ANNENBERG
BROWN UNIVERSITY

EdWorkingPaper No. 24-893

A Scalable Approach to High-Impact Tutoring for Young Readers: Results of a Randomized Controlled Trial

Kalena E. Cortes
Texas A&M University

Karen Kortecamp
George Washington University

Susanna Loeb
Stanford University

Carly D. Robinson
Stanford University

This paper presents the results from a randomized controlled trial of Chapter One, an early elementary reading tutoring program that embeds part-time tutors into the classroom to provide short bursts of 1:1 instruction. Eligible kindergarten students were randomly assigned to receive supplementary tutoring during the 2021-22 school year (N=818). The study occurred in a large Southeastern district serving predominantly Black and Hispanic students. Students assigned to the program were over two times more likely to reach the program's target reading level by the end of kindergarten (70% vs. 32%). The results were largely homogenous across student populations and extended to district-administered assessments. These findings provide promising evidence of an affordable and sustainable approach for delivering personalized reading tutoring at scale.

VERSION: January 2024

Suggested citation: Cortes, Kalena E., Karen Kortecamp, Susanna Loeb, and Carly D. Robinson. (2024). A Scalable Approach to High-Impact Tutoring for Young Readers: Results of a Randomized Controlled Trial. (EdWorkingPaper: 24-893). Retrieved from Annenberg Institute at Brown University: https://doi.org/10.26300/pbn6-s898

A scalable approach to high-impact tutoring for young readers: Results of a randomized controlled trial

Kalena E. Cortes*
Texas A&M University

Karen Kortecamp
George Washington University

Susanna Loeb
Stanford University

Carly D. Robinson
Stanford University

Abstract

This paper presents the results from a randomized controlled trial of Chapter One, an early elementary reading tutoring program that embeds part-time tutors into the classroom to provide short bursts of 1:1 instruction. Eligible kindergarten students were randomly assigned to receive supplementary tutoring during the 2021-22 school year (N=818). The study occurred in a large Southeastern district serving predominantly Black and Hispanic students. Students assigned to the program were over two times more likely to reach the program's target reading level by the end of kindergarten (70% vs. 32%). The results were largely homogenous across student populations and extended to district-administered assessments. These findings provide promising evidence of an affordable and sustainable approach for delivering personalized reading tutoring at scale.

* Cortes: The Bush School of Government and Public Service, 4220 TAMU, 2088 Allen Building, Texas A&M University, College Station, TX 77843 (e-mail: kcortes@tamu.edu). Kortecamp: Graduate School of Education and Human Development, The George Washington University, 2136 G Street NW, Washington, DC 20052 (e-mail: karenkor@gwu.edu). Loeb: Stanford Graduate School of Education, 482 Galvez Mall, Stanford, CA 94305 (e-mail: sloeb@stanford.edu). Robinson: Stanford Graduate School of Education, 482 Galvez Mall, Stanford, CA 94305 (e-mail: carly_robinson@stanford.edu). We thank Melissa Gentry, Danila Serra, Sara White, seminar and conference participants at Texas A&M University, the Association for Education Finance and Policy, the American Educational Research Association for helpful feedback. Institutional support was provided by Brown University's Annenberg Institute for School Reform, Texas A&M University, and Stanford University.

I. INTRODUCTION

A primary goal of early elementary education is developing literacy skills (Fiester, 2010), yet two-thirds of U.S. students will not be proficient readers by the time they reach 4th grade (U.S. Department of Education, 2022). The COVID-19 pandemic rallied the nation around the need to accelerate student learning and, in particular, to address pandemic-induced "lost learning." Research provides unusually strong evidence pointing to the promise of specific practices to advance learning for struggling young readers, including intensive individualized instruction from an adult who knows the child well and can inspire and motivate them to learn (Herrera et al., 2021; Neitzel et al., 2022). However, children across the country are not receiving the quality or quantity of support they need to meet grade-level benchmarks. In part, this failing may stem from the difficulty of implementing some of the most promising practices within schools as they are currently structured, and as a result, questions remain about the ability to use these research-based practices at the scale needed to address the breadth of need in early reading. Scaling promising practices in the short-run likely requires them fit within the routines of early elementary school and at a cost that is viable within school budgets.

In this article, we present results from a randomized controlled trial of an early elementary reading tutoring program designed to be feasible at scale. During the 2021-22 school year, we randomly assigned over eight hundred kindergarten students in the Broward County Public School District in Florida to receive, or not receive, supplementary tutoring from an early literacy program, called Chapter One. Overall, Chapter One served approximately 10,000 students in Broward County that year, within and outside of the randomized controlled trial (RCT). The Chapter One program embeds part-time tutors into the classroom to provide short bursts of instruction to individual students each week over the course of the school year. With the support

of technology, tutors deliver a sequenced curriculum to students. The consistent presence of the tutors in the classroom is designed to allow them to build strong relationships with students and meet students' individual needs at the moment they might most benefit from personalized instruction.

The results of the study show that students who participated in Chapter One's program were over twice as likely to reach the target reading stage by the end of kindergarten (a 120% increase). While this outcome was pre-identified as the main outcome for this study, because the assessors were not blind to treatment status, we focus our results on other assessments with less potential for bias and find meaningful positive effects on those outcomes as well. The results at the end of the first year of implementation provide promising evidence of an affordable and sustainable approach for delivering one-on-one personalized reading tutoring at scale. We also supplement the quantitative analysis of the first year of the intervention with insights from a concurrent qualitative investigation. The findings from the interviews conducted with district and tutoring provider staff provide insights into the mechanisms behind the success of the program.

II. BACKGROUND

A. The Promise and Challenges of High-Impact Tutoring for Early Literacy

Even prior to the pandemic, millions of students across the country were not learning to read through classroom instruction alone (Lesnick et al., 2010). Decades of evidence point to an effective intervention to help struggling readers: one-on-one or small group tutoring (Neitzel et al., 2022). Research consistently demonstrates that tutoring interventions have substantial positive effects on student learning—often translating to an additional 3-15 months of schooling (Nickow et al., 2020). The evidence base for early elementary tutoring in reading is particularly strong,

although the effectiveness of individual programs can vary greatly (Heinrich et al., 2014; Nickow et al., 2020; Wanzek et al., 2016).

The documented variations in tutoring-program effectiveness may be, in part, due to the wide range of interventions that people refer to as tutoring. While some tutoring takes the form of homework help and drop-in support (Robinson et al., 2022), reading tutoring interventions that provide students with one-on-one, personalized reading instruction over an extended period of time consistently demonstrate the largest improvements in reading achievement (Cavanaugh et al., 2004; Gersten et al., 2020; Neitzel et al., 2022; Slavin et al., 2011; Wanzek et al., 2016, 2018). Specific programs may differ in delivery or approaches, but most effective reading tutoring programs involve students meeting for 20-60 minute sessions several times a week with a consistent educator who uses evidence-based reading curricula (Galuschka et al., 2014; Wanzek et al., 2016, 2018). These features align with the definition of "high-impact" tutoring, which involves substantial time each week spent in required tutoring; sustained and strong relationships between students and their tutors; close monitoring of student knowledge and skills; alignment with school curriculum; and oversight of tutors to assure quality interactions (Robinson & Loeb, 2021).

High-impact tutoring programs drive the large effect sizes cited in the literature, but they can be hard to scale and require substantial resources to implement (Groom-Thomas et al., 2023). Successful tutoring programs often require dedicated tutoring blocks within the school schedule and cost over $1000 per student (Guryan et al., 2023; Sirinides et al., 2018). Given the large expected effect sizes, high-impact tutoring still can be quite cost-effective for improving student learning outcomes (Guryan et al., 2023). However, the urgent and growing demand for high-impact tutoring programs to build children's reading skills (U.S. Office of the Press Secretary,

2022) and common implementation issues (Carbonari et al., 2022; Groom-Thomas et al., 2023), may prompt district leaders to search for lower-cost programs that fit within existing school routines.

The early grades stand out as an opportunity to provide the benefits of high-impact tutoring – the close relationships that engage and motivate students with the high-quality instruction aimed at each student's particular needs and strengths – at a lower cost. Younger students have shorter attention spans, so they may benefit from the high-impact tutoring approach for shorter session times, which could reduce the cost if the costs of starting sessions, ending sessions and transitioning tutors to different students is not a large portion of the tutoring program time.

B. Early Literacy Policy Solutions

Students' early literacy skills at the start of kindergarten strongly predict their later reading proficiency in third grade (Herring et al., 2022). Students who traditionally lack educational opportunities, such as due to structural inequalities based on race or socioeconomic status, are less likely to be proficient readers by third grade, even when they start kindergarten with the same literacy skills (Herring et al., 2022). Ensuring kindergarteners make adequate gains in literacy during the kindergarten year may have an outsized impact on their future achievement.

As of 2021, 41 states and the District of Columbia had adopted early literacy policies to improve student literacy by the end of third grade (Cunningham & Stanovich, 1997; ExcelinEd, 2022; Sparks et al., 2014). Research provides evidence that state-level early literacy policies can positively affect student performance on third grade high-stakes assessments and reduce gaps in achievement between groups of students (Westall & Cummings, 2022). States that had comprehensive early literacy policies, which included having a plan for delivering interventions to students in-need, showed gains even on low-stakes tests (Westall & Cummings, 2022). Tutoring

has been a popular intervention choice, especially since the onset of the pandemic (see Hubbard et al., 2022)

C. Leveraging Close Relationships and Technology to Provide Short Bursts of Instruction

Although tutoring is not a novel approach to improving literacy, the evaluation of the Chapter One program is among the first to provide evidence that early elementary students can benefit from frequent, short bursts of reading instruction from consistent tutors embedded in the classroom. The program leverages technology and the close relationship tutors build with their students to personalize instruction, dosage, and session length to meet the individual needs of each child to develop a strong foundation in phonics and build reading fluency.

Many early literacy interventions deploy classroom teachers to deliver literacy instruction, which often involves carving out additional time for instruction for struggling readers leaving others students without adult support (Herrera et al., 2021; Nickow et al., 2020). Chapter One uses a "push-in" model that provides districts with part-time tutors, or Early Literacy Interventionists (ELIs), who meet with students one-on-one in the back of the classroom or in an adjacent room over the course of a school year. One ELI can serve multiple classrooms in the school and provides tutoring to individual students in five to ten minute sessions during blocks of reading instruction or other opportune moments. At the end of each session with a student, the departing student brings the next student to the ELI to minimize interruptions of classroom instruction.

These short sessions leverage young students' short attention spans and allow for each session to focus on a progression of discrete skills (Ehri et al., 2001). Students move through stages of phonics development, learning to segment and blend short and long vowel sounds, learn sight words, and learn strategies to fluently read both decodable and noncontrolled texts. The curriculum

draws on a strong evidence base on teaching young children to read (Ehri et al., 2001) and is designed to match learning and instruction with a child's developmental level (Vygotsky, 1978).

The program not only adjusts the focus of instruction to each students' current understandings, but also adjusts the length of each session and the number of sessions per week for each student based on need and rate of progress. Students who are making adequate progress may only meet with their tutor once or twice a week, whereas students who the tutors identify as in need of more support may meet daily.

To provide this tailored approach, the Chapter One program uses technology to support instruction, as well as to direct student independent practice. ELIs follow a digital curriculum to conduct each session, which facilitates the assessment and tracking of student performance over time. In addition to using the technology in one-on-one sessions, students are scheduled to spend 15 minutes each day independently practicing using Chapter One's software on program-provided tablets. All assessments sync in real time with individual student tablets, so that when a student uses the practice software after the one-on-one session, they practice items that are precisely aligned to their most recent tutored instruction. ELIs also regularly meet with teachers, reading coaches, and principals to review online reports of student progress.

D. Labor and Program Costs

The structured curriculum and technological support allow for a wide range of people to serve as ELIs. Some ELIs are former classroom teachers, however most do not have a teaching certification. All ELIs have earned at least a Bachelor's degree and undergo an extensive series of online training courses with associated assessments that they must pass to proceed in the training plan. ELIs are compensated substantially above minimum wage and also receive ongoing support and development over the term of their employment.

The program costs school districts approximately $375 per student, which includes the ELI, student technology (tablets - Kindle Fires), background check, training time, Chromebook for the ELI, reinforcement materials for the ELI vetted to align with the model, and indirect costs for implementing the program. In implementations that involve over 5,000 students, Chapter One asks the district to fund the cost of district-wide managers which increases the cost per student to approximately $450. This cost per student is far below typical costs for most intensive tutoring programs. For example, in Massachusetts, the cost of early literacy tutoring services during the same academic year ranged from $925 to $1,909 per student (Hubbard et al., 2022). Even in large implementations, Chapter One's cost is substantially lower than the vast majority of other tutoring programs and does not require districts to coordinate complicated logistical arrangements.

E. The Present Study

Our evaluation explores the effect of receiving Chapter One tutoring in kindergarten and first grade on reading proficiency through early elementary school. In this article we present the results from year one of the study, in which we assess the intermediate impact of Chapter One tutoring on kindergarten students' reading development. Specifically, we ask whether students receiving Chapter One tutoring in kindergarten reach the program's targeted Reading Foundation Stage (stage 4) at the end of kindergarten. We also estimate the impact of the program on other assessments and explore how the intervention differentially affected students based on their initial characteristics. We will continue to assess student progress through the end of third grade to measure the long-term impact of the intervention.

The implementation of early literacy supports often determines how successful they are (Herrera et al., 2021), and tutoring programs are no exception (Groom-Thomas et al., 2023). We

draw on insights from qualitative interviews with educators in the district to contextualize the enabling features of Chapter One's program design.

III. METHODS

A. Study Details

During the 2021-22 school year, Chapter One partnered with Broward County Public Schools to conduct a randomized controlled trial of the program with early elementary students. Fifty-six percent of students' families in the district qualify for free and reduced priced lunch and 13% of students are English learners / multi-language learners (ELs). The district identified 49 kindergarten classes across 13 schools to participate in the evaluation. Tutoring by Chapter One started in early November 2021 in some of the schools and was rolled out to all classrooms over the course of the next two months. The first year of the program lasted through the end of the kindergarten school year, in May 2022. Students who remained in their schools were expected to receive Chapter One tutoring in first grade during the 2022-23 school year, as well.

Sample and Randomization

The study consisted of 818 kindergarten students in 13 schools. These schools served 7,891 students with an average enrollment of 607 students per school. The percentage of students eligible for free- and reduced-priced lunch ranged from 79% to 92% (with an average of 84%). Students were considered eligible for the study if they were classified as Emergent Readers on the district's kindergarten screener and they had parental consent for participating in Chapter One. Panel A of Table 1 provides information on the demographics of the students in the RCT sample. We conducted a student-level randomization stratified by classroom. Within each of the 49 kindergarten classrooms, we randomly assigned the 818 eligible students to the treatment group

(i.e., to receive Chapter One tutoring; N = 420) or to the control group (i.e., to receive business-as-usual instruction; N = 398).

B. Data

We collected administrative data from the school district and Chapter One, including data on gender, race/ethnicity, English learner status, and whether students qualify for special education services.

Baseline Reading Skill

As a proxy for baseline reading skill, we use the district's administration of the Florida Kindergarten Readiness Screener (FLKRS), which was the Renaissance Star Early Literacy measure in Fall of 2021. The FLKRS must be administered to all public-school kindergarten students within the first 30 days of each school year. The literacy classifications for the scores are as follows: Early Emergent Reader (300 - 487), Late Emergent Reader (488 - 674), Transitional Reader (675 - 774), and Probable Reader (775 - 900). The Chapter One program included students who scored in either of the first two levels.

Primary Outcome

The primary outcome for the present study, which we pre-registered, is a binary indicator for whether students reached Reading Foundation Stage (RFS) 4 or higher at the end of their kindergarten year. While binary outcomes provide less information than continuous test scores, this benchmark was the stated goal that the district had for the program and we chose it as the main outcome in partnership. Chapter One follows a child's progression through six Reading

Foundation Stages. Upon mastering the Reading Foundation Stages, students continue to work with ELIs to practice oral reading and adaptive phonics content. Reading Foundation Stage 4 entails segmenting and blending CVC words (consonant-vowel-consonant, such as "cat and hot") and recognizing 30 common words by sight. Students who master Reading Foundation Stage 4 have learned the sounds for short vowels and most consonants. This level of reading is approximately equivalent to the Fountas & Pinnell Reading Level C and is the end-of-year target for kindergarteners.

Additional Outcomes

In addition to the binary Reading Foundation Stage 4 or higher indicator, we assessed students' average Stage level, as well as their Oral Reading Fluency (ORF) scores (standardized within our sample) and for a subset of schools for which it is available, students' District Reading Level tests. These latter two assessments are not the primary outcomes for kindergarten students and include domains not covered by Chapter One in Kindergarten, but will be key outcomes in future analyses and provide a continuous measure of learning.

The ORF assessment was created by Dr. Karen Kortecamp from George Washington University, and is similar to the Dynamic Indicators of Basic Early Literacy Skills (DIBELS) 8. To administer the assessment, the ELI begins by explaining that the student is to do their best reading of the passage aloud. If they are stuck on a word, the ELI would tell them the word so that they could keep reading. After checking that the student understands the directions, the ELI reads the title, then points to the first word and says "Begin". The ELI starts the timer when the student says the first word of the passage. If the student fails to say the first word after 3 seconds, the ELI tells them the word and marks it incorrect and starts the timer. The maximum time for each word is 3 seconds. At the end of 1 minute the ELI notes the last word read by marking it with a bracket. If the stop time falls mid-sentence the ELI allows the student to complete reading the sentence but does not record

scores for any words read beyond the stop bracket. The ELI records the words correct per minute and calculates accuracy by dividing the words correct per minute by the total words read and multiplying by 100. ELIs are told that they must follow the same script for every assessment without variation. ORF scores may have floor effects for kindergarten students.

C. Analysis

We preregistered our study design, hypotheses, and analytic plan on the Social Science Registry prior to conducting the primary analysis.[1] We use the following model to evaluate the difference between the treatment and control groups:

$$Y_{ijk} = \alpha + \beta_1 \cdot Treatment_i + \beta_2 \cdot FLKRS_i + \pi \cdot \boldsymbol{X}_i + \gamma \cdot T_i + \varepsilon_{ijk}$$

where Y_{ijk} is the outcome for student i in classroom j in school k; $Treatment_i$ is an indicator for whether student i was assigned to Chapter One; $FLKRS_i$ is a student's beginning of the year FLKRS score (included as a control for baseline achievement); $\boldsymbol{X}_i$ is a vector of student-level characteristics (i.e., indicators for gender, race, English learner status, and special education services); T_i is a classroom fixed effect; and ε_{ijk} is an error term. We also run models with only the classroom fixed effect, since randomization should eliminate selection bias. We add controls to the main model to increase power. We calculated the minimum detectable effect size (MDES) using https://powerupr.shinyapps.io/index/. Based on conservative assumptions (e.g., 33% of the variation is explained by covariates and baseline achievement), we have 80% power to detect an MDES of 0.164-standard deviations.

Additionally, we conducted exploratory analyses that study the heterogeneity of the treatment effects by pre-intervention characteristics and student demographics. Specifically, we looked at

[1] See: https://www.socialscienceregistry.org/trials/10810.

outcomes for students with different reading skills at the beginning of kindergarten (comparing Early Emergent Readers to students who scored as Late Emergent Readers or greater) and whether the program differentially impacts English-Language Learners and native English speakers. We also explored the variation in the treatment effect by school site.

D. Qualitative Data Details and Analysis

During the 2021-22 school year, members of our research team partnered with school districts, tutoring providers, and quarterback organizations that support implementation across districts to learn from their efforts (see Groom-Thomas et al., 2023). The goal of the study was to provide insight into common barriers to implementing tutoring and the ways that districts have overcome these barriers. Overall, the team conducted 112 semi-structured interviews with 90 interviewees. Nine interviews were conducted with Broward County and Chapter One stakeholders between April and June of 2022 (District Administrator – 2, School Leader – 1, Teacher – 3, Tutor – 1, Provider – 2). The research team conducted semi-structured interviews using protocols developed for each of the stakeholder groups from January through June 2022. As part of the broader research study, the team developed an a priori coding scheme grounded in prior literature (e.g., Burch et al., 2007; Heinrich et al., 2010; Jacob et al., 2015; Robinson & Loeb, 2021; Worthy et al., 2003).

For this study of Chapter One, we conducted a sequential analysis in which we focused on excerpts from interviewees that were particularly relevant to the success of Chapter One, including educator perceptions of the design, effectiveness, and scalability of the program. Specifically, we focused on excerpts related to Tutoring Goals and Outcomes (sub-codes include: Influence on students, and key design elements for program success), Curriculum and Instruction (sub-codes include: Student-Tutor Relationship, Personalized Approach, Connection to Tier 1, and Curriculum

Materials), and Funding and Costs. We use the findings from our qualitative exploration to supplement our discussion as to why the program was successful and may hold promise for scale.

IV. RESULTS

A. Descriptive Statistics, Balance Check, and Attrition

Table 1 provides details on the sample descriptive statistics and how the two conditions compared at baseline. The treatment group is slightly less likely to be White, more likely to be female, and more likely to be classified as an English learner. The treatment group also has slightly lower initial FLKRS baseline scores than the control group. Our preregistration specified that we would control for these features in our final model, and this should help to account for any covariate imbalance. The results are qualitatively similar with and without these controls.

We aimed to retain all students in our final analytic sample; however, we did see some attrition during the first year of the study. All treated individuals have start dates, with the exception of two students. One of those students withdrew from school and the other was moved to an Autism Spectrum Disorder classroom before their original class began the treatment. An additional 61 students withdrew from the participating schools before the end of the program (31 in the treatment condition, 30 in the control condition). We retained the students in our analytic sample if they had outcome data (e.g., three withdrawn students had end-of-year Reading Foundation Stage data and were included in our analysis).

Due to attrition and students missing tests, we have some missingness in the data. At baseline, 79 students are missing FLKRS scores. For all analyses, we provide two sets of results: (1) excluding students who do not have FLKRS baseline data and (2) imputing missing FLKRS scores with the sample mean and including an indicator for missing the score. At the end of the year, 74 students are

missing end-of-year Reading Foundation Stage data and 82 students are missing Oral Reading Fluency (ORF) assessments. Additionally, only 274 kindergarten students in the sample took the District Reading Level assessment because it was given in some schools and not in others. Table 2 shows that attrition from the sample due to missing data is equal across experimental conditions.

B. Reading Foundation Stage Results

We present our primary results in Table 3 and Figure 1. As Figure 1 illustrates, students who received Chapter One during kindergarten were 38 percentage points more likely to reach RFS Level 4 or higher by the end of kindergarten (70%) than students in the control group (32%). As Table 3 shows, this increase stems from students in the treatment group being, on average, about one Reading Foundation Stage ahead of students in the control group at the end of the year. Table 3 shows that the average student in the control group is at a Reading Foundation Stage 3 (M = 2.96) whereas students in the treatment group are at a Stage 4 (M = 3.97).

C. Other Reading Assessment Results

Table 4 shows the impact of being assigned to Chapter One on the Oral Reading Fluency measure and on the district assessment of early reading. Not all students took the assessments and, in some cases, students were excused from completing assessments if they were not considered academically ready. Attrition was equal across conditions, so we model the results with the outcome variable coded as "missing" (and therefore students missing the specific outcome are excluded from the analysis).

Chapter One scored, on average, 0.23 standard deviations higher on Oral Reading Fluency assessments. Only six of the 13 participating schools administered the District Reading Level assessment to kindergarten students. Among the schools that did, we find some evidence that students

receiving Chapter One scored 0.312 points (translating to 0.11 standard deviations) higher than the treatment students.

Literacy interventions often improve performance on outcomes developed by researchers or providers, which often represent skills similar to those taught in the program, more than they improve performance on standardized outcomes (Herrera et al., 2021). Although these results are exploratory, they are promising indicators that Chapter One has impacts beyond its own internal metrics.

D. Heterogeneity Analysis

Baseline Reading Ability

We conducted a heterogeneity analysis to understand the extent to which the effect of Chapter One differed based on students' baseline reading abilities. Table 5 shows the effect Chapter One had on kindergarten students who were classified as Early Emergent Readers at the beginning of the year compared to those who had more advanced reading abilities at the outset of the program. Overall, we see that the treatment effect estimates are largely consistent across the two ability groups. As Figure 2 illustrates, Early Emergent Readers who received the Chapter One program were 37 percentage points more likely to reach Reading Foundation Stage 4 or higher and more advanced readers were 44 percentage points more likely to reach the target stage. Because students are making equivalent gains no matter their baseline ability levels, those scoring higher at the outset ultimately achieve more advanced reading levels: 89% of students who were at least Late Emergent Readers reached Stage 4, compared to 60.9% of Early Emergent Readers. Similarly, both Early and Late Emergent readers in the treatment group scored higher on the Oral Reading Fluency Measure, by 20 percent and 30 percent of a standard deviation respectively.

English Language Learners

Table 6 and Figure 3 shows the impact of Chapter One on English-Language Learners and native English speakers. Again, the treatment effect estimates are consistent between the two groups of students. Students classified as English learners and native English speakers both were almost 40 percentage points more likely to reach the target Reading Foundation Stage after participating in Chapter One's program. The estimated effects of Chapter One on Oral Reading Fluency was positive and statistically significant for both groups, with estimates even higher for students classified as English learners, 0.384 standard deviations in comparison to 0.182 standard deviations for non-English learners.

V. DISCUSSION

We found that implementing the Chapter One program in kindergarten can meaningfully improve the reading ability of students. Almost 70% of students who received Chapter One tutoring reached the goal for kindergarten students, Reading Foundation Stage 4, by the end of the year. Comparatively, only 32% of students in the control group reached Stage 4. By reaching Stage 4, where the students can segment and blend consonant-vowel-consonant words, these students can "hit the ground reading" at the beginning of first grade. Students who enter first grade unable to decode consonant-vowel-consonant words may be at risk of failing to be fluent readers at the end of first grade. The results we found were largely homogenous across student populations and extended to other assessments, less likely to be affected by rater bias. We will continue to track students' progress through third grade, but the results from the first year of the evaluation are encouraging.

Our sequential analysis of the qualitative data resulted in identifying three primary factors that may have contributed to the early success of the program and speak to its potential for scale. These

three factors highlight the program features that appear to contribute to its effectiveness (1:1 instruction and classroom integration), as well those that might lead to scaling going forward (affordability and the staffing model).

Personalized, 1:1 Literacy Instruction Promoted Learning and Positive Relationships

We conducted the interviews before the results of the first year of the study were available, but the district personnel's perceptions of Chapter One aligned with the findings: the program was universally viewed as improving students' literacy. Overall, staff at Broward County were effusive about the benefits of students receiving 1:1 instruction in literacy. Much of the academic success of the program was credited to tutors being able to deliver instruction at the right level without other distractions.

The teachers in the sample noted that the personalized nature of the program was beneficial for young students: "Students, especially at this level… they're tiny and they're easily distracted…But when they're 1:1 with a tutor, no distraction, that's when they grasp the instruction the most… I cannot overstate how valuable it is because of that 1:1 that the children get with the tutor."

Another key benefit of the 1:1 sessions with a consistent tutor was that tutors were able to bond with the students over the course of the year and that led to students' continued growth. Tutors were present every day and became a familiar—and often adored—face for students. One teacher believed that the strong positive tutor-student relationships fueled their students' success and shared, "My kids were—I'm not even exaggerating—excited [to go to their tutoring sessions]." Because the sessions are personalized, tutors are able to provide students with a lot of positive reinforcement and praise, which led to students asking to "go next" to receive tutoring.

The visibility of the program's benefits for students appeared to promote teacher and staff buy-in—every person interviewed in the district expressed that they hoped the program would continue: "I would just love for this to continue because…I have really seen a lot of improvement in my students and that is really all that matters; that they are progressing at a good pace." And, although the features of Chapter One were discussed as being beneficial for students in general, it often came up how critical the program was for addressing the educational disruptions due to the COVID-19 pandemic. Educators were generally concerned that missing opportunities to develop literacy skills in kindergarten could hinder academic progress in later grades, and they saw individualized instruction as a way to combat unrealized learning.

Integrating Tutoring into the Classroom Facilitated Implementation and Buy-In

Providing Chapter One as an integrated aspect of students' classroom experience increased alignment between educators and tutors, which ultimately resulted in a streamlined learning experience for students and buy-in among staff. Because tutors spend the year working with particular classrooms, they are able to develop connections with teachers. For example, teachers who had Chapter One tutors in their classrooms viewed the program as supporting their work with students, as opposed to an add-on: "It takes a village, you know, to raise a child and indeed it with the help of this program, and us working together and [the tutors], who were here, you know, they really supported me supported my goals and the goals of the students and we work together as a team." This response by teachers reflects the goals of the district: as one district administrator put it, "they were trying really hard not just to add something else to [the teacher's] plate."

The consistent presence of the tutors, which allowed for tutors to cultivate positive relationships with students, also facilitated working relationships between the tutor and the teacher that not only increased buy-in but also improved students' educational experiences. The interviewed

stakeholders all highlighted how important the flow of information between the tutor and the teacher (or the school's literacy coach) was for helping students progress. Chapter One leadership sees working with the teachers as one of the crucial aspects of their model. One member of the organization talked about the importance of relationship building day one, even before the children arrive on campus.

Classroom integration of Chapter One facilitated open lines of communication between tutors and teachers, which ultimately set the conditions for student learning. Moreover, because teachers were able to see the gains students were making as a result of their participation in Chapter One, they found value in the program and communicated to leadership that they hoped it would continue in future years.

Scaling Effective Programs That Do Not Require Additional Hiring and Training of Staff

During the first year of the study, the Chapter One program was funded through the district with Elementary and Secondary School Emergency Relief (ESSER) funding. The available funds led many schools to be eager to take on the logistics of implementing the program. The school leader interviewed shared that their initial enthusiasm for the program stemmed, in part, from the fact that the program was not coming out of the school budget and they did not feel they were incurring additional costs. At less than $500 per student, which covers all aspects of the program, Chapter One can be an attractive option for districts looking for cost-effective interventions to ensure students are reading by third grade.

The cost of the program covers staffing, which saves districts from recruiting, hiring, and training part-time tutors. The district administrator was very clear that Chapter One's hiring model was a major contributor to the district using the program and its potential for scale: "[Chapter One] provides [tutors] with the professional development that they need, all the resources they need, and

the flexibility that they get to determine their schedule." The part-time tutors are thoughtfully placed near where they live and are compensated for the work that they do.

Providing 1:1 instruction is inherently costly and U.S. schools are not currently designed to offer this type of individualized support to all students (Bloom, 1981). Without overhauling traditional school staffing models, programs like Chapter One—which provide trained educators at reasonable costs—may be a promising solution for providing 1:1 early literacy interventions that result in student learning gains.

V. CONCLUSION

As policymakers look for solutions to ensure students are proficient readers by third grade, they might consider how the features of the Chapter One program might be applied more broadly. Chapter One's unique combination of short bursts of 1:1 instruction by trained staff, together with independent practice on digital devices precisely synched to the 1:1 instruction, delivers a program that is highly affordable and scalable. The program is also likely to be less obtrusive to classroom instruction than tutoring programs that pull out students for greater amounts of time. The program aligns with beginning reading curricula and is provided on a turnkey basis that appears to be easily implemented by districts and schools. Early literacy tutoring programs that leverage technology to support in-class tutors may be a low-cost, scalable, and effective way to deliver literacy intervention to students who need it most.

References

Bloom, B. S. (1981). All our children learning: A primer for parents, teachers and other educators.

Burch, P., Steinberg, M., & Donovan, J. (2007). Supplemental Educational Services and NCLB: Policy Assumptions, Market Practices, Emerging Issues. *Educational Evaluation and Policy Analysis*, *29*(2), 115–133.

Carbonari, M., Davison, M., DeArmond, M., Dewey, D., Dizon-Ross, E., Goldhaber, D., Hashim, A., Kane, T., McEachin, A., & Morton, E. (2022). *The Challenges of Implementing Academic COVID Recovery Interventions: Evidence from the Road to Recovery Project*.

Cavanaugh, C. L., Kim, A.-H., Wanzek, J., & Vaughn, S. (2004). Kindergarten Reading Interventions for At-Risk Students: Twenty Years of Research. *Learning Disabilities: A Contemporary Journal*, *2*(1), 9–21.

Cunningham, A. E., & Stanovich, K. E. (1997). Early reading acquisition and its relation to reading experience and ability 10 years later. *Developmental Psychology*, *33*(6), 934.

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.

ExcelinEd. (2022). *Comprehensive early literacy policy: State-by-state analysis of fundamental principles.* https://excelined.org/wpcontent/uploads/2021/10/ExcelinEd_PolicyToolkit_EarlyLiteracy_StatebyStateAnalysis_2021.pdf

Fiester, L. (2010). Early Warning! Why Reading by the End of Third Grade Matters. KIDS COUNT Special Report. In *Annie E. Casey Foundation*. Annie E. https://eric.ed.gov/?id=ED509795

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.

Gersten, R., Haymond, K., Newman-Gonchar, R., Dimino, J., & Jayanthi, M. (2020). Meta-Analysis of the Impact of Reading Interventions for Students in the Primary Grades. *Journal of Research on Educational Effectiveness*, *13*(2), 401–427.

Groom-Thomas, L., Leung, C., Loeb, S., Pollard, C., Waymack, N., & White, S. (2023). *Challenges and Solutions: Scaling Tutoring Programs*. Inter-American Development Bank. http://dx.doi.org/10.18235/0005070

Guryan, J., Ludwig, J., Bhatt, M. P., Cook, P. J., Davis, J. M., Dodge, K., Farkas, G., Fryer Jr, R. G., Mayer, S., & Pollack, H. (2023). Not too late: Improving academic outcomes among adolescents. *American Economic Review*, *113*(3), 738–765.

Heinrich, C. J., Burch, P., Good, A., Acosta, R., Cheng, H., Dillender, M., Kirshbaum, C., Nisar, H., & Stewart, M. (2014). Improving the Implementation and Effectiveness of Out-of-School-Time Tutoring. *Journal of Policy Analysis and Management*, *33*(2), 471–494.

Heinrich, C. J., Meyer, R. H., & Whitten, G. (2010). Supplemental education services under No Child Left Behind: Who signs up, and what do they gain? *Educational Evaluation and Policy Analysis*, *32*(2), 273–298.

Herrera, S., Phillips, B. M., Newton, Y.-C., Dombek, J. L., & Hernandez, J. A. (2021). Effectiveness of Early Literacy Instruction: Summary of 20 Years of Research. REL 2021-084. In *Regional Educational Laboratory Southeast*. Regional Educational Laboratory Southeast. https://eric.ed.gov/?id=ED613947

Herring, W. A., Bassok, D., McGinty, A. S., Miller, L. C., & Wyckoff, J. H. (2022). Racial and Socioeconomic Disparities in the Relationship Between Children's Early Literacy Skills and Third-Grade Outcomes: Lessons From a Kindergarten Readiness Assessment. *Educational Researcher*, *51*(7), 441–450.

Hubbard, D., Gross, K., Holod, A., Wilson, E., Lachowicz, M., & Ackman, E. (2022). *Massachusetts Early literacy Tutoring Study Final Report*. https://www.doe.mass.edu/research/reports/2022/06early-literacy-tutoring-study.pdf

Jacob, R., Armstrong, C., & Willard, J. (2015). Mobilizing volunteer tutors to improve student literacy: Implementation, impacts, and costs of the Reading Partners Program.

Lesnick, J., Goerge, R., Smithgall, C., & Gwynne, J. (2010). Reading on grade level in third grade: How is it related to high school performance and college enrollment. *Chicago, IL: Chapin Hall at the University of Chicago*, *1*, 12.

Neitzel, A. J., Lake, C., Pellegrini, M., & Slavin, R. E. (2022). A Synthesis of Quantitative Research on Programs for Struggling Readers in Elementary Schools. *Reading Research Quarterly*, *57*(1), 149–179.

Nickow, A., Oreopoulos, P., & Quan, V. (2020). The impressive effects of tutoring on prek-12 learning: A systematic review and meta-analysis of the experimental evidence.

Robinson, C. D., Bisht, B., & Loeb, S. (2022). *The inequity of opt-in educational resources and an intervention to increase equitable access*. Ed Working Paper. https://www.edworkingpapers.com/ai22-654

Robinson, C. D., & Loeb, S. (2021). High-Impact Tutoring: State of the Research and Priorities for Future Learning. In *EdWorkingPapers.com*. Annenberg Institute at Brown University. https://www.edworkingpapers.com/ai21-384

Sirinides, P., Gray, A., & May, H. (2018). The Impacts of Reading Recovery at Scale: Results From the 4-Year i3 External Evaluation. *Educational Evaluation and Policy Analysis*, *40*(3), 316–335.

Slavin, R. E., Lake, C., Davis, S., & Madden, N. A. (2011). Effective Programs for Struggling Readers: A Best-Evidence Synthesis. *Educational Research Review*, *6*(1), 1–26.

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*, *27*(1), 189–211.

U.S. Department of Education. (2022). National Assessment of Educational Progress (NAEP) 2022 Long-Term Trend Assessment Results: Reading and Mathematics. https://www.nationsreportcard.gov/highlights/ltt/2022/

U.S. Office of the Press Secretary. (2022). *FACT SHEET: Biden-Harris Administration Launches National Effort to Support Student Success*. https://www.whitehouse.gov/briefing-room/statementsreleases/2022/07/05/fact-sheet-biden-harris-administration-launches-national-effort-to-supportstudentsuccess/?utm_content=&utm_medium=email&utm_name=&utm_source=govdelivery&utm_term=

Vygotsky, L. S. (1978). Mind in Society: Development of Higher Psychological Processes. Harvard University Press.

Wanzek, J., Stevens, E. A., Williams, K. J., Scammacca, N., Vaughn, S., & Sargent, K. (2018). Current Evidence on the Effects of Intensive Early Reading Interventions. *Journal of Learning Disabilities*, *51*(6), 612–624.

Wanzek, J., Vaughn, S., Scammacca, N., Gatlin, B., Walker, M. A., & Capin, P. (2016). Meta-analyses of the effects of tier 2 type reading interventions in grades K-3. *Educational Psychology Review*, *28*, 551–576.

Westall, J., & Cummings, A. (2022). The Effects of Early Literacy Policies on Student Achievement. *Available at SSRN 4427675*.

Worthy, J., Prater, K., & Pennington, J. (2003). "It's a Program That Looks Great on Paper": The Challenge of America Reads. *Journal of Literacy Research*, *35*(3), 879–91.

Figure 1.
Percent of Kindergarten Students Achieving Target Reading Stage by Treatment Status

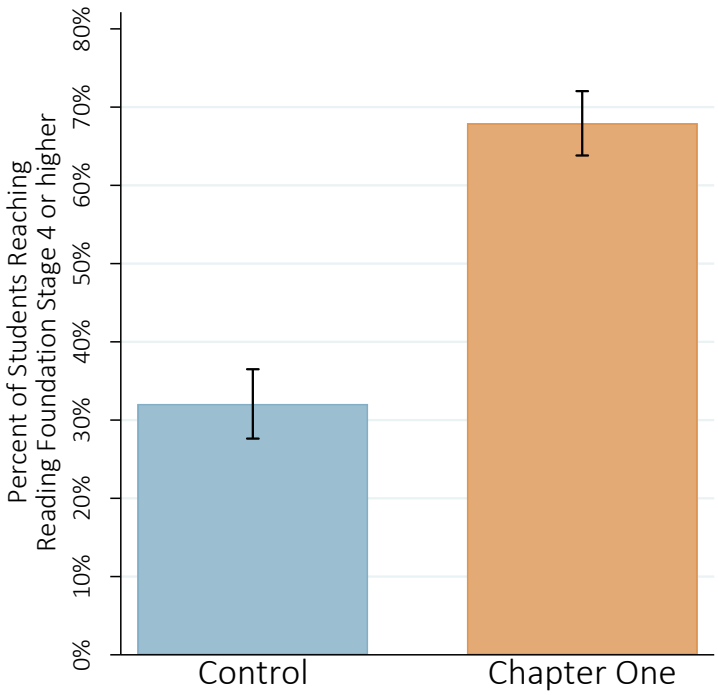

Figure 2.
Percent of Students Achieving Target Reading Stage by Treatment Status and Literacy Classification at the Beginning of the School Year

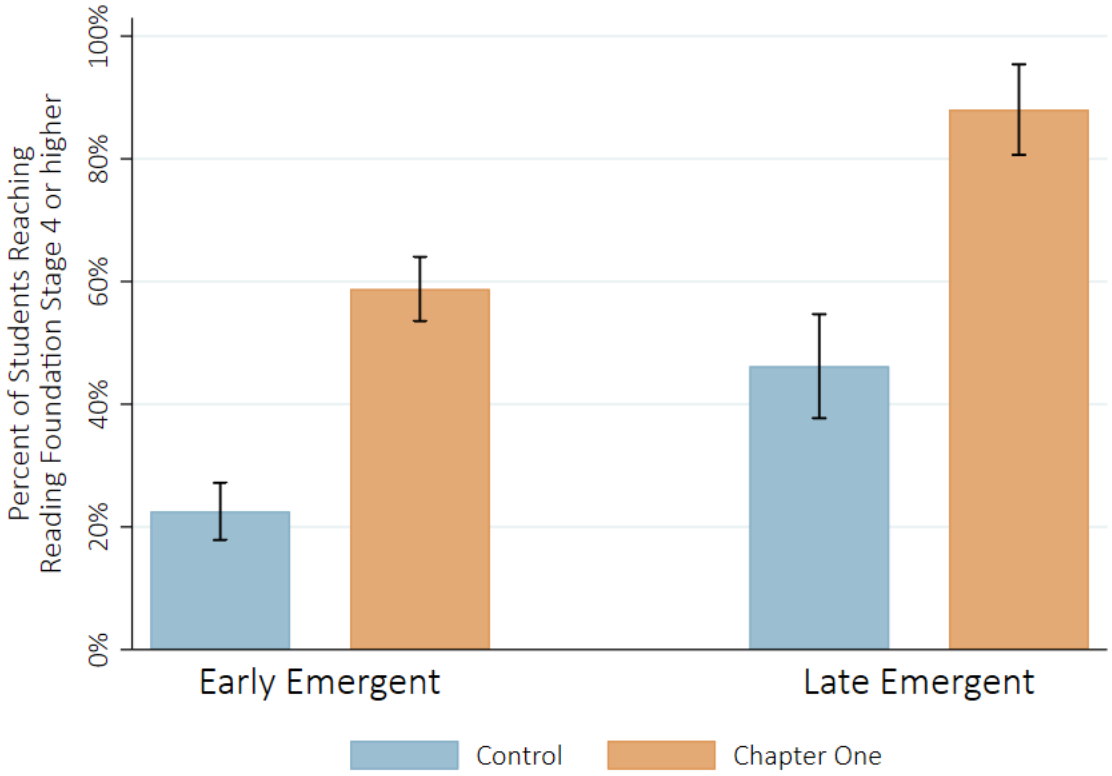

Figure 3.
Percent of Students Achieving Target Reading Stage by Treatment Status and English-Language Learner Status

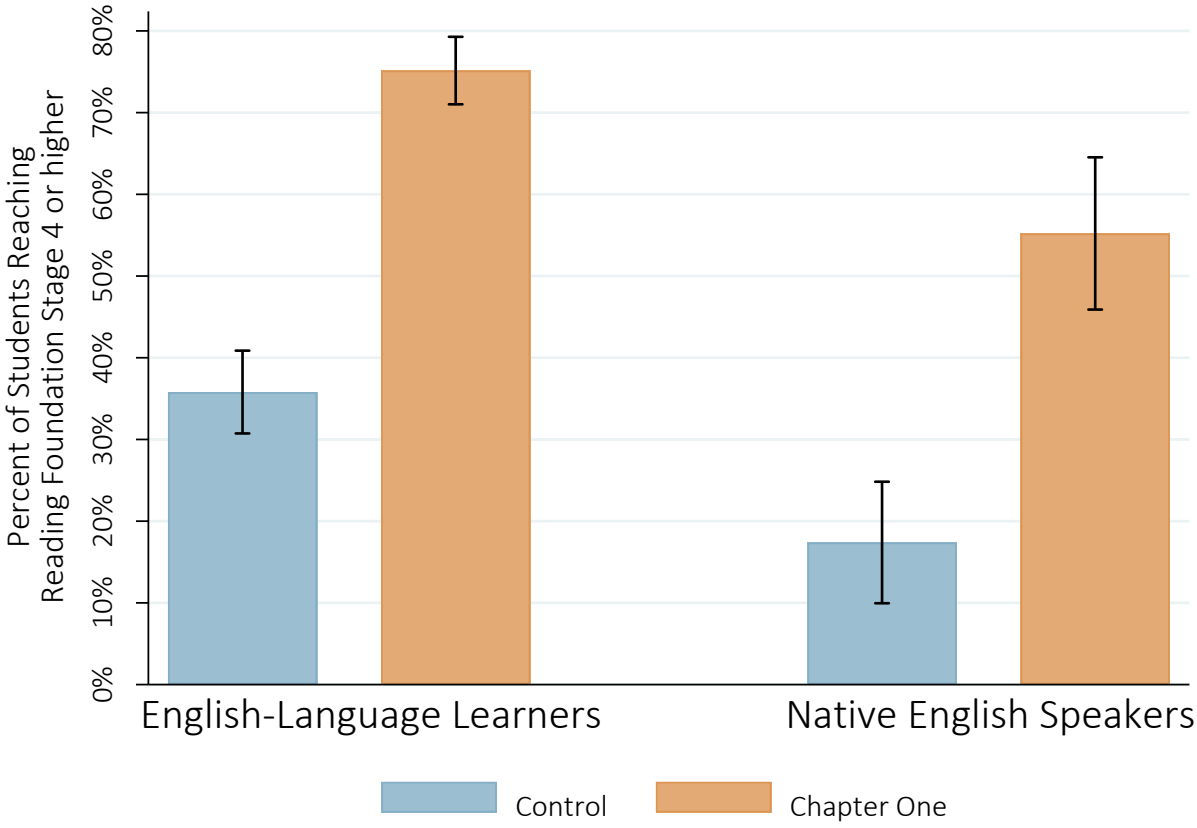

Table 1. Sample Descriptive Statistics and Balance Test for Student Baseline Measures

	Panel A: Overall			Panel B: Treatment			Panel C: Control					
	Mean	Std. Dev.	N	Mean	Std. Dev.	N	Mean	Std. Dev.	N	Diff	SE	
White	0.04		818	0.02		420	0.07		398	-0.05	0.01	***
Black	0.72		818	0.73		420	0.71		398	0.02	0.03	
Hispanic	0.21		818	0.22		420	0.19		398	0.02	0.02	
Other Race	0.03		818	0.03		420	0.03		398	0.00	0.01	
Female	0.47		818	0.50		420	0.44		398	0.07	0.04	+
English-Language Learner	0.28		818	0.31		420	0.25		398	0.06	0.02	*
Special Education	0.11		818	0.11		420	0.11		398	0.00	0.02	
Indicator for Missing Demographics	0.00		818	0.00		420	0.01		398	0.00	0.01	
FLKRS Scaled Score	452.75	93.78	739	445.31	90.63	381	460.66	96.52	358	-15.24	6.59	*
Indicator for Imputed FLKRS	0.10		818	0.09		420	0.10		398	-0.01	0.02	

Notes: Standard errors clustered at the classroom-level. The FLKRS is a screening instrument, known as the Florida Kindergarten Readiness Screener (FLKRS), which must be administered to all public school kindergarten students within the first 30 days of each school year. *** p<0.001, ** p<0.01, * p<0.05, + p<0.10.

Table 2. Attrition Analysis on Outcome Measures

	Attrition (End of the School Year Outcomes)			
	Reading Foundations	Reading Foundations Growth	Oral Reading Fluency (ORF)	District Reading Level
Treatment	0.0188	0.0196	0.0151	0.00520
	(0.0214)	(0.0205)	(0.0206)	(0.0108)
Constant	0.900***	0.919***	0.892***	0.332***
	(0.0110)	(0.0105)	(0.0106)	(0.00553)
Observations	818	818	818	818
R^2	0.073	0.058	0.075	0.921
Control Group: Mean	0.919	0.939	0.907	0.337
Classroom FE's	Yes	Yes	Yes	Yes

Notes: Standard errors clustered at the classroom-level are shown in parentheses. All regressions include classroom fixed effects. *** p<0.01; ** p<0.05; * p<0.10.

Table 3. The Effect of Chapter One on Reading Foundation Stage Levels

	Panel A: Reading Foundation Stage 4 or Higher				Panel B: Reading Foundation Stage			
	(1)	(2)	(3)	(4)	(1)	(2)	(3)	(4)
Treatment	0.363***	0.367***	0.385***	0.383***	0.912***	0.978***	1.014***	1.012***
	(0.0432)	(0.0452)	(0.0443)	(0.0431)	(0.0902)	(0.0931)	(0.0910)	(0.0903)
Constant	0.318***	0.313***	0.307***	0.252***	2.961***	2.923***	2.907***	2.837***
	(0.0325)	(0.0235)	(0.0245)	(0.0787)	(0.108)	(0.0486)	(0.0485)	(0.191)
Observations	744	682	744	744	744	682	744	744
R^2	0.132	0.386	0.385	0.396	0.107	0.514	0.498	0.505
Control Group: Mean			0.318				2.961	
FLKRS Control	No	Yes	Yes	Yes	No	Yes	Yes	Yes
Imputed FLKRS	No	No	Yes	Yes	No	No	Yes	Yes
Student-level Controls	No	No	No	Yes	No	No	No	Yes
Classroom FE's	No	Yes	Yes	Yes	No	Yes	Yes	Yes

Notes: Standard errors clustered at the classroom-level are shown in parentheses. Student-level controls include dummy variables for female, white (omitted category), black, Hispanic, other race, English Language Learner, and special education. The FLKRS is a screening instrument, known as the Florida Kindergarten Readiness Screener (FLKRS), which must be administered to all public school kindergarten students within the first 30 days of each school year. *** p<0.01, ** p<0.05, * p<0.10.

Table 4. The Effect of Chapter One on Other Reading Achievement Outcomes

	Panel A: Panel A: Oral Reading Fluency (Z-score)				Panel B: District Reading Level			
	(1)	(2)	(3)	(4)	(1)	(2)	(3)	(4)
Treatment	0.123*	0.177**	0.217***	0.226***	0.103	0.204	0.309	0.312*
	(0.0671)	(0.0722)	(0.0665)	(0.0673)	(0.237)	(0.226)	(0.203)	(0.177)
Constant	-0.0639	-0.0973**	-0.111***	-0.169	3.508***	3.262***	3.206***	4.016**
	(0.0562)	(0.0379)	(0.0357)	(0.172)	(0.347)	(0.130)	(0.114)	(1.590)
Observations	736	676	736	736	274	260	274	274
R^2	0.004	0.439	0.418	0.427	0.000	0.502	0.495	0.508
Control Group: Mean			-0.064				3.508	
FLKRS Control	No	Yes	Yes	Yes	No	Yes	Yes	Yes
Imputed FLKRS	No	No	Yes	Yes	No	No	Yes	Yes
Student-level Controls	No	No	No	Yes	No	No	No	Yes
Classroom FE's	No	Yes	Yes	Yes	No	Yes	Yes	Yes

Notes: Standard errors clustered at the classroom-level are shown in parentheses. Student-level controls include dummy variables for female, white (omitted category), black, Hispanic, other race, English Language Learner, and special education. The FLKRS is a screening instrument, known as the Florida Kindergarten Readiness Screener (FLKRS), which must be administered to all public school kindergarten students within the first 30 days of each school year. *** p<0.01, ** p<0.05, * p<0.10.

Table 5. Heterogeneity Analysis by Literacy Classification at Beginning of School Year

	Panel A: Early Emergent Readers				Panel B: Late Emergent Readers or More Advanced			
	Achieve Stage 4 or Higher	Reading Foundation Stage	Reading Foundation Growth	Oral Reading Fluency (Z-score)	Achieve Stage 4 or Higher	Reading Foundation Stage	Reading Foundation Growth	Oral Reading Fluency (Z-score)
Treatment	0.374***	1.060***	1.222***	0.205***	0.436***	1.018***	1.088***	0.304**
	(0.0474)	(0.102)	(0.0883)	(0.0761)	(0.0760)	(0.173)	(0.186)	(0.146)
Constant	0.344***	3.138***	1.238***	-0.246	0.339	3.010***	0.501	0.165
	(0.0800)	(0.182)	(0.196)	(0.185)	(0.205)	(0.539)	(0.401)	(0.453)
Observations	480	480	493	474	264	264	267	262
R^2	0.400	0.514	0.544	0.343	0.477	0.523	0.524	0.402
Control Group: Mean	0.236	2.627	0.929	-0.395	0.449	3.493	1.305	0.457
FLKRS Control (imputed)	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Student-level Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Classroom FE's	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Notes: Standard errors clustered at the classroom-level are shown in parentheses. Over 70% of the students were classified as "Early Emergent" readers which is the lowest level. According to the definition, "an Early Emergent Reader is beginning to understand that printed text has meaning. The student is learning that reading involves printed words and sentences, and that print flows from left to right and from the top to the bottom of the page. The student is also beginning to identify colors, shapes, numbers, and letters." Student-level controls include dummy variables for female, white (omitted category), black, Hispanic, other race, English Language Learner, and special education. *** p<0.01, ** p<0.05, * p<0.10.

Table 6. Heterogeneity Analysis by English-Language Learner Status

	Panel A: English-Language Learners				Panel B: Native English Speakers			
	Achieve Stage 4 or Higher	Reading Foundation Stage	Reading Foundation Growth	Oral Reading Fluency (Z-score)	Achieve Stage 4 or Higher	Reading Foundation Stage	Reading Foundation Growth	Oral Reading Fluency (Z-score)
Treatment	0.397***	1.320***	1.285***	0.384***	0.389***	0.965***	1.073***	0.182**
	(0.0843)	(0.148)	(0.149)	(0.109)	(0.0431)	(0.0964)	(0.104)	(0.0737)
Constant	0.223	2.918***	1.130***	-0.321	0.255**	2.801***	0.739***	-0.154
	(0.241)	(0.470)	(0.399)	(0.233)	(0.0989)	(0.225)	(0.172)	(0.199)
Observations	208	208	215	205	536	536	545	531
R^2	0.432	0.604	0.606	0.505	0.424	0.505	0.447	0.419
Control Group: Mean	0.184	2.471	0.924	-0.509	0.362	3.118	1.124	0.078
FLKRS Control (imputed)	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Student-level Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Classroom FE's	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Notes: Standard errors clustered at the classroom-level are shown in parentheses. Student-level controls include dummy variables for female, white (omitted category), black, Hispanic, other race, English Language Learner, and special education. *** p<0.01, ** p<0.05, * p<0.10.