups. The seven techniques are described below.
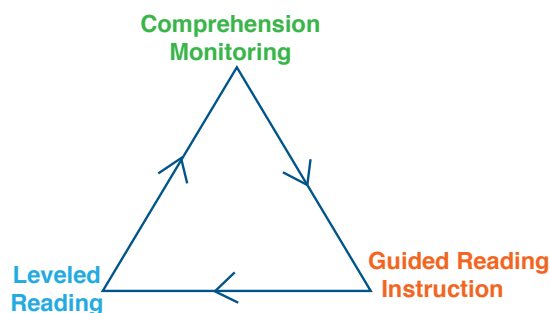
Comprehension Monitoring

Research shows comprehension monitoring, a strategy whereby a teacher provides immediate feedback to a student while they read, is effective in improving student reading comprehension (National Reading Panel, 2000). Furthermore, teaching students the meanings of words while they read a passage can support their reading comprehension (Wright & Cervetti, 2017). Likewise, improving students' knowledge about topics they are reading about can also help improve reading comprehension (Cabell & Hwang, 2020; Cervetti et al., 2016; Connor et al., 2017).

Guided Reading Instruction

Guided reading is an early reading instructional approach for primary grades that is widely implemented across the U.S. (Fountas & Pinnell, 2012). This approach is designed to promote independent reading through the use of small groups of readers with similar reading and comprehension skills before, during, and after reading (Young, 2018). Guided reading can support the development of students' word reading, especially among students with reading difficulties (Denton et al., 2014), and it is associated with increased phonemic awareness (Tobin & Calhoon, 2009) and improved reading comprehension over typical school instruction (Denton et al., 2014).

Writing Instruction

Recent research has come to regard reading and writing as reciprocal processes. In a meta-analysis of students in grades 1–12, Graham and Herbert (2011) found that writing interventions improved reading outcomes, and in a separate meta-analysis of preschool to grade 12 students, Graham and Santangelo (2014) found that spelling instruction improved reading comprehension scores. With respect to the effect of reading on writing, additional meta-analyses have found that reading improves spelling (Graham, 2000; Share, 1995). Finally, preschool through grade 12 literacy programs that balanced both reading and writing simultaneously improved both outcomes (Graham et al., 2018).

Leveled Reading

Not all students begin each school year at their expected literacy level. A strategy to accelerate literacy development is to teach within students' Zone of Proximal Development (ZPD) (Vygotsky, 1978). The ZPD refers to an area of learning where a student can almost complete a task independently but needs assistance from a teacher or instructor to complete the task successfully.

One strategy that educators use to ensure that instruction occurs in the ZPD is through small group, leveled reading practices. Leveled reading is the practice of matching the difficulty of a text to the reading level, or just slightly above the reading level, of the reader. Teacher support during leveled reading activities in the ZPD can be effective in improving reading comprehension (Paris & Paris, 2007). Furthermore, differentiating instruction based on the students' reading levels may further improve reading comprehension (Allington et al., 2015; Ankrum, 2022; Halladay, 2012; Pearson, 2007).

Systematic Phonics Instruction

Systematic phonics instruction (i.e., sequenced phonics instruction; Shanahan, 2005) teaches students that different letters have different sounds, which are combined to form words familiar to the child. This instructional approach systematically uses a sequential set of phonics elements and stands in contrast to incidental phonics instruction that occurs haphazardly and infrequently (National Reading Panel, 2000). The extensiveness of scholarly literature concerning the efficacy of systematic phonics instruction has eclipsed the arguments favoring the efficacy of whole language instruction (i.e., instruction that focuses on the meaning of words instead of letter-sound correspondences) for teaching students to read (see Kim, 2008, and Pearson, 2004, for reviews). Several meta-analyses demonstrate the efficacy of systematic phonics instruction (Adesope et al., 2011; Camilli et al., 2003, 2006; Ehri et al., 2001; Galuschka et al., 2014; Han, 2010; McArthur et al., 2012; National Reading Panel, 2000; Sherman, 2007; Suggate, 2010, 2016; Torgerson et al., 2006, inter alia). Although still contested, the majority of meta-analytic studies confirm systematic phonics instruction as more effective than whole language teaching for students' reading development.

Appropriately Challenging Texts

Not only are teacher practices instrumental in student reading comprehension, increased exposure to texts can also improve comprehension (Mol & Bus, 2011). Students who had access to texts that are numerous, engaging, and available at a range of difficulty levels have demonstrated higher reading comprehension levels than those without it (Hoffman et al., 2004). Moreover, variation in the amount of exposure to different texts can predict students' reading comprehension beyond their current reading level. Sparks and colleagues (2014) found that differences in exposure to texts in elementary school predicted differences on standardized measures of reading comprehension (i.e., the Indiana Statewide Testing for Educational Progress [ISTEP] Reading and Language Tests) in grade 10.

Small, Flexible Groups

Students learn best when instruction matches their learning needs (Al Otaiba & Fuchs, 2006; Kamps & Greenwood, 2005), and reading instruction is no different. Small-group reading instruction—a strategy that groups students by reading level and skill—can provide teachers with an efficient and effective way to deliver personalized reading instruction (Gersten et al., 2008).

Small-group reading interventions can be effective for students at risk for reading failure in early elementary (Kamps et al., 2008; Nielsen & Friesen, 2012), upper elementary (Faggella-Luby & Wardwell, 2011), middle school (Fagella-Luby & Wardwell, 2011; Vaughn et al., 2010), and high school (Bemboom & McCaster, 2013). Additionally, a 2018 meta-analysis of reading interventions found that small group reading formats had a moderate effect on elementary school student literacy outcomes (Hall & Burns, 2018).

Problem Statement

Despite research indicating positive relationships between and among instruction, format, access to text, and reading comprehension, many teachers do not employ these strategies in classrooms. For some teachers, it is a lack of training; in one study, preservice teachers reported receiving little training on reading instruction during preservice teacher education (Begeny & Martens, 2006).

Other studies suggest that, at least among elementary teachers, expressed confidence in one's ability to provide reading instruction fails to correspond to the use of evidence-based instructional practices with students (National Reading Panel, 2000; Wijekumar et al., 2019). In some classrooms, inconsistent access to resources such as challenging texts and reading materials is an obstacle—one that varies by ethnicity (Nelson, 2009) and socioeconomic status (Funge et al., 2017; Teravainen-Goff & Clark, 2017; Tuck & Holmes, 2016). The American Association of School Librarians recommended providing electronic access to texts and educational materials aligned with appropriate students' reading levels, indicating that digital access should be continuously available to students (AASL, 2011). Soulen and Tedrow (2021) showed that access to reading materials diminished during the COVID-19 pandemic, which makes this suggestion salient to the current landscape of literacy education.

The Literacy Footprints Guided Reading System

To address the issues described, Pioneer Valley Books (PVB) developed the LFP guided reading system to improve teacher practices, students' access to books, and students' reading achievement. The LFP guided reading system is founded on a research-based, comprehensive, small-group literacy framework that provides explicit and systematic guided reading instruction. Literacy Footprints takes a multipronged approach, addressing a multiplicity of skills to improve student reading. Finally, guided by the academic literature, Literacy Footprints encourages small group instruction, challenging texts, and responsive instructor feedback, which differs from other literacy education materials.

The LFP guided reading system includes both print and *Digital Reader*[2] versions. Each classroom receives at least one kit, and the teachers lead students through guided reading and can assign texts for independent reading. When kindergarten through grade 6 students use the LFP guided reading system, they can access Literacy Footprints grade-appropriate kits (i.e., each grade level has a uniquely designed kit) in print. Each kit comprises six to 11 levels of texts and lessons, depending on the grade level, and includes assigned and independent readings, lesson cards, and pre-recorded lessons for each reading. Each student reading level contains five to 23 texts. For grade 3, which is the focus of this study, there are six levels and 10 or 11 texts for each lesson, for a total of 62 texts. The lessons are based on the Next Step lesson format (Richardson, 2009).

PVB also provides training and professional learning opportunities for teachers to effectively use their products. The LFP guided reading system kits include materials to support teachers in delivering lessons that are easy to follow to ensure high-quality instruction for student users.

---

[2] *The Digital Reader* was developed to facilitate student access during the COVID-19 pandemic. The print version was the primary focus of the study, though schools also had access to the *Digital Reader*.

Study Description and Setting

PVB contracted McREL to examine the effects of their LFP guided reading system on grade 3 English language arts outcomes. To accomplish this goal, McREL designed and conducted a cluster randomized controlled trial during school year (SY) 2021/22 in central Florida's School District of Manatee County (SDMC). Schools were randomly assigned to either the treatment group, which purchased LFP guided reading system kits during SY 2021/22, or the control group, which purchased kits for SY 2022/23.

The study was conducted entirely within public schools in the SDMC (National Center for Education Statistics [NCES] District ID: 1201230). A single school district was used in this study to maintain consistency in the general educational practices and policies of the schools and to ensure consistent outcome measures for all participating students.

SDMC is in a large suburban area and serves approximately 52,000 students, as of 2022. The district consists of 66 schools, including 31 elementary schools and two K–8 schools (SDMC, n.d.). Of the 66 schools, 25 (37.9 percent) are Title 1 schools. In terms of student demographics, 44 percent of students are white, 36 percent are Hispanic or Latino, and 13 percent are Black or African American compared to 35 percent, 36 percent, and 22 percent, respectively, for all grade 3 public school students in Florida. In addition, 19 percent of students are reported as English language learners (ELL) and 54 percent[3] as economically disadvantaged, whereas statewide 10 percent of K–12 students are reported as English language learners and 63 percent as economically disadvantaged.[4]

Intervention and Control Groups

Schools in the treatment group purchased grade 3 LFP guided reading print kits for all enrolled grade 3 students from August 2021 to June 2022 (i.e., the end of the school year) and access to the *Digital Reader*. Teachers in the treatment group schools received training from PVB at the beginning of the school year on how to use the LFP guided reading system kits, and follow-up support was provided throughout the school year.

Because of the COVID-19 pandemic, PVB primarily conducted this teacher training virtually instead of in person (in-person teacher training is standard practice). Teachers were offered access to recorded webinars about how to use the kits. Some schools, however, did request in-person demonstration lessons and coaching support, which was also provided by PVB.

Teachers in the treatment group schools were not required to use the LFP guided reading system kits but were encouraged to incorporate them into their regular reading instruction. To track how often teachers used the kits, teachers were asked to log each lesson they assigned to students in an online platform. Teachers were inconsistent in how often they used the kits and how often they logged their use. PVB provided additional virtual follow-up support to schools with lower implementation rates throughout the school year, per their standard practice.

---

[3] When including all grades, 59 percent of SDMC students are reported as economically disadvantaged.

[4] More information about Florida student demographics may be found on the Florida Department of Education (FLDOE) website (https://www.fldoe.org/accountability/data-sys/edu-info-accountability-services/pk-12-public-school-data-pubs-reports/index.stml)

As previously noted, schools assigned to the control group were delayed in receiving their LFP guided reading system kits until SY 2022/23. During the study period, the control group schools conducted business-as-usual reading instruction for grade 3 students.

During spring 2022, in accordance with normal district practices, both treatment group and control group schools administered the Florida Statewide Assessment (FSA) in English language arts (ELA).

Study Design and Methodology

Research Question

The following research question guides this study:

> *To what extent does using Literacy Footprints guided reading system affect grade 3 English language arts standardized test scores?*

Design

McREL used a matched-pair cluster randomized controlled trial to address the research question. In August 2021, prior to the beginning of the school year, 31 schools from the SDMC (out of all 33 eligible schools in the district [those serving grade 3 either via a K–5 or K–8 program])[5] were matched into pairs based on their Title 1 status, total enrollment, average percentage improvement on grade 3 i-Ready reading diagnostic assessments[6] in spring 2021, and average grade 3 FSA in ELA scale scores in spring 2021. After the schools were paired, one school in each pair was randomly assigned to the treatment group (received the LFP kits) and the other to the control group (did not receive the kits). Since the total number of participating schools was an odd number, one cluster included a single treatment group school matched with two control group schools.

Analytic Sample

Of the original 31 randomly assigned schools, two schools (one matched treatment and comparison school) were removed from the study, which is equivalent to 6.5 percent overall cluster attrition and 0 percent differential cluster attrition. Within the participating schools, 3,133 students (that is, 1,552 students from treatment group schools and 1,581 students from control group schools) were enrolled when schools were randomly assigned. The school that a student was enrolled in at the time of randomization was used in the analysis, regardless of whether the student switched schools during the study period. No joiners to the schools were included in the analytic sample. Of the 3,133 original students, 3,071 (1,518 students from treatment group schools and 1,553 students from control group schools) had available outcome data at the end of the study, which corresponds to 1.9 percent individual-level overall attrition and 0.5 percent differential attrition. Students with missing outcome data were excluded from the analytic sample. Thus, the study's analytic sample included 3,071 grade 3 students enrolled at 29 SDMC schools (14 treatment group schools and 15 control group schools).

---

[5] Two schools opted not to join the study and thus were not among those randomly assigned.

[6] The i-Ready diagnostic assessments were administered as part of the district's standard practice. More information on the assessments can be found here: https://www.curriculumassociates.com/programs/i-ready-assessment/diagnostic

Demographic data for the students in the analytic sample are provided in Table 1. This sample represents 87.9 percent of SDMC schools. There are some statistically significant differences in student demographics between the treatment and control groups.[7] Treatment schools had a lower percentage of Hispanic or Latino, Title 1, and ELL students compared with control schools, but the two groups of schools had comparable percentages of students' reported races and genders.[8]

Table 1. Demographics of Analytic Sample

Sample Characteristics (N=3,071)	Treatment—Literacy Footprints (frequency, percentage) [n=1,518]	Control (frequency, percentage) [n=1,553]	Chi-square p-value
Race			0.49
American Indian or Alaska Native	8 (0.5%)	11 (0.7%)	
Asian	49 (3.2%)	39 (2.5%)	
Black or African American	311 (20.5%)	291 (18.7%)	
Native Hawaiian or Pacific Islander	10 (0.7%)	10 (0.6%)	
White	1231 (81.1%)	1294 (83.3%)	
Ethnicity			<0.001
Hispanic/Latino	490 (32.3%)	613 (39.5%)	
Not Hispanic/ Latino	1028 (67.7%)	940 (60.5%)	
Gender			0.92
Male	782 (51.5%)	804 (51.8%)	
Female	736 (48.5%)	749 (48.2%)	
Title 1	685 (45.1%)	793 (51.1%)	0.0011
ELL	279 (18.4%)	373 (24.0%)	<0.001

Data Collection

SDMC provided data for the study. The outcome measure for the study was the grade 3 FSA in ELA that was administered at all schools in the district in spring 2022 using standard procedures.

The results of the FSA in ELA were provided to the district by the Florida Department of Education (FLDOE).

---

[7] Appendix A provides a sensitivity analysis that examines how including student-level covariates, such as demographics, affects the estimated average treatment effect.

[8] While this sample represents a large percentage of the district's grade 3 students, the results from this sample may not be generalizable to other communities.

Outcome Measures

This study used one outcome measure, student scale scores from the spring 2022 grade 3 FSA in ELA. This assessment is Florida's standardized achievement test and was designed and scored by the FLDOE and administered by the SDMC. The assessment measures student performance in terms of the Florida Standards in English language arts and contains four general categories (Florida Department of Education, 2022):

- Key Ideas and Details
  In this category, students are expected to read closely to comprehend, analyze, and summarize essential information and concepts, referencing evidence from the text to support inferences and conclusions.

- Craft and Structure
  In this category, students are expected to interpret literal and nonliteral meanings of words/phrases, determine how text structures and text features impact meaning, and distinguish personal point of view from that of the narrator or author.

- Integration of Knowledge and Ideas
  In this category, students are expected to integrate and analyze content presented in diverse media formats and analyze treatment of similar themes or topics.

- Language and Editing
  In this category, students are expected to demonstrate command of the conventions of standard English grammar, usage, capitalization, punctuation, and spelling.

An external psychometric validity study of the state assessment in 2015 found it to be consistent with established testing practices and provided limited recommendations for improvement (Wiley et al., 2015). Descriptive statistics for the study sample on the outcome measure are provided in Table 2.

Table 2. Study Sample Descriptive Statistics for the Outcome Measure

Measure	Treatment—Literacy Footprints			
	Number of students	Number of clusters	Unadjusted mean	Unadjusted standard deviation
Grade 3 FSA in ELA scale scores	1,518	14	297.7	22.3
Control (Business as Usual)				
Grade 3 FSA in ELA scale scores	1,553	15	293.5	23.0

Empirical Strategy

To find the average treatment effect of Literacy Footprints on grade 3 FSA in ELA scale scores, student-level data were analyzed using a design-based difference in means estimator using the R language package *estimatr* (Blair et al., 2022). The design-based difference in means estimator accounts for both school-level clustering and school matched-pairs when estimating the differences in means between the treatment and control groups. The estimated difference in means, $\hat{\tau}$, is calculated by the following equation.

$$\hat{\tau} = \sum_{j=1}^{J} \frac{N_j}{N} \hat{\tau}_j$$

J is the number of matched-pairs; N_j is the size of the matched-pairs; N is the total number of units; and $\hat{\tau}_j$ is the estimated difference in means in each matched-pair. The variance for the estimated difference in means, $\hat{\mathbb{V}}[\hat{\tau}]$, is estimated by the following equation.

$$\hat{\mathbb{V}}[\hat{\tau}] = \frac{J}{(J-1)N^2} \sum_{j=1}^{J} \left(N_j \hat{\tau}_j - \frac{N\hat{\tau}}{J} \right)^2$$

Confidence intervals (*CI*) for the estimated difference in means are calculated by the following equation.

$$CI^{1-\alpha} = \hat{\tau} + t^{df}_{\frac{\alpha}{2}} \sqrt{\hat{\mathbb{V}}[\hat{\tau}]} + t^{df}_{1-\frac{\alpha}{2}} \sqrt{\hat{\mathbb{V}}[\hat{\tau}]}$$

In the above equation, α is 0.05, and degrees of freedom (*df*) are defined as *J*−1.

Findings

The results of the design-based difference in means estimator indicate that there is a statistically significant difference in average grade 3 FSA in ELA scale scores between the treatment and comparison groups (Table 3), with treatment students outperforming their control student counterparts. That is, students enrolled at treatment group schools had, on average, a 4.3-point higher grade 3 FSA in ELA scale score in spring 2022 than students enrolled in control group schools. These results are also robust to the observed differences in student demographics between the treatment and control groups, albeit slightly smaller in magnitude when controlling for those student-level differences (see Appendix A). While there is a statistically significant difference, the estimated effect size (Hedges' *g*) was small (0.19).

Table 3. Design-based Difference in Means Estimator Results

Outcome	Estimated Difference in Means	Hedges' g*	Std. Error	t-value	df	p-value	CI Lower	CI Upper
Grade 3 FSA in ELA scale scores	4.31	0.19	0.27	16.21	13	<0.001	3.46	5.16

*The Hedges' *g* value was not calculated using the design-based estimator and does not take into account school-level or matched-pair clustering.

Note: School-level intraclass correlation (ICC) = 0.22 and matched-pair ICC = 0.05.

Conclusions and Discussion

This study results show that LFP guided reading system had a positive, but small, effect on grade 3 ELA scores for SDMC students. The small effect size, however, was anticipated, given the challenges in properly implementing LFP guided reading system in the treatment group schools due to pandemic-related logistical complications. While the true effect of a nearly perfect implementation cannot be estimated, this study provides evidence that, despite being inconsistently implemented across treatment schools in a district, the LFP guided reading system can still improve ELA standardized test scores.

Although the results from this study show promising results for LFP guided reading system, further research is needed to examine the extent to which these results may be generalizable to other school districts and contexts.

Appendix A. Sensitivity Analysis

Researchers conducted a sensitivity analysis in which student-level demographic differences between the treatment and control groups were controlled for. The research team used Bayesian additive regression trees (BART), which generally demonstrate high predictive performance and model non-linear patterns, heteroscedastic data, and complex variable interactions without requiring pre-specification (Hill, 2011). The R package *bartCause* (Dorie & Hill, 2020) was used to calculate the BART model, which included student-level demographics (gender, race, ethnicity, Title 1 status, and ELL status) and random effect grouping by the matched pairs of schools. The model's response surface was calculated using a three-step process:

1. Calculate the response surface using BART.

2. Calculate the treatment effect estimates using a propensity score weighted average.

3. Correct the estimates using the targeted minimum loss-based estimation (TMLE) adjustment.

To ensure common support between students in both the treatment and control group schools, the model was specified to exclude observations using the "chisq" rule.[9] However, no observations were removed using this method. The model finally was run with four chains that each had 4,000 samples and 2,000 burn-in samples.

The results of the BART model indicate that the average treatment effect for LFP guided reading system on grade 3 FSA in ELA scale scores is 2.78 (sd = 0.73, CI = [1.35, 4.22]) after controlling for student-level demographics. The results of this sensitivity analysis support the main finding, which controlled for only school-level demographic differences, that the LFP guided reading system had a positive effect on grade 3 FSA in ELA scale scores for students in the study, though the effect is somewhat smaller than the main finding after controlling for student-level demographics.

---

[9] Any observations with a *p-value* less than 0.05 from a chi-square test of the ratio of the posterior variance of the predicted counterfactual to the posterior variance of the observed condition are removed.

References

Adesope, O. O., Lavin, T., Thompson, T., & Ungerleider, C. (2011). Pedagogical strategies for teaching literacy to ESL *immigrant students: A meta-analysis. British Journal of Educational Psychology, 81*(4), 629–653. https://doi.org/10.1111/j.2044-8279.2010.02015.x

Ahmed, Y., Wagner, R.K., & Lopez, D. (2014). Developmental relations between reading and writing at the word, sentence, and text levels: A latent change score analysis. *Journal of Educational Psychology, 106*(2), 419–434. https://doi.org/10.1037/a0035692

Al Otaiba, S., & Fuchs, D. (2006). Who are the young children for whom best practices in reading are ineffective? *Journal of Learning Disabilities, 39*(5), 414–431. https://doi.org/10.1177/00222194060390050401

Allington, R. L., McCuiston, K., & Billen, M. (2015). What research says about text complexity and learning to read. *The Reading Teacher, 68*(7), 491–501. https://doi.org/10.1002/trtr.1280

American Association of School Librarians (AASL). (2011). *AASL Urban Schools Task Force survey report* [Report]. Urban Schools Taskforce. https://alair.ala.org/handle/11213/15889

Ankrum, J. W. (2022). Complex texts or leveled readers for the primary grades? Yes and yes! *Early Childhood Education Journal, 50*, 605–611. http://doi.org/10.1007/s10643-021-01178-1

Begeny, J. C., & Martens, B. K. (2006). Assessing pre-service teachers' training in empirically-validated behavioral instruction practices. *School Psychology Quarterly, 21*(3), 262–285. https://doi.org/10.1521/scpq.2006.21.3.262

Bemboom, C. M., & McMaster, K. L. (2013). A comparison of lower- and higher-resourced tier 2 reading interventions for high school sophomores. *Learning Disabilities Research and Practice, 28*(4), 184–195. https://doi.org/10.1111/ldrp.12020

Berninger, V.W., & Abbott, R.D. (2010). Listening comprehension, oral expression, reading comprehension, and written expression: Related yet unique language systems in grades 1, 3, 5, and 7. *Journal of Educational Psychology, 102*(3), 635–651. https://doi.org/10.1037/a0019319

Blair, G., Cooper, J., Coppock, A., Humphreys, M., & Sonnet, L. (2022). *Estimatr: Fast Estimators for Design- Based Inference* [R package version 1.0.0]. https://CRAN.R-project.org/package=estimatr

Bowers, J. S. (2020). Reconsidering the evidence that systematic phonics is more effective than alternative methods of reading instruction. *Educational Psychology Review, 32*, 681–705. https://psycnet.apa.org/record/2020-02143-001

Buckingham, J. (2020.) Systematic phonics instruction belongs in evidence-based reading programs: A response to Bowers. *The Educational and Developmental Psychologist, 37*(2), 105–113. https://doi.org/10.1017/edp.2020.12

Cabell, S. Q., & Hwang, H. (2020). Building content knowledge to boost comprehension in the primary grades. *Reading Research Quarterly, 55*(S1), S99–S107. https://doi.org/10.1002/rrq.338

Camilli, G., Vargan, S., & Yurecko, M. (2003). Teaching children to read: the fragile link between science and federal education policy. *Education Policy Analysis Archives, 11*(15), 1–51. https://epaa.asu.edu/index.php/epaa/article/view/243/369

Camilli, G., Wolfe, M., & Smith, M. L. (2006). Meta-analysis and reading policy: Perspectives on teaching children to read. *The Elementary School Journal, 107*(1), 27–36. https://doi.org/10.1086/509525

Castles, A., Rastle, K., & Nation, K. (2018). Ending the reading wars: Reading acquisition from novice to expert. *Psychological Science in the Public Interest, 19*(1), 5–51. https://journals.sagepub.com/doi/10.1177/1529100618772271

Cervetti, G. N., Wright, T. S., & Hwang, H. (2016). Conceptual coherence, comprehension, and vocabulary acquisition: A knowledge effect? *Reading and Writing, 29*(4), 761–779. https://link.springer.com/article/10.1007/s11145-016-9628-x

Clay, M. M. (2005). *Literacy lessons designed for individuals: Part two: Teaching procedures*. Heinemann.

Connor, C. M., Dombek, J., Crowe, E. C., Spencer, M., Tighe, E. L., Coffinger, S., Zargar, E., Wood, T., & Petscher, Y. (2017). Acquiring science and social studies knowledge in kindergarten through fourth grade: Conceptualization, design, implementation, and efficacy testing of content-area literacy instruction (CALI). *Journal of Educational Psychology, 109*(3), 301–320. https://doi.org/10.1037/edu0000128

Darling-Hammond, L. (2000). Teacher quality and student achievement. *Education Policy Analysis Archives, 8*(1), 1–44. https://doi.org/10.14507/epaa.v8n1.2000

Denton, C. A., Fletcher, J. M., Taylor, W. P., Barth, A. E., & Vaughn, S. (2014). An experimental evaluation of guided reading and explicit interventions for primary-grade students at-risk for reading difficulties. *Journal of Research on Educational Effectiveness, 7*(3), 268–293. http://doi.org/10.1080/19345747.2014.906010

Dorie, V., & Hill, J. (2020). *Causal inference using Bayesian additive regression trees (bartCause)* [R package version 1.0-4]. https://CRAN.R-project.org/package=bartCause

Ehri, L. C., Nunes, S. R., Stahl, S. A., & Willows, D. M. (2001). Systematic phonics instruction helps students learn to read: evidence from the National Reading Panel's meta-analysis. *Review of Educational Research, 71*(3), 393–447. https://doi.org/10.3102/00346543071003393

Faggella-Luby, M., & Wardwell, M. (2011). RTI in a middle school: Findings and practical implications of a Tier 2 reading comprehension study. *Learning Disability Quarterly, 34*(1), 35–49. https://doi.org/10.1177/073194871103400103

Fitzgerald, J., & Shanahan, T. (2000). Reading and writing relation and their development. *Educational Psychologist, 35*(1), 39–50. https://doi.org/10.1207/S15326985EP3501_5

Fletcher, J. M., Savage, R., & Vaughn, S. (2021). A commentary on Bowers (2020) and the role of phonics instruction in reading. *Educational Psychology Review, 33*(3), 1249-1274. https://eric.ed.gov/?id=EJ1310133

Florida Department of Education. (2022). *Understanding Florida Statewide Assessment reports 2022.* https://www.fldoe.org/core/fileparse.php/5663/urlt/USFL-SWAssessReports.pdf

Fountas, I. C., & Pinnell, G. S. (2012). Guided reading: The romance and the reality. *The Reading Teacher, 66*(4), 268–284. http://doi.org/10.1002/TRTR.01123

Funge, S. P., Sullivan, D. J., & Tarter, K. (2017). Promoting positive family interactions: Evaluating a free early childhood book distribution program. *Early Childhood Education Journal, 45*(5), 603–611. https://doi.org/10.1007/s10643-016-0815-9

Galuschka, K., Ise, E., Krick, K., & Schulte-Körne, G. (2014). Effectiveness of treatment approaches for children and adolescents with reading disabilities: a meta-analysis of randomized controlled trials. *PLoS One, 9*(2), e89900. https://doi.org/10.1371/journal.pone.0089900

Gersten, R., Compton, D., Connor, C. M., Dimino, J., Santoro, L., Linan-Thompson, S., & Tilly, W. D. (2008). *Assisting students struggling with reading: Response to Intervention (RtI) and multi-tier intervention for reading in the primary grades (NCEE 2009-4045).* National Center for Education Evaluation and Regional Assistance, Institute of Education Sciences, U.S. Department of Education. https://ies.ed.gov/ncee/wwc/docs/practiceguide/rti_reading_pg_021809.pdf

Graham, S. (2000). Should the natural learning approach replace traditional spelling instruction? *Journal of Educational Psychology, 92*(2), 235–247. https://doi.org/10.1037/0022-0663.92.2.235

Graham, S., & Hebert, M. (2011). Writing to read: A meta-analysis of the impact of writing and writing instruction on reading. *Harvard Educational Review, 81*(4), 710–744. https://doi.org/10.17763/haer.81.4.t2k0m13756113566

Graham, S., & Santangelo, T. (2014). Does spelling instruction make students better spellers, readers, and writers? A meta-analytic review. *Reading and Writing, 27*(9), 1703–1743. https://www.researchgate.net/publication/265296738_Does_spelling_instruction_make_students_better_spellers_readers_and_writers_A_meta-analytic_review

Graham, S., Liu, X., Aitken, A., Ng, C., Barlett, B., Harris, K. R., & Holzapfel, J. (2018). Effectiveness of literacy programs balancing reading and writing instruction: A meta-analysis. *Reading Research Quarterly, 53*(3), 729–304. https://www.jstor.org/stable/26622547

Hall, M., & Burns, M. (2018). Meta-analysis of targeted small-group reading interventions. *Journal of School Psychology, 66*, 54–66. https://doi.org/10.1016/j.jsp.2017.11.002

Halladay, J. L. (2012). Revisiting key assumptions of the reading level framework. *The Reading Teacher, 66*(1), 53–62. https://doi.org/10.1002/TRTR.01093

Han, I. (2010). *Evidence-based reading instruction for English language learners in preschool through sixth grades: A meta-analysis of group design studies.* University of Minnesota Digital Conservancy. http://hdl.handle.net/11299/54192

Hill, J. L. (2011). Bayesian nonparametric modeling for causal inference. *Journal of Computational and Graphical Statistics, 20*(1), 217–240. https://doi.org/10.1198/jcgs.2010.08162

Hoffman, J. V., Sailors, M., Duffy, G. R., & Beretvas, S. N. (2004). The effective elementary classroom literacy environment: Examining the validity of the TEX-IN3 Observation System. *Journal of Literacy Research, 36*(3), 303–334. https://doi.org/10.1207/s15548430jlr3603_3

IRA/NICHD Panel on Research on Reading–Writing Connections. (2012). *Report of the IRA/NICHD Panel on Research on Reading–Writing Connections.* International Reading Association; National Institute of Child Health and Human Development. https://lra2012.sched.com/event/160us1I/research-directions-a-report-on-the-iranichd-expert-panel-on-research-on-the-reading-writing-connection

Kamps, D., Abbott, M., Greenwood, C., Wills, H., Veerkamp, M., & Kaufman, J. (2008). Effects of small-group reading instruction and curriculum differences for students most at risk in kindergarten: Two-year results for secondary- and tertiary-level interventions. *Journal of Learning Disabilities, 41*(2), 101–114. https://doi.org/10.1177/0022219407313412

Kamps, D., & Greenwood, C. (2005). Formulating secondary-level reading interventions. *Journal of Learning Disabilities, 38*(6), 500–509. https://doi.org/10.1177/00222194050380060501

Kim, J. S. (2008). Research and the reading wars. In F. M. Hess (Ed.), *When research matters: How scholarship influences education policy* (pp. 89–111). Harvard Education Press.

Lacour, M. & Tissington, L. D. (2011). The effects of poverty on academic achievement. *Educational Research and Reviews, 6*(7), 522–527. https://eric.ed.gov/?id=EJ936666

Langer, J.A., & Applebee, A.N. (1987). *How writing shapes thinking.* National Council of Teachers of English.

Mather, N., Bos, C., & Babur, N. (2001). Perceptions and knowledge of preservice and in-service teachers about early literacy instruction. *Journal of Learning Disabilities, 34*(5), 472–482. https://doi.org/10.1177/002221940103400508

McArthur, G., Eve, P. M., Jones, K., Banales, E., Kohnen, S., Anandakumar, T., Larsen, L., Marinus, E., Wang, H.-C., & Castles, A. (2012). Phonics training for English-speaking poor readers. *The Cochrane Database of Systematic Reviews, 12*, CD009115. https://doi.org/10.1002/14651858.CD009115.pub2

McCombes-Tolis, J., & Feinn, R. (2008). Comparing teachers' literacy-related knowledge to their state's standards for reading. *Reading Psychology, 29*(3), 236–265. https://doi.org/10.1080/02702710801982258

Mol, S. E., & Bus, A. G. (2011). To read or not to read: A meta-analysis of print exposure from infancy to early adulthood. *Psychological Bulletin, 137*(2), 267–296. https://doi.org/10.1037/a0021890

National Reading Panel. (2000). *Report of the National Reading Panel. Teaching children to read: An evidence-based assessment of the scientific research literature on reading and its implications for reading instruction: Reports of the subgroups* (NIH Publication No. 00-4754). https://www.nichd.nih.gov/publications/pubs/nrp/smallbook

Nelson, J. V. (2009). *Leadership through middle school library programming: Impact on Latino student achievement* (Order No. 3349790). [Doctoral dissertation, University of California, San Diego]. ProQuest Dissertations and Theses Global.

Nelson, N., & Calfee, R. (1998). The reading–writing connection. In N. Nelson & R. Calfee (Eds.), *Ninety-seventh yearbook of the National Society for the Study of Education* (Part 2, pp. 1–52). National Society for the Study of Education.

Nielsen, D. C., & Friesen, L. D. (2012). A study of the effectiveness of a small-group intervention on the vocabulary and narrative development of at-risk kindergarten children. *Reading Psychology, 33*(3), 269–299. https://doi.org/10.1080/02702711.2010.508671

Paris, A. H., & Paris, S. G. (2007). Teaching narrative comprehension strategies to first graders. *Cognition and Instruction, 25*(1), 1–2. https://www.tandfonline.com/doi/full/10.1080/07370000709336701

Pearson, P. D. (2004). The reading wars. *Educational Policy, 18*(1), 216–252. https://doi:10.1177/0895904803260041

Pearson, P. D. (2007). An endangered species act for literacy education. *Journal of Literacy Research, 39*(2), 145–162. https://doi.org/10.1080/10862960701331878

Piasta, S. B., Connor, C. M., Fishman, B. J., & Morrison, F. J. (2009). Teachers' knowledge of literacy concepts, classroom practices, and student reading growth. *Scientific Studies of Reading, 13*(3), 224–248. https://doi.org/10.1080/10888430902851364

Richardson, J. (2009). *The next step in guided reading: Focused assessments and targeted lessons for helping every student become a better reader.* Scholastic.

School District of Manatee County (SDMC). (n.d.). *Academic dashboard.* Assessment Accountability and Research. https://www.manateeschools.net/academicdb

Shanahan, T. (2005). *The national reading panel report: Practice advice for teachers.* https://files.eric.ed.gov/fulltext/ED489535.pdf

Shanahan, T. (2006). Relations among oral language, reading, and writing development. In C.A. MacArthur, S. Graham, & J. Fitzgerald (Eds.), *Handbook of writing research* (pp. 171–183). Guilford.

Shanahan, T. (2016). Relationships between reading and writing development. In C.A. MacArthur, S. Graham, & J. Fitzgerald (Eds.), *Handbook of writing research* (2nd ed., pp. 194–207). Guilford.

Shanahan, T., Callison, K., Carriere, C., Duke, N. K., Pearson, P. D., Schatschneider, C., & Torgesen, J. (2010). *Improving reading comprehension in kindergarten through 3rd grade: A practice guide* (NCEE 2010-4038). https://ies.ed.gov/ncee/wwc/practiceguide/14

Share, D.L. (1995). Phonological recoding and self-teaching: Sine qua non of reading acquisition. *Cognition, 55*(2), 151–218. https://doi.org/10.1016/0010-0277(94)00645-2

Sherman, K. H. (2007). *A meta-analysis of interventions for phonemic awareness and phonics instruction for delayed older readers* (Order No. 3285626) [Doctoral dissertation, University of Oregon]. ProQuest Dissertations and Theses Global.

Soulen, R. R., & Tedrow, L. (2021). Students' frequency of access to school library materials in transformative times. *Journal of Librarianship and Information Science, 00*(0), 1–18. https://doi.org/10.1177/09610006211037721

Sparks, R. L., Patton, J., & Murdoch, A. (2014). Early reading success and its relationship to reading achievement and reading volume: Replication of '10 years later.' *Reading and Writing: An Interdisciplinary Journal, 27*(1), 189–211. https://doi.org/10.1007/s11145-013-9439-2

Suggate, S. P. (2010). Why what we teach depends on when: grade and reading intervention modality moderate effect size. *Developmental Psychology, 46*(6), 1556–1579. https://doi.org/10.1037/a0020612

Suggate, S. P. (2016). A meta-analysis of the long-term effects of phonemic awareness, phonics, fluency, and reading comprehension interventions. *Journal of Learning Disabilities, 49*(1), 77–96. https://doi.org/10.1177/0022219414528540

Teravainen-Goff, A., & Clark, C. (2017). *School libraries: A literature review on current provision and evidence of impact.* The National Literacy Trust. https://www.researchgate.net/publication/328539120_School_libraries_A_literature_review_o n_current_provision_and_evidence_of_impact

Tierney, R.J., & Pearson, P.D. (1983). Toward a composing model of reading. *Language Arts, 60*(5), 568–580. https://eric.ed.gov/?id=ED233329

Tierney, R.J., & Shanahan, T. (1991). Research on the reading–writing relationship: Interactions, transactions, and outcomes. In R. Barr, M.L. Kamil, P. Mosenthal, & P.D. Pearson (Eds.), *Handbook of reading research* (Vol. 2, pp. 246–280). Longman.

Tobin, K. G., & Calhoon, M. B. (2009). A comparison of two reading programs on the reading outcomes of first-grade students. *Journal of Direct Instruction, 9*(1), 35–46. https://eric.ed.gov/?id=EJ877260

Torgerson, C. J., Brooks, G., & Hall, J. (2006). *A systematic review of the research literature on the use of phonics in the teaching of reading and spelling* (DfES research rep. 711). Department for Education and skills, University of Sheffield.

Tuck, K. D., & Holmes, D. R. (2016). *Library/media centers in U.S. public schools: Growth, staffing, and resources. Full report.* National Education Association. https://files.eric.ed.gov/fulltext/ED581061.pdf

Vaughn, S., Wanzek, J., Wexler, J., Barth, A., Cirino, P., Fletcher, J., Romain, M., Denton, C., Roberts, G., & Francis, D. (2010). The relative effects of group size on reading progress of older students with reading difficulties. *Reading and Writing, 23*, 931–956. https://doi.org/10.1007/s11145-009-9183-9

Vygotsky, L. S. (1978). *Mind in society: The development of higher psychological processes*. Harvard University Press.

Wijekumar, K., Beerwinkle, A. L., Harris, K. R., & Graham, S. (2019). Etiology of teacher knowledge and instructional skills for literacy at the upper elementary grades. *Annals of Dyslexia, 69*(1), 5–20. https://doi.org/10.1007/s11881-018-00170-6

Wiley, A., Hembry, T. R., Buckendahl, C. W., Forte, E., Towles, E., & Nebelsick-Gullett, L. (2015). *Independent Verification of the Psychometric Validity for the Florida Standards Assessment*. https://www.fldoe.org/core/fileparse.php/5306/urlt/FSA-Final-Report_08312015.pdf

Wright, T. S., & Cervetti, G. N. (2017). A systematic review of the research on vocabulary instruction that impacts text comprehension. *Reading Research Quarterly, 52*(2), 203–226. https://doi.org/10.1002/rrq.163

Wyse, D. & Bradbury, A. (2022). Reading wars or reading reconciliation? A critical examination of robust research evidence, curriculum policy and teachers' practices for teaching phonics and reading. *Review of Education, 10*(1), e3314. https://doi.org/10.1002/rev3.3314

Young, C. (2018). Increased frequency and planning: A more effective approach to guided reading in grade 2. *The Journal of Educational Research, 112*(1), 121–130. https://doi.org/10.1080/00220671.2018.1451814