Published Citation:
Cortes, K. E., Kortecamp, K., Loeb, S., & Robinson, C. D. (2025). A scalable approach to high-impact tutoring for young readers. *Learning and Instruction, 95*, 102021.
https://doi.org/10.1016/j.learninstruc.2024.102021

A scalable approach to high-impact tutoring for young readers

Kalena E. Cortes
Texas A&M University

Karen Kortecamp
George Washington University

Susanna Loeb
Stanford University

Carly D. Robinson[1]
Stanford University

Abstract

Learning to read is foundational to student success in early elementary school, yet many students are not proficient readers by third grade. A large body of research suggests high-impact tutoring is the most effective intervention to help struggling readers, however it can be hard to implement and scale. This study presents results from a randomized controlled trial of an early elementary reading tutoring program designed to be feasible at scale. Participants were 818 kindergarten students in a large southeastern district in the US who were classified as emergent readers on the district's screening tool. Within kindergarten classrooms, eligible students were randomly assigned to receive supplementary early literacy tutoring during the 2021-22 school year. The program embeds part-time tutors in the classroom to provide short bursts of instruction to individual students over the course of the school year. With the support of technology, tutors delivered a sequenced curriculum to students. At the end of the school year, students completed program and district literacy assessments. 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%) and scored 0.23-standard deviations higher on an oral reading fluency test than the control group. The results were largely homogenous across student populations and extended to district-administered assessments. 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.

Keywords: Tutoring; Early Literacy; Randomized Controlled Trial; Personalized Instruction; Educator-Student Relationships

---

[1] Corresponding author. E-mail address: carly_robinson@stanford.edu

Acknowledgements: We are grateful to our partners in this research, Chapter One and Broward County. We especially thank Seth Weinberger, Kim Whitten, and Lori Canning for their invaluable contribution to project implementation, as well as Sara White, Melissa Gentry, and Danila Serra for their research assistance. We received insightful feedback and support from the National Student Support Accelerator team, in particular Sarah Novicoff and Nancy Waymack. We thank the Smith Richardson Foundation and Arnold Ventures/Accelerate for their support of our full research program.

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, particularly 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. In the short run, scaling promising practices likely requires that they 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 (RCT) 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 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 when they might benefit most from personalized instruction.

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 (Elbaum et al., 2000; Dietrichson et al., 2017; Neitzel et al., 2022; Slavin et al., 2011). Research consistently demonstrates that tutoring interventions have substantial positive effects on student learning (average effect size = 0.29-standard deviations), often translating to an additional 3-15 months of schooling (Nickow et al., 2024). 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., 2024; 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 (e.g., Robinson et al., 2022), several syntheses of research on

early reading interventions converge to suggest providing students with personalized reading instruction over an extended time period consistently demonstrate the largest improvements in reading achievement. However, these syntheses slightly differ on the magnitude of the effects and what characteristics drive effectiveness in early literacy interventions.

Elbaum and colleagues (2000), who reviewed 29 studies of one-on-one reading tutoring programs in the late twentieth century, found a large effect size of 0.41-standard deviations on reading outcomes. The analysis revealed that trained college students and community members could serve as effective tutors. Slavin and colleagues (2011) examined 70 research studies on programs for struggling readers and found that, after whole-class interventions, one-on-one, phonetic tutoring from a teacher was the most effective strategy for improving student reading performance (ES = 0.38). More recently, the meta-analysis by Nietzel and colleagues (2022) confirmed that one-on-one tutoring interventions produced the largest effect sizes (ES = 0.41), however teachers were not necessarily more effective than teaching assistants. The meta-analysis by Gersten and colleagues (2020) also identified that one-on-one reading interventions tended to have the highest effect sizes, but that difference was not statistically distinguishable from the average student learning gains from small group instruction.

Cavanaugh and colleagues (2004) reviewed 27 reading interventions from the late 20th century and the early 2000s. They found that the majority of tested interventions were delivered in small groups (2-7 students), compared to one-on-one or whole class focused, and those small group interventions had the most consistently positive effect sizes (Cavanaugh et al., 2004). Wanzek and colleagues' two meta-analyses focused on Tier 2 (27 studies; 2016) and intensive supports (25 studies; 2018) and found no consistent evidence of differences in effect sizes based on the group size or intervention type; however, all effective reading interventions tended to have a high level of

standardization via well-prescribed lessons, instructional content addressing foundational reading skills, and delivered by instructors other than the teacher (Wanzek et al., 2016; 2018). Notably, all these syntheses focused on struggling readers or students with reading disabilities (Cavanaugh et al., 2004; Elbaum et al., 2000; Gersten et al., 2020; Neitzel et al., 2022; Slavin et al., 2011; Wanzek et al., 2016, 2018).

In summary, although the researchers conducting the syntheses had different inclusion criteria or foci, they all identified one-on-one or small group personalized instruction as an effective way to improve student reading achievement, particularly for students who may be struggling with reading or at-risk of falling behind (Cavanaugh et al., 2004; Gersten et al., 2020; Neitzel et al., 2022; Slavin et al., 2011; Wanzek et al., 2016, 2018). Therefore, although programs may differ in delivery or approaches, the 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). This manuscript extends the existing literature in two important ways. First, most of the studies focus on at-risk learners, whereas Chapter One targets all students in a classroom—both at-risk and general education students. Second, the meta-analyses rely on studies evaluating more traditional tutoring programs, whereas the present program delivers instruction in short bursts (versus longer sessions) and relies heavily on technology to guide independent practice and instruction (discussed more in Section 2.3).

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). Even so, 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. Early elementary-age children are also often more responsive to educational interventions, particularly literacy interventions, than children in older grades (Hill et al., 2008). Thus, early literacy tutoring initiatives have the potential to be an especially cost effective way to improve student achievement.

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 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 other students without adult support (Herrera et al., 2021; Nickow et al., 2024). 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, the departing student notifies the next student to join the ELI to minimize interruptions of classroom instruction. Although the ELIs instruction follows the Chapter One curriculum (and therefore does not necessarily reflect the specific lessons a classroom teacher might deliver on a given day), they are expected to coordinate with the classroom teacher to share information about student progress and needs.

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 understanding, 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.

Prior research on early reading interventions that involved computer-assisted instruction alone found few effects on reading achievement (Dynarski et al., 2007; Slavin et al., 2009; 2011). However, there is evidence that a blended learning tutoring model, or combining human tutors with computer-assisted instruction, can effectively improve student reading achievement. In most cases, programs focused on technology-assisted one-to-one tutoring (e.g., Chambers et al., 2008) or technology-assisted small-group tutoring (e.g., Chambers et al., 2011; Madden & Slavin, 2017). One study examined the impact of alternating human tutoring with computer-assisted instructional technology on high school math student outcomes (Bhatt et al., 2024) which, programmatically, is the most similar to Chapter One's blended learning tutoring model. The results showed that the effects of the alternating model were on par with human tutoring alone (Bhatt et al., 2024; Guryan et al., 2023), however we are not aware of any studies of early literacy tutoring programs that deploy human tutors alongside an independent computer-assisted instruction component for students.

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, though 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 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 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 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 compared to those

not assigned to the program (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 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 in these schools ranged from 79% to 92% (with an average of

84%). All kindergarten classrooms in participating schools took part in the study. Students were considered eligible for the study if they had parental consent to participate 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). The randomization was designed such that, if there was an unequal number of students in a classroom, the extra student was assigned to the treatment group.

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).

Primary Outcome

The primary outcome for the present study is a binary indicator for whether students reached Reading Foundation Stage (RFS) 4 or higher at the end of their kindergarten year. All students, regardless of condition assignment, completed the RFS assessment at the beginning and end of the

school year. 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. While binary outcomes provide less information than continuous test scores, this benchmark was the stated goal that the district had for the program. We present results for both the binary and continuous outcome. We received a single summative Reading Foundation Stage score for each student and, therefore, are not able to calculate the internal reliability for the sample involved in this study. The assessment has been used in prior studies that suggest it correlates strongly with other nationally normed assessments (Kortecamp & Peters, 2023).

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. Both assessments were only administered once (if at all, in the case of the District Reading Level test) at the end of the school year.

The Oral Reading Fluency 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. All students in the study completed the ORF at the end of the school year. 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 will tell them the word so that they can 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. Oral Reading Fluency scores may have floor effects for kindergarten students. Prior research demonstrates high correlations with nationally validated ready assessments (see Kortecamp & Peters, 2023).

During this school year, Broward County used the Benchmark Assessment System (BAS) from Fountas and Pinnell Leveled Literacy Intervention System as its end-of-year District Reading Level assessment. In schools that administered the assessment to kindergarten students, educators observe, code, and analyze students' reading behavior while the student reads orally. In our sample, student scores ranged from 0 (pre-A) to 21 (U) (M = 3.56, SD = 2.74). We received only a single summative score from the district and therefore do not have reliability evidence for our sample. A

prior study using the test found a test re-test reliability ranging from 0.93-0.97 and high correlations with other reading assessments (Ransford-Kaldon et al., 2013), although prominent literacy researchers suggest it may be inappropriate as a research measure (see Hanford & Peak, 2021).

C. Analysis

We preregistered our study design, hypotheses, and analytic plan on the Social Science Registry prior to conducting the primary analysis (see: BLINDED; appended). We used the following model to evaluate the difference between the treatment and control groups:

$$Y_{ijk} = \alpha + \beta_1 Treatment_i + \beta_2 FLKRS_i + \pi \boldsymbol{X_i} + \gamma T_j + \varepsilon_{ijk}$$

where Y 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 and an indicator if the score was missing and had to be imputed); $\boldsymbol{X_i}$ is a vector of student-level characteristics (i.e., indicators for gender, race, English learner status, and special education services); T_j is a classroom (or strata) fixed effect; and ε_{ijk} is an error term. We also ran models uncontrolled and with only the classroom fixed effect, since randomization should eliminate selection bias (although we did see an imbalance on observables in practice). We added controls to the main model to increase power and add precision to the final estimate. The results are the same using a multi-level modeling approach. We calculated the minimum detectable effect size (MDES) using PowerUp. 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 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.

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. In Table 2, we present correlations among the assessments administered at the beginning of year (FLKRS Scaled Score) and the end of the year.

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, except 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 were missing end-of-year Reading Foundation Stage data and 82 students were missing Oral Reading Fluency 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 3 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 4 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 4 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 4 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 5 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).

Students assigned to Chapter One scored, on average, 0.23 standard deviations higher on Oral Reading Fluency assessments ($p < .01$). 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 ($p < 0.10$). 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. Panel A of Table 6 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. Table 6, Panel B formally models the interaction and shows there is no statistically significant interaction between assignment to treatment and prior reading abilities. 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 7 and Figure 3 shows the impact of Chapter One on English-Language Learners and native English speakers. Again, Panel A of Table 7 shows the treatment effect estimates are consistently positive for the two groups of students. Students classified as English learners and native English speakers were both almost 40 percentage points more likely to reach the target Reading Foundation Stage after participating in Chapter One's program. The estimated effect 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. Table 7, Panel B demonstrates that English learners experienced greater gains on the Oral Reading Fluency assessment when assigned to treatment compared to Native English speakers.

E. Qualitative Results

Perceptions of the 1:1 Literacy Instruction & 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.

Integration of the Tutoring Program into the Classroom

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.

Promising Features that Lead to Scaling Effective Programs

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 eagerly 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 they do.

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.

The findings from this study extend the already rich evidence base demonstrating that early literacy tutoring can have large positive effects on student achievement (Austin et al., 2017; Elbaum, 2000; Neitzel et al., 2022; Nickow et al., 2024). Although only half the children in a classroom were assigned tutoring, in practice Chapter One is a whole classroom approach to improving literacy. While much of the evidence for the effectiveness of early literacy interventions focuses on Tier 3 approaches (e.g., Austin et al., 2017) and at-risk students (e.g., Elbaum, 2000; Nietzel et al., 2022), this model of tutoring can be implemented at the classroom level and the results suggest that all students can benefit from receiving the program.

The Chapter One program differs from other studied early literacy tutoring initiatives in two important ways. First, the tutoring sessions occur in short bursts ranging from five to just over ten minutes of instruction. Second, ELIs are expected to adjust the number of sessions students receive per week based on individual student needs and data. However, the guidance on high-impact tutoring to date has clearly indicated that tutoring must be scheduled and occur several times per week (Robinson & Loeb, 2021). This evaluation suggests that district expectations for a program's dosage and frequency might need to be evaluated based on what is developmentally appropriate for the student, the curriculum, and the way in which students and tutors interact.

Our qualitative analysis ultimately identified the program features that appear to contribute to its effectiveness (1:1 instruction and classroom integration), as well as those that might lead to scaling going forward (affordability and the staffing model). Overall, the opportunity for students to receive 1:1 instruction in literacy during kindergarten was universally well-received by the staff. In particular, people noted that the tutors' consistent, embedded presence in the classroom allowed them to better personalize the instruction and coordinate with teachers.

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). However, there are many different sources districts can use to fund tutoring programs (NSSA, 2023). Without overhauling traditional school staffing models, programs that provide trained educators at reasonable costs may be a promising solution for providing 1:1 early literacy interventions that result in student learning gains.

A. Limitations & Future Directions

Our findings should be interpreted in light of the limitations of the study. One limitation was that the classroom ELIs administered the Reading Foundation Stage level and Oral Reading Fluency assessments to both treatment and control students and, thus, were not blind to students' assigned condition. As a result, there is the possibility that the assessors could bias the results. The qualitative data indicated that the ELIs generally became familiar with all the students in the classroom due to their physical presence. However, an ELI might feel motivated to see the students they worked with closely score well, potentially biasing the estimates for the Treatment group. Conversely, ELIs might hold these students to higher standards, which could bias the results towards the Control group. The district administered the District Reading assessment was a

standard district assessment and, thus, we would not expect any bias. However, for the 2021-22 school year, the district used the Fountas and Pinnell BAS, which has demonstrated inconsistent reliability in the field (EdReports, 2021; Ransford-Kaldon et al., 2013) and only a subset of the schools chose to administer the assessment to Kindergarteners. Thus, although all the estimates suggest a positive impact of the program, the future studies will benefit from exploring the impact on standardized district and state assessments.

Another limitation is that parental consent was required for students to be eligible to receive tutoring as part of this study. Although the vast majority of students' parents consented due to an intensive outreach period from the organization, these findings may not generalize to the population of students. Given that students from families with low income or whose first language is not English are less likely to actively consent (Liu et al., 2017), not including these students might result in more conservative estimates of the program's effects, especially given that we saw English learners were more responsive to the program.

Additionally, we see several opportunities for future research to develop this line of work. First, we intend to follow participating students' outcomes through third grade. So, while this manuscript only explores the impact at the end of kindergarten and does not provide information on long-term, our research team plans to share our longitudinal analyses after the 2024-25 school year. More research on whether early literacy tutoring has lasting impacts will help to illuminate the cost-effectiveness of the approach. Second, this study cannot disentangle whether students benefit more from short, frequent tutoring sessions or longer, pre-scheduled sessions (e.g., Robinson et al., 2024). The recent meta-analysis by Nickow and colleagues (2024) only looks at outcomes based on whether tutoring sessions were less or more than 30 minutes long, providing no guidance on the differential effectiveness between sessions that are 15 or 30 minutes, for instance. Given the

positive results reported in this manuscript and how positively the program was received by the district, we believe studies that aim to answer this question would better allow districts to consider the trade-offs between deploying different types of interventions to support young students. Third, a burgeoning field of research has already started to consider how technology can support human tutors (e.g., Bhatt et al., 2024; Madden & Slavin, 2017) and future studies might explore the most efficient way to combine technology and human tutoring to improve young students' reading outcomes.

VI. 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

Bhatt, M. P., Guryan, J., Khan, S. A., LaForest-Tucker, M., & Mishra, B. (2024). *Can Technology Facilitate Scale? Evidence from a Randomized Evaluation of High Dosage Tutoring* (No. w32510). National Bureau of Economic Research.

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. doi: 10.3102/01623737073020

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.

Dynarski, M., Agodini, R., Heaviside, S., Novak, T., Carey, N., Campuzano, L., ... & Sussex, W. (2007). Effectiveness of reading and mathematics software products: Findings from the first student cohort.

EdReports. (2021). Summary of Alignment & Usability: Fountas & Pinnell Classroom. https://www.edreports.org/reports/overview/fountas-pinnell-classroom-2020

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/wp-content/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. doi: 10.1257/aer.20210434

Hanford, E. & Peak, C. (2021). Influential authors Fountas and Pinnell stand behind disproven reading theory. *APMReports*. https://www.apmreports.org/story/2021/11/19/fountas-pinnell-disproven-childrens-reading-theory

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. doi: 10.3102/016237371036164

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. doi: 10.3102/0013189X221091535

Hill, C. J., Bloom, H. S., Black, A. R., & Lipsey, M. W. (2008). Empirical benchmarks for interpreting effect sizes in research. *Child development perspectives*, *2*(3), 172-177.

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. *MDRC*. https://files.eric.ed.gov/fulltext/ED558508.pdf

Kortecamp, K., & Peters, M. L. (2023). The Impact of a High-Dosage Tutoring Program on Reading Achievement of Beginning Readers: A Multi-Level Analysis. *Journal of Education for Students Placed at Risk (JESPAR)*, 1-19. doi: 10.1080/10824669.2023.2179056

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.

Liu, C., Cox Jr, R. B., Washburn, I. J., Croff, J. M., & Crethar, H. C. (2017). The effects of requiring parental consent for research on adolescents' risk behaviors: A meta-analysis. *Journal of Adolescent Health*, *61*(1), 45-52.

National Student Support Accelerator (2023). Funding for High-Impact Tutoring. Brief Report. https://studentsupportaccelerator.org/briefs/funding-high-impact-tutoring

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. (2024). The Promise of Tutoring for PreK–12 Learning: A Systematic Review and Meta-Analysis of the Experimental Evidence. *American Educational Research Journal*, *61*(1), 74-107.

Ransford-Kaldon, C., Ross, C., Lee, C., Sutton Flynt, E., Franceschini, L., & Zoblotsky, T. (2013). Efficacy of the Leveled Literacy Intervention System for K–2 urban students: An empirical evaluation of LLI in Denver Public Schools. *Memphis, TN: Center for Research in Educational Policy, University of Memphis*.

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

Robinson, C. D., Pollard, C., Novicoff, S., White, S., & Loeb, S. (2024). The Effects of Virtual Tutoring on Young Readers: Results from a Randomized Controlled Trial. (EdWorkingPaper: 24-955). Retrieved from Annenberg Institute at Brown University: https://doi.org/10.26300/rtxg-5111

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., Chambers, B., Cheung, A., & Davis, S. (2009). Effective Reading Programs for the Elementary Grades: A Best-Evidence Synthesis. Review of Educational Research, 79(4), 1391-1466. https://doi.org/10.3102/0034654309341374

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

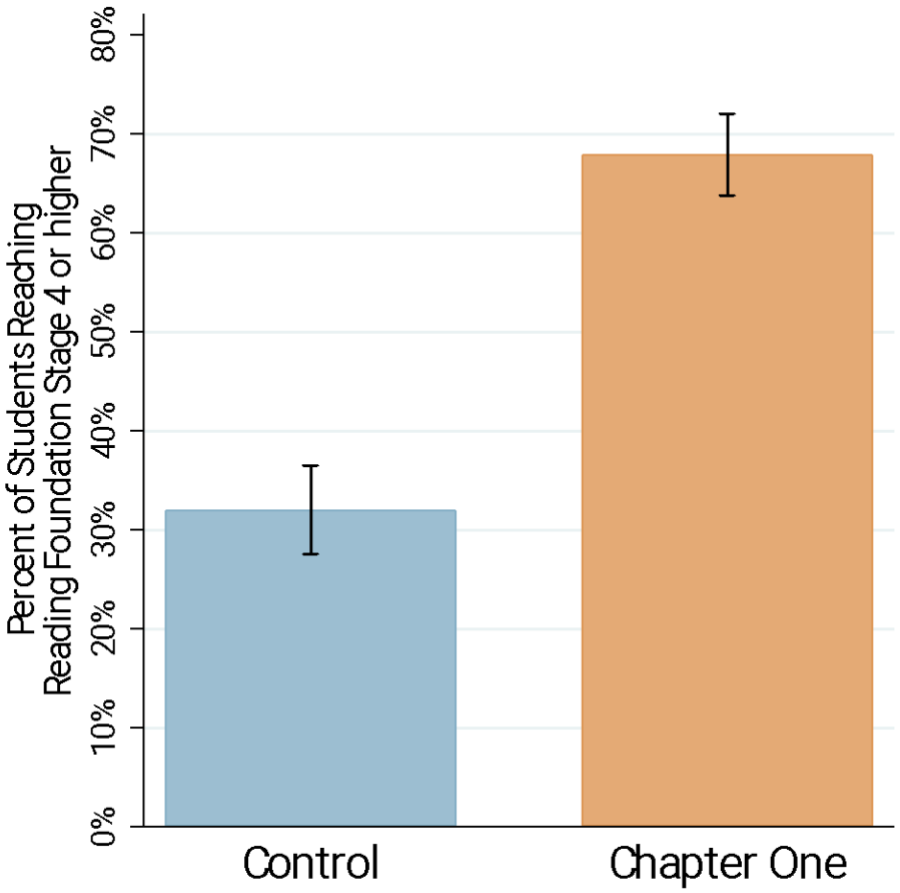

Figure 2.
Percent of Students Achieving Target Reading Stage by Treatment Status and Literacy Classification at the Beginning of the School Year

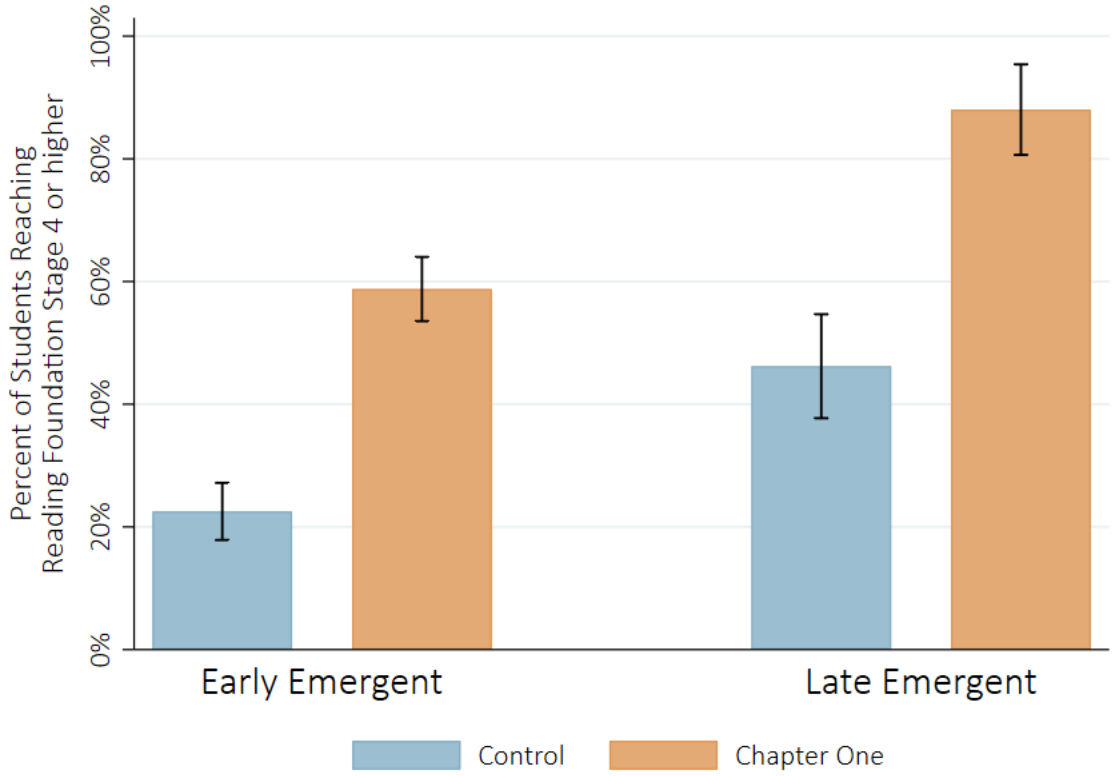

Figure 3.
Percent of Students Achieving Target Reading Stage by Treatment Status and English-Language Learner Status

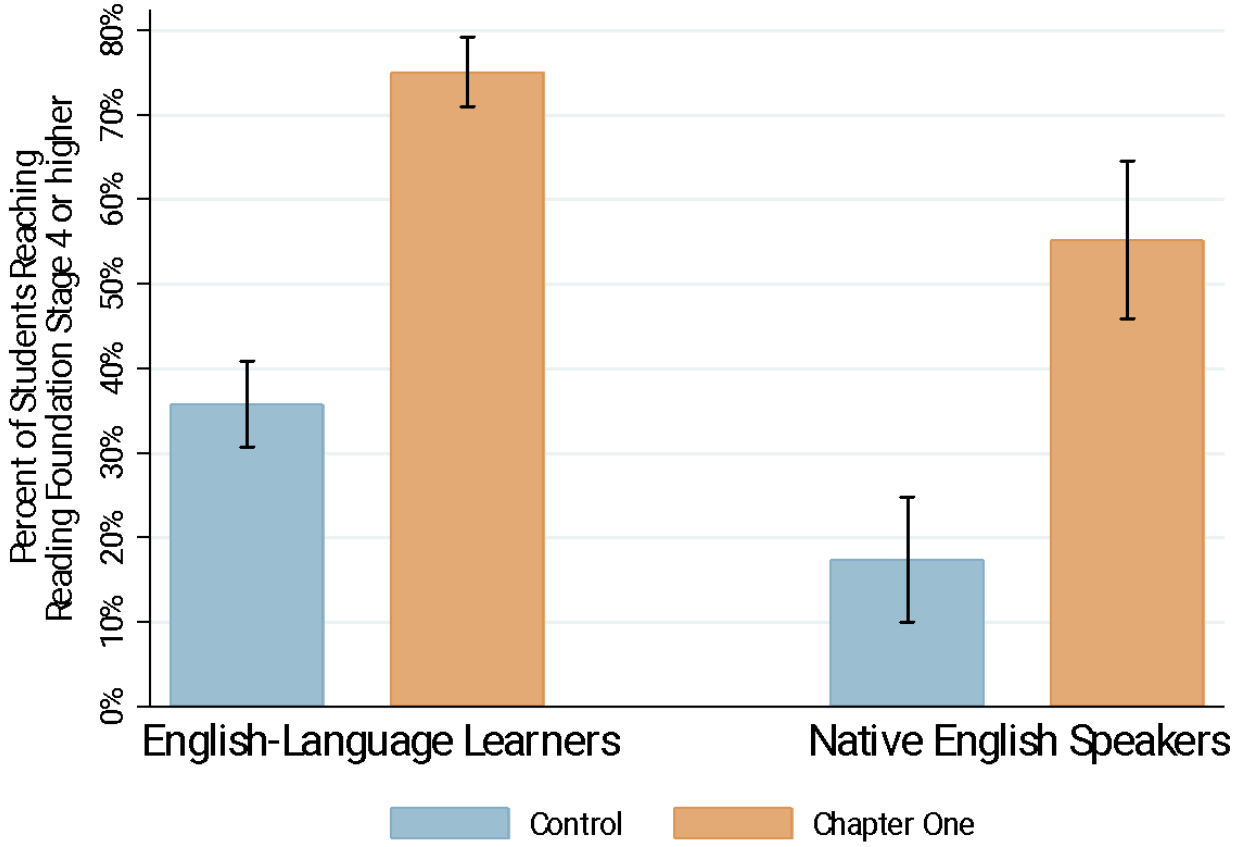

Table 1. Sample Descriptive Statistics and Balance Test for Student Baseline Measures

	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		36	0.02		9	0.07		27	-0.05	0.01	***
Black	0.72		586	0.73		305	0.71		281	0.02	0.03	
Hispanic	0.21		168	0.22		91	0.19		77	0.02	0.02	
Other Race	0.03		24	0.03		13	0.03		11	0.00	0.01	
Female	0.47		383	0.50		210	0.44		173	0.07	0.04	+
English-Language Learner	0.28		227	0.31		128	0.25		99	0.06	0.02	*
Special Education	0.11		87	0.11		45	0.11		42	0.00	0.02	
Indicator for Missing Demographics	0.00		4	0.00		2	0.01		2	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		79	0.09		39	0.10		40	-0.01	0.02	
Total Observations			818			420			398			

Notes. EL is English learner and SPED is special education. 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. Table 1. Sample Descriptive Statistics and Balance Test for Student Baseline Measures. + p<0.10, * p<0.05, ** p<0.01, *** p<0.001.

Table 2. Correlation matrix.

	(1) FLKRS Scaled Score	(2) Reading Foundation Stage	(3) Oral Reading Fluency (std)
(1) FLKRS Scaled Score	--		
(2) Reading Foundation Stage	0.45***	--	
(3) Oral Reading Fluency (std)	0.56***	0.67***	--
(4) District Reading Level	0.54***	0.57***	0.81***

Notes: The FLKRS was administered at the beginning of the 2021-22 school year. The other assessments were all administered at the end of the 2021-22 school year. *** p<0.01; ** p<0.05; * p<0.10.

Table 3. Attrition Analysis on Outcome Measures

	(1) Reading Foundation Stage	(2) Oral Reading Fluency	(3) District Reading Level
Treatment	0.019	0.015	0.005
	(0.021)	(0.020)	(0.010)
Classroom FE	Yes	Yes	Yes
R^2	0.073	0.075	0.921
Control Mean	0.919	0.907	0.337
Observations	818	818	818

+ p<0.10, * p<0.05, ** p<0.01, *** p<0.001.

Table 4. The Effect of Chapter One's Program on Reading Foundation Stage Levels

	Achieved RFS 4 or higher				End-of-year RFS			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Treatment	0.363***	0.358***	0.385***	0.383***	0.912***	0.913***	1.014***	1.012***
	(0.043)	(0.043)	(0.043)	(0.042)	(0.090)	(0.091)	(0.088)	(0.087)
FLKRS Control	No	No	Yes	Yes	No	No	Yes	Yes
Student Controls	No	No	No	Yes	No	No	No	Yes
Classroom FE	No	Yes	Yes	Yes	No	Yes	Yes	Yes
R^2	0.132	0.294	0.385	0.396	0.107	0.331	0.498	0.505
Control Mean	0.318				2.961			
Observations	744	744	744	744	744	744	744	744

Notes: We impute FLKRS scores for students who are missing baseline assessment data and include an indicator variable for students missing the baseline assessment in models 3, 4, 7, and 8. Student-level controls include dummy variables for female, white (omitted category), black, Hispanic, other race, English Language Learner, and special education. Robust standard errors clustered at the classroom level in parentheses. + p<0.10, * p<0.05, ** p<0.01, *** p<0.001.

Table 5. The Effect of Chapter One's Tutoring Program on Other Reading Achievement Outcomes

	(1) ORF (std)	(2) District Reading Level
Treatment	0.226**	0.312+
	(0.065)	(0.171)
FLKRS Control	Yes	Yes
Student Controls	Yes	Yes
Classroom FE	Yes	Yes
R^2	0.427	0.508
Control Mean	-0.117	3.398
Observations	736	274

Notes: We present the estimates from our preferred model that includes classroom fixed effects, baseline FLKRS scores (imputed for students who are missing baseline assessment data), an indicator variable for students missing a baseline score, and student-level controls (including dummy variables for female, white (omitted category), black, Hispanic, other race, English Language Learner, and special education). Robust standard errors clustered at the classroom level in parentheses. + p<0.10, * p<0.05, ** p<0.01, *** p<0.001.

Table 6. Heterogeneity Analysis by Literacy Classification at Beginning of School Year

Panel A.

	Early Emergent Readers			Late Emergent Readers or More Advanced		
	Achieve RFS 4 or higher	ORF (std)	District Reading Level	Achieve RFS 4 or higher	ORF (std)	District Reading Level
	(1)	(2)	(3)	(4)	(5)	(6)
Treatment	0.374***	0.205**	0.478	0.436***	0.304*	0.418
	(0.045)	(0.072)	(0.283)	(0.069)	(0.132)	(0.400)
FLKRS Control	Yes	Yes	Yes	Yes	Yes	Yes
Student Controls	Yes	Yes	Yes	Yes	Yes	Yes
Classroom FE	Yes	Yes	Yes	Yes	Yes	Yes
R^2	0.400	0.343	0.489	0.475	0.393	0.453
Control Mean	0.219	-0.419	2.464	0.452	0.404	4.597
Observations	480	474	167	262	260	106

Panel B.

	Achieve RFS 4 or higher	ORF (std)	District Reading Level
	(1)	(2)	(3)
Treatment	0.407***	0.216	0.098
	(0.0657)	(0.135)	(0.465)
Early Emergent	-0.252***	-0.977***	-2.318***
	(0.0571)	(0.123)	(0.528)
Early Emergent x Treatment	-0.0514	-0.0337	0.140
	(0.0667)	(0.156)	(0.688)
Constant	0.480***	0.705***	6.013***
	(0.0969)	(0.212)	(1.871)
FLKRS Control	Yes	Yes	Yes
Student Controls	Yes	Yes	Yes
Classroom FE	Yes	Yes	Yes
R^2	0.378	0.343	0.420
Observations	744	736	274

Notes: We present the estimates from our preferred model that includes classroom fixed effects, baseline FLKRS scores (imputed for students who are missing baseline assessment data), an indicator variable for students missing baseline scores, and student-level controls (including dummy variables for female, white (omitted category), black, Hispanic, other race, English Language Learner, and special education). Robust standard errors clustered at the classroom level in parentheses. + p<0.10, * p<0.05, ** p<0.01, *** p<0.001.

Table 7. Heterogeneity Analysis by English-Language Learner Status

Panel A.

	English Language Learners			Native English Speakers		
	Achieve RFS 4 or higher	ORF (std)	District Reading Level	Achieve RFS 4 or higher	ORF (std)	District Reading Level
	(1)	(2)	(3)	(4)	(5)	(6)
Treatment	0.397***	0.384***	0.450	0.388***	0.182*	0.281
	(0.074)	(0.096)	(0.338)	(0.041)	(0.071)	(0.216)
FLKRS Control	Yes	Yes	Yes	Yes	Yes	Yes
Student Controls	Yes	Yes	Yes	Yes	Yes	Yes
Classroom FE	Yes	Yes	Yes	Yes	Yes	Yes
R^2	0.419	0.501	0.534	0.421	0.418	0.490
Control Mean	0.153	-0.574	2.175	0.364	0.052	3.857
Observations	204	201	72	534	529	198

Panel B.

	Achieve RFS 4 or higher	ORF (std)	District Reading Level
	(1)	(2)	(3)
Treatment	0.386***	0.171**	0.240
	(0.0432)	(0.0735)	(0.215)
ELL	-0.105**	-0.321***	-0.338
	(0.0514)	(0.0815)	(0.289)
ELL x Treatment	-0.0138	0.202**	0.276
	(0.0655)	(0.0988)	(0.465)
Constant	0.251***	-0.147	4.036**
	(0.0778)	(0.170)	(1.587)
FLKRS Control	Yes	Yes	Yes
Student Controls	Yes	Yes	Yes
Classroom FE	Yes	Yes	Yes
R^2	0.396	0.429	0.508
Observations	744	736	274

Notes: We present the estimates from our preferred model that includes classroom fixed effects, baseline FLKRS scores (imputed for students who are missing baseline assessment data), and student-level controls (including dummy variables for female, white (omitted category), black, Hispanic, other race, English Language Learner, and special education). Robust standard errors clustered at the classroom level in parentheses. + p<0.10, * p<0.05, ** p<0.01, *** p<0.001.