James S. Kim
Harvard University, Graduate School of EducationJonathan Guryan[1]
Northwestern University, Institute for Policy Research

Thomas G. White
The Center for Advanced Study of Teaching and Learning
University of Virginia, Charlottesville, VA, USA

David M. Quinn
Harvard University, Graduate School of Education

Lauren Capotosto
College of the Holy Cross

Helen Chen Kingston
Harvard University, Graduate School of Education

[1]Corresponding concerning this article should be addressed to Jonathan Guryan, Institute for

Policy Research, Northwestern University, 2040 Sheridan Road, Evanston, IL 60208.

Email: j-guryan@northwestern.edu

Delayed Effects of a Low-Cost and Large-Scale Summer Reading Intervention on Elementary
School Children's Reading Comprehension

Abstract

To improve the reading comprehension outcomes of children in high poverty schools,

policymakers need to identify reading interventions that show promise of effectiveness at scale.

This study evaluated the effectiveness of a low-cost and large-scale summer reading intervention

that provided comprehension lessons at the end of the school year and stimulated home-based

summer reading routines with narrative and informational books. We conducted a randomized

controlled trial involving 59 elementary schools, 463 classrooms, and 6,383 second and third

graders and examined outcomes on the North Carolina End-of-Grade (EOG) reading

comprehension test administered nine months after the intervention, in the children's third- or

fourth-grade year. We found that on this delayed outcome, the treatment had a statistically

significant impact on children's reading comprehension, improving performance by .04 SD

(standard deviation) overall and .05 SD in high poverty schools. We also found, in estimates

from an instrumental variables analysis, that children's participation in home-based summer

book reading routines improved reading comprehension. The cost-effectiveness ratio for the

intervention compared favorably to existing compensatory education programs that target high

poverty schools.

Keywords: Reading interventions, reading comprehension, scale-up, summer reading

Nearly 80% of low-income children in fourth-grade fail to read proficiently on the

National Assessment of Educational Progress (NAEP), compared to 50% of middle-income

children (U. S. Department of Education, 2012). One way to address income-based gaps in

reading comprehension is to implement school-based (school year) interventions such as Success

for All and Reading Recovery, and there is emerging evidence from experimental evaluations

that this can be done on a large scale (May et al., 2014; Quint et al., 2015). However, research

that compares children's reading gains during school year and summer months indicates that the

gap in reading comprehension by family income increases significantly during the summer

months when children are not in school, particularly in the elementary grades (Alexander,

Entwisle, & Olson, 2007; Downey, von Hippel, & Broh, 2004; McCoach, O'Connell, Reis, &

Levitt, 2006). Therefore, to reduce and prevent income-based gaps in reading comprehension,

we need cost-effective and scalable interventions that begin in the elementary grades, target high

poverty schools, and expand the quality of literacy learning opportunities in the summer months

as well as the school year.

This article presents the results of a large randomized controlled trial of a summer

reading intervention called READS (Reading Enhances Achievement During Summer). Our

study was designed (a) to estimate treatment effects on a measure of reading comprehension

administered nine months after the intervention ended and (b) to examine whether children's

participation in what we call "home-based summer book reading routines" improved

performance on the delayed posttest.

Previous Studies of READS

READS has two core features. First, in the summer months children receive books that

are matched to their reading levels and interests. Second, teachers implement comprehension

lessons at the end of the school (Kim, 2006; Kim & White, 2008; White, Kim, Kingston, & Foster, 2014). We hypothesize that the comprehension lessons are a critical program component, because in prior work (Kim & White, 2008), we showed that providing matched books alone or matched books and oral reading fluency instruction was insufficient to improve comprehension outcomes. The purpose of the lessons is to encourage children to complete the home-based summer book reading routines.

Previous experimental studies have yielded mixed evidence of the effectiveness of READS in improving short-term reading comprehension outcomes. Two smaller efficacy trials demonstrated positive short-term effects on fall reading comprehension (Kim, 2006; Kim & White, 2008). We evaluated a READS intervention that replicated the core model and added summer scaffolding at home in the form of teacher phone calls (White et al., 2014). This study found no main effect of either the core treatment or the enhanced condition for a sample of third grade students in 19 high and moderate poverty schools. However, for high poverty schools, defined as schools where 75% to 100% of the students were receiving free or reduced-price lunches (FRL), there were positive effects ($d = .08$ and .11 for the core and enhanced treatment conditions, respectively).

More recently, we studied the effects of READS with a large sample of second and third grade students in 59 North Carolina schools (Guryan, Kim, & Quinn, 2014). In addition to the core components of lessons and books, the intervention included an afterschool family literacy event in which participating parents received suggestions for ways to support their child's summer reading activities. On a posttest administered in the fall, there was no significant difference overall between students who were randomly assigned to receive READS and control

students who received math lessons and no summer books. However, there was a significant treatment effect for third grade girls ($d = .07$).

These findings suggest that it may be possible to reduce summer loss and income-based achievement gaps by implementing a summer reading program like READS, especially for third grade girls, students in high poverty schools, and for older children with sufficiently strong word reading skills. However, the existing studies are limited in three important ways. First, with the exception of the experiment by Guryan, Kim, and Quinn (2014), the studies were small, involving 1 to 19 schools in a single school district. Second, they examined only short-term effects on a researcher-administered reading comprehension test given immediately after the intervention. Third, while they showed that the intervention increased self-reported summer book reading, they offered no evidence of a link between reading outcomes and the specific kinds of reading activities that the lessons and family nights aimed to increase.

The current longitudinal study addresses these limitations by extending our experimental investigation (Guryan et al., 2014). We followed children who were randomly assigned to READS or the control condition for nine months after the intervention ended and analyzed their scores on a reading comprehension test that was administered as part of federal and state accountability requirements. Additionally, we measured students' participation in *home-based summer book reading routines,* defined as reading the books and using comprehension routines taught in the lessons (described later), with assistance from their parents who participated at least to the extent of mailing back to us a document we call a "trifold" (see Figure 2 and 3 in the supplementary online materials). And we assessed the extent to which children who participated to a greater extent in more home-based summer book reading routines had better reading comprehension outcomes at the end of the academic school year.

Theoretical and Empirical Background

Some scholars have hypothesized that income-based gaps in reading grow more during the summer because the "resource faucet," including access to books and support for reading activities, is closed (Entwisle, Alexander, & Olson, 2000; Guryan, Hurst, & Kearney, 2008). There are two common approaches to keeping the learning faucet open and promoting summer learning for low-income children: school (classroom)-based interventions and home-based interventions (Kim, & Quinn, 2013). A third approach is to combine features of both, as we do in READS. In school-based summer interventions, the critical mechanism that promotes children's reading achievement is teacher-directed instruction (Connor, Morrison, & Katch, 2004; Tseng & Seidman, 2007). School-based interventions sponsored by public school districts typically employ certified teachers to implement structured literacy and math programs that are designed to remediate children's academic weaknesses and prepare them for mastering academic content to be presented in the upcoming school year (McCombs et al., 2011). Other school-based interventions sponsored by community-based organizations rely on college and graduate students rather than certified teachers to implement academic and enrichment activities (Chaplin & Capizzano, 2006; McCombs et al., 2014).

Home-based summer interventions are designed to stimulate child-initiated book reading activities as well as family and parent-supported literacy activities at home (Allington, et. al., 2010; Senechal & Young, 2008). Home interventions share several common program features and aim (a) to increase children's access to a wide variety of narrative and informational texts and increase their exposure to print, (b) to promote children's intrinsic motivation to read books at home during the summer months, and (c) to increase family involvement (Mol & Bus, 2011).

READS Program Theory and Evidence Base

The READS intervention can be thought of as a hybrid approach that relies primarily on child-initiated and parent-supported summer book reading routines while capturing some of the benefits of teacher-directed reading instruction. The current version of READS includes (a) end-of-year reading instruction provided by classroom teachers, (b) an after-school family literacy event at the end of the school year, and (c) home-based summer reading routines with appropriately challenging and interesting books. The reading instruction in the classroom lessons and family night include activities based on a trifold that prompts children to use a comprehension routine with each summer book.

The program theory for READS is displayed in Figure 1. The program theory is supported, first, by research suggesting that the match between children's skill levels and interests and the texts they are reading is an important ingredient in an effective summer reading program (Kim & Guryan, 2010; Guthrie & Humenick, 2004). Second, teachers can strengthen the efficacy of summer reading programs by instructing children to use reading comprehension routines with narrative and informational books that they apply when reading independently at home without teacher assistance (Kim, 2007; Kim & White, 2008; Kim & Guryan, 2010). Furthermore, parents can encourage their children's home-based book reading activities if they understand the reading comprehension routines that were taught to their children (Senechal & Young, 2008). Accordingly, we implement an afterschool family literacy event that is designed to strengthen the school and home microsystems by providing information to parents about the reading comprehension routines that children are encouraged to use with the books that are mailed home during the summer (Bronfenbrenner & Morris, 2006; Lareau, 1989).

The access to matched books and teacher-directed lessons in classrooms and the family literacy event are intended to work in unison to support the goal of increasing the amount of

home-based summer book reading routines. As described by other researchers, reading engagement is a multi-dimensional construct that captures aspects of children's individual behavior, motivation, and cognitive engagement (Fredricks, Blumenfeld, & Paris, 2004). Engaged readers read widely and independently; enjoy reading about personally interesting topics; are internally motivated to read; and use comprehension routines to think deeply about text (Guthrie, Wigfield, & You, 2012; Wigfield et al., 2008). We did not directly measure student engagement but did measure the extent to which students participated in the summer reading program using return of the trifolds as a proxy for participation in the home-based book reading routines.

Effectiveness and Costs of School-Based and Home-Based Summer Interventions

Research suggests that home-based interventions and home- and school-based interventions like READS are apt to be more cost-effective than school-based interventions. A previous meta-analysis (Kim & Quinn, 2013) showed that school-based and home-based summer interventions had roughly comparable short-term effects of about .10 SD on reading achievement. The estimate for school-based interventions did not include a recent large randomized controlled trial of school-based voluntary summer learning programs that found no significant effects on reading achievement (McCombs et al., 2014). Regarding costs, another study by RAND researchers (McCombs et al., 2011) revealed that the per child costs of school-based summer learning programs ranges from $1,109 to $2,801 per student, compared to $245 for home-based programs.

Summary of the Literature

Recent studies of school-based, home-based, and school- and home-based summer interventions leave the field at a crossroads. Large-scale school-based voluntary summer

programs may not be effective, and they are clearly expensive. Home-based and particularly

home- and school-based summer interventions programs may be a more cost-effective approach

for improving reading comprehension, but the existing studies involve mostly small samples and

immediate posttests and shed no light on the specific home-based summer book reading routines

that may improve reading comprehension outcomes at the end of the school year.

Research Questions

In this study, we aim to provide the most definitive evidence to date on whether a large-

scale and low-cost summer reading intervention can produce durable and positive effects on

reading comprehension outcomes and, if it works, to provide an explanation of why. We address

two research questions:

(1) What is the intent-to-treat effect of READS on a delayed posttest, a state-required reading

comprehension test administered in third- and fourth-grade?

 (2) Does the extent to which children participate in home-based summer book reading routines

lead to better reading comprehension outcomes at the end of the school year?

METHOD

Setting and Participants

A total of 59 elementary schools in 7 North Carolina public school districts participated

in this study, including 39 high poverty schools (75-100% free- and reduced-price lunch; FRL)

and 20 moderate poverty schools (61-74% FRL). The participating schools were located in 2

metropolitan school districts including suburban and central city schools, 3 mid-sized urban

districts, and 2 rural districts. Overall, 71% of the children (n = 6,383 children) received parental

consent to participate in the study.   The baseline sample included children from mostly low-

income and minority families. After pretesting, we implemented a two-step random assignment

procedure, as described below.

Random Assignment Procedures and Baseline Equivalence

The procedures for teacher and student random assignment were designed to eliminate confounds related to both teacher and student characteristics. Random assignment proceeded in two steps. First, 463 classroom teachers were randomly assigned, within grade-by-school blocs, to either deliver math lessons to control students at the end of the school year or READS lessons to treatment students. Teachers randomly assigned to teach the math and READS lessons were balanced on race, gender, and the selectivity of the college they attended as an undergraduate; teachers assigned to teach the math classes were older on average. Second, within each homeroom class, treatment status for children was determined by randomly assigning each student to one of the teachers delivering READS lessons for their end-of-year lessons or to one of the control teachers teaching math lessons. Because student random assignment took place within homeroom class, students' teacher quality for both the prior school year and the end-of-year lessons was balanced for the treatment and control groups.

Table 1 presents means of several child characteristics measured at baseline, separately for the treatment and control groups, along with *p*-values from tests of the differences in means. In the left panel of Table 1, we report tests of baseline group equivalence for the full sample. As expected, random assignment generated balance in observables across the treatment and control groups on both demographic measures and baseline reading pretests. The right panel of Table 1 also reports balance tests for the analysis sample of 5,569 children who completed the reading posttests and pretests. Because all children in North Carolina were required to take the end-of-grade reading test (EOG), the EOG test scores were available for any child in our study who remained in the state in spring 2014. For the analytic sample, there was also balance on the baseline measures, including the reading pretests. These balance tests suggest that the pretest

characteristics of treatment and control children in the analytic sample were statistically equivalent.

Approximately 9% of the children (n = 573) who began the study did not complete the spring 2014 EOG reading test. Attrition rates were equivalent for the treatment (8.4%) and control groups (8.7%), χ^2=.15, p=.69. Among the 573 children with missing EOG reading posttests, the mean standardized spring ITBS score for the treatment group (M = -.57, SD = 1.06) was equivalent to the mean pretest reading score of the control group (M = -.59, SD = .96), p = .89. Children in the analytic sample were balanced on pretest characteristics, attrition rates were unrelated to condition, and the pretest reading scores of children with missing EOG reading scores were statistically equivalent across conditions. In summary, there were no major threats to internal validity.

Measures

Iowa Test of Basic Skills, Reading Comprehension (ITBS). At baseline in spring 2013, children were administered the reading comprehension section of the Iowa Test of Basic Skills. Children in second grade were administered Level 8, Form A and children in third grade were administered Level 9, Form A. The ITBS is a reliable assessment with reported KR-20 coefficients above .93 and equivalent form estimates of .86 or higher (Hoover et al., 2003). The ITBS pretest scores yielded both a developmental standard score (M = 176, SD = 22, Min = 114, Max = 260), national percentile ranks (M = 46, SD = 27) and Lexile scores (M = 529, SD = 172, Min = 25, Max = 1100). We standardized the developmental standard score to have a mean of 0 and a standard deviation of 1 within each grade and used this measure as the pretest covariate. We used the child Lexile score to facilitate our procedures for identifying books with Lexile levels that were within each child's independent reading level (Wright & Stenner, 2000).

North Carolina End-of-Grade (EOG) reading comprehension. Nine months after the intervention in spring 2014, children were administered the North Carolina End-of-Grade (EOG) reading comprehension test, a state-required test. The third- and fourth-grade EOG reading tests included 52 multiple-choice items that are scaled based on a 3-parameter IRT (item-response theory) model. The test was administered over three 60 minute blocks of time. Reported KR-20 reliabilities ranged from .88 to .92 (North Carolina Department of Public Instruction, 2014). To facilitate interpretation of the coefficients as effect sizes, we standardized the EOG scaled scores to have a mean of 0 and a standard deviation of 1 within each grade.

Amount of home-based summer reading routines with READS books. The trifold served as proxy for the amount of home-based summer reading routines involving a combination of child-initiated and parent-supported book reading. Each READS book came with a trifold that included three multiple-choice comprehension questions about the book. Figure 2 and 3 provides sample trifolds (see supplementary online materials). We created three measures that served as a proxy for the amount of home-based summer book reading routines: (1) the number of trifolds returned; (2) a binary indicator, high or low number of trifolds returned, where "high" and "low" were defined as above or below the median number of trifolds returned; and (3) the number of trifolds returned with at least one comprehension question answered correctly. Overall, the 3,174 treatment group children returned 11,497 trifolds ($M = 3.62$, $Mdn = 3.00$, $SD = 3.60$, $Min = 0$, $Max = 10$) and answered an average of 7.5 total questions correctly ($M = 7.50$, $Mdn = 6.00$, $SD = 7.92$, $Min = 0$, $Max = 30$). Since children returned 3.62 trifolds on average and there were three questions on each trifold, the overall percentage of comprehension questions answered correctly was 70 percent (7.5 / (3.62 * 3)).

Amount of summer book reading. All children in the study completed a survey when they returned to school in the fall. To measure reading amount during summer vacation, we used an item and response options that have been included in the National Assessment of Education Progress, prior studies on reading habits outside school, and our previous experiment studies (White et al., 2014; Cunningham & Stanovich, 1991; Donahue et al., 2001). The survey included an item that asked them to report the total number of books they had read independently during the summer: "During summer vacation, about how many books did you read?" with answer choices "0-1," "2-3," up to "20 or more." We assigned students the mid-point value for their answer choice (e.g. students who reported having read 2-3 books were given a value of 2.5), and students who selected "20 or more" were given a value of 20. For both the treatment group and the control group, the responses were used to create two measures of the amount of summer book reading: (1) the self-reported number of books read, and (2) a binary indicator of high or low number of self-reported books read, where "high" and "low" were above or below the self-reported number of books read.

Procedures for Implementing End-of-Year Lessons and Afterschool Family Literacy Event

Matched and interesting books and lesson books. We consulted with three elementary school teachers in North Carolina to identify 519 books (366 narrative, 153 informational) for this study. Narrative text structures (Stein & Glenn, 1979) included characters, settings, plots, problems, resolutions, and themes. The majority of the informational texts included books about animals, natural science topics, or biographies.

Similar to our three previous studies, we used a computer algorithm to identify books that were matched to each student's reading skill level and reading interests. Each child's reading skill levels were measured using the spring baseline reading comprehension test (ITBS), which

were translated into Lexiles, a proprietary system designed to align reading skills with the difficulty of children's books. Lexiles are vertically-scaled measures that place each child's reading ability and text difficulty on a common scale to facilitate efforts to find appropriately challenging books with reading levels that are within a child's independent reading level (Wright & Stenner, 2000). A linking study generated a conversion table that allowed each student's ITBS comprehension score to be reported as a Lexile (Metametrics, 2011). For this study, children were also asked questions about the types of books they would like to read by completing a reading preference survey. Using an algorithm, eight books were then chosen that best match each student's interest among those that are the appropriate difficulty given the student's baseline reading skill level. In addition to the eight matched books, children received two lesson books for a total of 10 summer books.

The end-of-year lesson books were then selected by the consultant teachers, using several criteria. The books were high-interest, not commonly read by children, had appropriate text difficulty, and sufficiently rich and complex for content-based comprehension lessons. In the third-grade lessons, teachers selected a three-chapter narrative text, *Look Out, Jeremy Bean!* by Alice Schertle and the three informational texts were *A Picture Book of Helen Keller*, by David A. Adler; *Tornado Alert*, by Franklyn M. Branley; and *Owls*, by Gail Gibbons. In the second-grade lessons, teachers selected a three-chapter narrative text *Frog and Friends: Party at the Pond*, by Eve Bunting, and three lessons using the informational texts *Wilma Rudolph*, by Victoria Sherrow*; The Big Dipper*, by Franklyn M. Branley; and *National Geographic Readers: Frogs!*, by Elizabeth Carney. In addition to the eight matched books, children also received two lesson books (the fiction book and one of the non-fiction books) in the mail over the summer.

Teacher training. Teachers in the treatment group participated in a two-hour afterschool training, and training was led by lead teachers who were trained by the research staff. During the training, teachers received a lesson plan for each of the six lessons and accompanying materials. The trainers began by explaining the goals and purpose of the lessons, walked teachers through the lesson procedures, modeled the lessons, answered questions, and offered the teachers an opportunity to ask follow-up questions via email or phone prior to and during the week when the lessons were scheduled. Trainers emphasized that the goal of the training was to help teachers adhere to the lesson scripts. Teachers were also instructed to complete lesson logs after each of the six lessons. Participating teachers received a $300 stipend.

End-of-year lessons. Using recommendations from an Institute of Education Sciences (IES) research-based practice guide for improving K-3 reading comprehension (Shanahan et al., 2010), we developed lessons in which teachers (a) modeled how to use a before, during, and after reading comprehension routine tailored to narrative or expository texts, (b) asked literal and inferential questions to promote discussion during read-alouds, (c) taught children to apply these activities to multiple genres of texts, and (d) created a motivational context that promoted children's engagement with lesson books. We used an iterative procedure to develop the lessons and incorporated recommendations from the IES K-3 practice guide, a panel of five literacy experts, and feedback from third-grade teachers who pilot-tested the lessons in the year before full implementation.

The content-based lessons incorporated different comprehension routines (Duke & Pearson, 2002) for the narrative and informational texts. In the first three lessons with the narrative text, *Look Out, Jeremy Bean!*, the routine was based on a "story impression" activity (McGinley & Denner, 1987). The story impression activity was an instructional tool that

prompted children to make a story prediction using content drawn from narrative text, to read the

story with their teachers, and then to check the prediction after reading the text. The lesson

materials included key words selected from the story and listed in the order they appeared. In the

pre-reading activity, the teacher read aloud the story impression words and phrases and prompted

students to use the words and phrases to make a story prediction based on common text

structures such as character, setting, goal, problem, plot or action, resolution, or theme. The

words and phrases were selected to direct children's attention to such text structures and to frame

their story predictions within these text structures. In the during-reading activity, the teacher

read the story aloud and asked literal and inferential text-based questions that helped students

activate and integrate text-based ideas. In the post-reading activity, the teacher led a discussion

comparing the story prediction with the actual content of the story. Thus the teacher provided

scaffolding for the content routine by helping students perform each part of the routine prior to

summer vacation (Graves & Graves, 2003). Over the course of the three lessons, the teacher

gradually released responsibility for the story impressions activity to students.

In the next three lessons with informational text, the comprehension routine was based on

an "information impressions" activity. The lessons were designed to help teachers use one

expository text structure (description) to help children understand the author's purpose, to

organize ideas in the order of their importance, and to facilitate children's recall and memory of

informational books (Kintsch, 1994; Meyer, 1987). Teachers explained that informational texts

communicate information about the social and natural world, and famous people (Duke &

Billman, 2009), that texts have central and non-central ideas, and that the most important facts

are called "main ideas." Teachers then defined the three types of informational books that

children might read in the summer months and followed scripted language to introduce types of

main ideas that the author might describe. For example, in the case of biographies, teachers explained: "In biographies, authors might describe main ideas such as: what this person's childhood was like; his or her background (for example, where they lived, what their family was like, etc.); why this person is famous."

As a parallel activity to the story prediction with narrative texts, teachers then implemented a before reading activity in which children viewed central ideas, or information impressions about a book, used the central ideas to predict the main idea being described by the author, and then made a "main idea" prediction. In the during reading activity, teachers read aloud from the text, asked literal and inferential questions, and concluded with a discussion on whether children's main idea prediction was accurate. As with the first three lessons, over the course of the three lessons, the teacher gradually released responsibility for the information impression activity to students. To support children's use of the comprehension routines at home without teacher assistance, teachers asked children to complete a trifold as part of a homework assignment given during the lessons. Details on the fidelity of lesson implementation are in the online supplementary materials.

Teachers in the control condition taught six math lessons. The math lessons were developed by a professor of mathematics education and were based on a unit using a problem-based learning framework. Children worked through a unit on the Zoo Improvement Plan, and developed a report for improving the North Carolina Zoo. The unit covered mathematics content and standards in the North Carolina State Standards.

Afterschool family literacy event. To help parents understand the goals of the READS program and the procedures in the READS Reading Routine, we implemented an afterschool family literacy event. The event was designed to strengthen the school and home microsystems

by providing information to parents about the comprehension routines that children were encouraged to use with the matched books that were mailed home. At the school-based event, children and their parents had another opportunity to use the content-based prediction routine prior to receiving matched books in the mail during the summer months.

School principals and teachers were asked to hold family nights after teachers had begun implementing classroom lessons, but before the end of the school year. The family night included two primary components. First, facilitators hired and trained by the research team informed parents about key components of the summer reading intervention. Specifically, parents learned that their children would receive books in the mail during the summer and they were expected to complete before- and after-reading activities on a trifold.

Second, facilitators asked parents to be involved with their child's home reading activities in four ways, as expressed by the acronym EATS: (a) Encourage your child to read their matched books received in the mail; (b) Ask whether they completed the trifold; (c) Talk with your child about the books; and (d) Send back the trifolds after your child reads the books and answers the comprehension questions. Children also recited a reading pledge and received medals. To increase attendance at family night, participating families were offered a ticket for an opportunity to win a $50 gift card. Forty percent of the children in the treatment group (n = 1,270) attended the family night event.

To ensure that participating parents understood the main goals of the program, we also implemented summer phone calls with children who were not participating in the program, as measured by the trifold data. In the summer months, we called parents whose children had not returned at least one trifold by July 1 or August 1. Eighty-one percent of parents whose children did not attend an afterschool family night were called once during the summer. Thus, most

parents received information about the program goals and the importance of having their

children return the tri-folds, either through the afterschool family literacy event or the summer

phone calls.

Analytic Models

We conducted an intent-to-treat analysis to address our first research question. The

intent-to-treat effects were estimated by comparing the posttest reading scores of children who

were randomly assigned to treatment or the control condition in spring 2013 (i.e., at the end of

their second- or third-grade year). A year later in spring 2014, children were administered the

North Carolina EOG test, at the end of their third- or fourth-grade year. We included the pretest

reading score as a control variable to improve the precision of the estimated treatment effect.

The models also accounted for the study design by including fixed-effects for the homeroom

class randomization blocs from which children were randomly assigned to conditions. With 59

schools and 5,569 children in the analytic sample, and pretest covariates that explain 62% of the

variation in posttest scores, the minimum detectable effect size (MDES) for this study was

approximately .05 with 80% power and an alpha level of .05 (two-tailed test).

In the standard intent-to-treat (ITT) analysis, we estimated the following model

(1) $y_i^{Post} = \alpha_0 + \alpha_1 R_i + \alpha_2 y_i^{Sp2013} + \lambda_c + \varepsilon_i$

where y_i^{Post} is the post-random assignment outcome measured in the spring of the school year, 9

months after the intervention ended, R_i is an indicator for being randomly assigned to the

treatment, y_i^{Sp2013} is a baseline reading comprehension test score, λ_c is a set of spring 2013

homeroom classroom fixed effects, α_0, α_1, and α_2 are parameters to be estimated, and ε_i is a

random error term. For all estimates reported in the paper, we clustered the standard errors by

school to account for within-school correlation in the error term ε_i. The parameter of interest, α_1

, measures the difference in average outcomes between students randomly assigned to the treatment and students in the control group.

We refer to the estimate of α_1 as the intent-to-treat (ITT) estimate because R_i indicates the actual random assignment rather than participation. However, since all students in the study received parent consent prior to random assignment, virtually every student who was randomly assigned to receive the treatment program did so. Thus, in practice our estimates of α_1 were very close to measures of the effect of participation, if participation was defined as receiving the spring lessons in the classroom, an invitation to an afterschool family literacy event and being mailed 10 books over the summer.

Assignment to the treatment condition was random within classrooms. Random assignment ensured that in large samples R_i was uncorrelated with ε_i (or equivalently with potential outcomes for y_i^{Post}), and as a result we can interpret estimates of α_1 as the causal effect of assignment to the treatment on y_i^{Post}. Because random assignment was conditional on classroom, we included classroom fixed effects in all models.[1] Inclusion of the baseline reading test as a control was not necessary for internal validity, but it improved precision because it explained a large portion of the residual variance.

In the instrumental variables analysis, we took advantage of the fact that there were essentially 463 different experiments, one in each second- and third-grade spring classroom. It would have been possible to estimate 463 treatment effects, each based on random assignment

---

[1] In practice, the probability of being assigned to the treatment group was close to 0.5 in each classroom. Since the treatment assignment probability was not correlated with classroom, excluding the classroom fixed effects should not cause any omitted variables bias, and they are therefore not necessary. Within each classroom, randomization was also stratified by whether students were limited English proficient (LEP) and whether they were eligible for free or reduced lunch (FRL). Within each classroom, the probability of assignment to the treatment was the same in the FRL/LEP and non-FRL/LEP strata, so in practice this level of stratification is ignorable. We report results from models that include classroom fixed effects because the subsequent analysis is at the classroom level, and models that include classroom, classroom-by-LEP/FRL status or no fixed effects are all unbiased. The results from all three sets of models yield the same qualitative findings.

within a specific classroom. Presumably, students responded to the treatment in different ways in different classrooms. Students may have been induced to participate in more of the home-based summer book reading routines that were stimulated by the treatment in some classrooms than others, for example. We wanted to know if in the classrooms where students participated in more home-based summer reading routines with the READS books, there were also larger treatment effects on reading test scores. The logic of this analysis followed directly from Kling, Liebman and Katz (2007). We used classroom-by-treatment fixed effects as instruments for potential mediators of the relationship between assignment to the treatment and test scores. This instrumental variables (IV) model took advantage of the random assignment and essentially asked whether delayed effects on reading test score effects were significantly larger in classrooms where treatment assignment caused the potential mediator to increase more relative to the control group.

The first stage model is

(2) $M_i = \pi_c(R_i \times C_i) + \pi_y y_i^{Sp2013} + \delta_c + \mu_i$

where M_i is the potential mediator, $(R_i \times C_i)$, the excluded instruments, are a set of interactions between the treatment assignment indicator and a full set of classroom indicators, π_c is a vector of classroom-specific first stage treatment effects, δ_c is a set of classroom fixed effects, and u_i is an error term.

The second stage model is

(3) $y_i^{Post} = \beta_0 + \beta_1 M_i + \beta_y y_i^{Sp2013} + \lambda_c + \varepsilon_i$

Classroom fixed effects were included in both the first and second stage models so that the variation identifying the effect of the mediator on end-of-year reading scores was treatment-control comparisons within classrooms. The identification of these mediator parameters is

vulnerable to the possibility that treatment effects on the mediators are correlated with an

unobservable cause of the test score treatment effects, what might be referred to as "omitted

mediator bias."[2] However, relative to the common alternative of adding controls for potential

mediators to the ITT test score model, the assumptions of the multi-site IV method are much

weaker. Whereas the commonly used approach is invalid if the level of unobservable

determinants of test scores are correlated with treatment effects across sites, the multi-site IV

relies on the less restrictive assumption that no treatment-control differences in unobservables

exist that are correlated with test score treatment effects across sites. Since treatment was

randomly assigned at each site (in our case classrooms), baseline unobservables should have

been balanced across treatment and control observations within sites. Thus, the identifying

assumption is that variation in treatment effects on unobserved *mediators* were not correlated

with treatment effects on the included mediator.[3]

RESULTS

Effects on the Amount of Independent Summer Book Reading

We used data from fall surveys to assess the amount of children's self-reported

independent summer book reading activities. Students in the control group reported reading 9.8

books during the summer on average, while students in the treatment group reported reading 10.9

books. The difference was statistically significant, suggesting that assignment to participate in

the treatment induced students to read an additional 1.1 books over the summer. In high poverty

schools, children in the treatment group ($M = 10.85$, $SD = 9.6$) reported reading approximately

1.25 more books than children in the control group ($M = 9.6$, $SD = 6.5$). In moderate poverty

schools, children in the treatment group ($M = 11.08$, $SD = 5.99$) reported reading approximately

---

[2] We thank Jens Ludwig for suggesting this term.
[3] Additional identifying assumptions apply, including the assumption that treatment effects and treatment compliance do not covary; see Reardon & Raudenbush (2013).

.68 more books than children in the control group ($M = 10.4$, $SD = 6.73$). These results suggest that treatment effects on the self-report measure of independent summer book reading were nearly twice as large in high poverty than in moderate poverty schools.

Intent-to-Treat Estimates on a Delayed Measure of Reading Comprehension

Table 2 presents ITT estimates of the treatment effect on EOG scores for the full analytic sample and separately for children in high poverty and moderate poverty schools. In the full sample, the treatment improved reading comprehension scores by 4% of a standard deviation ($d = .04$). Descriptively, the results in models 2 and 3 suggest that the magnitude of the effect size of $d = .05$ in high poverty schools (75-100% free lunch) was larger than the effect size of $d = .01$ in moderate poverty schools. We also fit a fourth model including a treatment by school poverty interaction. In this model the point estimates for high and moderate poverty schools were not reliably different from each other.

Instrumental Variables (IV) Estimates of the Effects of Participation in Home-Based Summer Book Reading Routines

Table 3 presents the estimates of mediator effects from the IV models for the combined sample.[4] Each row of the table shows results for a different potential mediator. Rows 1 to 3 show the results using the measures based on trifolds returned. A model based on a linear count of trifold returns is shown in row 1, and a model based on a median split indicator for having returned more than 3 trifolds is shown in row 2. In row 3, a model based on the number of trifolds returned in which children correctly answered at least one comprehension item is presented. Virtually all control students returned zero tri-folds, so the first stage estimates were

---

[4] In our previous study of short-term treatment effects (Guryan et al., 2014), we found positive intent-to-treat effects in reading in third-grade and thus we restricted the IV analyses to third-grade. However, since there was a positive main effect for the full sample of second- and third-graders on the delayed reading outcome, we examined the mechanism underlying this treatment effect for the full sample in the IVE analyses reported in table 3.

essentially the average number of trifolds returned by the treatment students in the classroom.[5]

The analysis therefore asked whether treatment effects on test scores were higher in classrooms

where the treatment increased the amount of home-based summer reading routines with the

READS books that were mailed home during the summer months.

The results suggest that only the measures of home-based summer book reading routines,

based on the trifold data, were associated with larger effects on the North Carolina EOG reading

tests, particularly for children in high poverty schools. In the specification with the linear

measure of trifold returns shown in row 1 and 3, the point estimate for children in the full sample

was 0.01, which indicates that in a classroom where the students read all 10 books that were

mailed home, the children in the treatment group scored approximately .10 standard deviations

higher in reading scores. In other words, these results suggest that, using the trifold returns as a

measure of the extent to which children participated in summer reading routines with their

READS books, the READS intervention improved reading comprehension skills for children in

classrooms where the students participated in the program as intended. This result suggests that

reading the READS books was a mediator of the treatment effect, though it is possible that there

was something else about the classrooms where the treatment students read more of the READS

books that explained the larger treatment effects on posttest reading scores in high poverty

schools.

Rows 4 and 5 show estimates in which the measures of self-reported amount of

independent summer book reading were treated as a potential mediator. Row 4 presents results

from a specification in which the self-reported number of books was the potential mediator; row

---

[5] There were a very small number of control students who lived in the same household as treatment students and who returned tri-folds that were mailed to the treatment student with whom they were living. This control group crossover would presumably tend to reduce any measured treatment effects since some control students received the treatment. Since this represents a very small portion of the sample, we suspect any understatement of treatment effects is small.

5 shows results for a median split of self-reported book reading, a binary indicator for having

read more than 10.5 books. A positive effect would indicate that classrooms where mailing

books caused a larger increase in summer book reading also had larger treatment effects on

reading comprehension skills. In both specifications using measures of the self-reported amount

of summer book reading, the point estimates for the mediator variables were essentially zero for

the full sample and by school poverty.

Sensitivity Analyses

The sensitivity analyses also indicated that treatment effects were robust among different

subgroups of students and different model specification (See table 4 in the supplementary online

materials).

DISCUSSION

In response to the national challenge of improving reading comprehension outcomes for

children in high poverty schools, we incorporated research-based components of both school-

and home-based summer programs in developing a large-scale and low-cost summer reading

intervention (READS). The intent-to-treat analyses revealed a positive effect on the North

Carolina End-of-Grade (EOG) reading test. Children in the treatment group scored 4 percent of

a standard deviation higher, on average, than did children in the control group on a delayed

posttest measure of reading comprehension. The effect size was also larger in high poverty

schools ($d = .05$) than moderate poverty schools ($d = .01$).

The magnitude of the intent-to-treat effects on the EOG was comparable in size to typical

effect sizes from elementary school interventions that target improvement on standardized tests

of reading comprehension. For example, Lipsey et al. (2012) reported a median effect size of .07

($SD = .27$; $N = 89$) from randomized trials of elementary school interventions that assess broad

domains of knowledge like reading comprehension. More recently, an i3 scale-up evaluation of Success for All in 37 schools yielded an effect size of .03 SD on the Woodcock-Johnson Passage Comprehension outcome (Quint et al., 2015). Thus, the magnitude of the effects observed in this large-scale randomized field trial of READS are comparable to impacts from literacy interventions that also target improvement in reading comprehension, particularly in high-poverty elementary schools.

Our most recent experiment (Guryan et al., 2014) found measured impacts on the ITBS reading comprehension test administered one month after the intervention, but only for the sub-sample of third-grade girls ($d = .07$). In contrast, this study showed a significant overall treatment effect combining boys and girls across both second- and third-grade on the North Carolina EOG reading comprehension test that was given nine months after the intervention when the participating students were completing third or fourth grade. The different results for the immediate and delayed posttests did not seem to result from differences in the samples. The sample size for the delayed posttest was larger by about 250 students because all students were required to take the state test: however, the Minimum Detectable Effect Sizes were virtually the same for the ITBS and North Carolina EOG. There are, however, other plausible explanations.

Since it is well-known that reading comprehension depends partly on oral reading fluency (Fuchs, Fuchs, Hosp, & Jenkins, 2001), one possible explanation for the stronger delayed effects is that the youngest children in the sample, who were second graders when they received the lessons, needed more time to develop reading fluency. However, our sensitivity analyses (see supplementary online materials) revealed a non-significant treatment effect on the third-grade fluency outcomes and helped to rule out improvement in children's ability to read connected text with speed and accuracy as the mechanism that led to better reading comprehension outcomes.

Thus, a more plausible explanation is that the EOG and ITBS may capture different component skills that underlie reading comprehension (Cutting & Scarborough, 2006; Keenan, 2012). Future research should examine whether the home-based summer book reading routines instantiated by the READS program are more likely to improve the component skills measured in reading comprehension tests like the EOG than ITBS. It is also important to note that the North Carolina EOG is untimed and requires about three hours to complete whereas the ITBS has time limits and requires about one hour to complete. Moreover, the EOG is more closely aligned with the Common Core State Standards for Literature than the ITBS, and emphasizes children's ability to comprehend narrative and informational text, and to answer literal and inferential questions (e.g., identify the main idea). Because the READS intervention encouraged children to participate in more home-based summer book reading routines with a variety of narrative and informational texts, more research is needed to understand whether gains in reading stamina (Hiebert, 2015) and higher-order comprehension processes (McNamara, Kintsch, Songer, & Kintsch, 1996; Rapp, van den Broek , McMaster, Kendeou & Espin, 2007) promoted transfer to the EOG comprehension test.

Results from this study indicate that children in high poverty schools in particular may benefit from participating in home-based summer reading routines and thus enjoy durable gains in reading comprehension. More broadly, our results are consistent with Raudenbush and Eschmann's (2015) hypothesis that low-SES children are likely to benefit more from efforts to improve educational and literacy learning opportunities outside school compared to their more affluent peers. The hypothesis was motivated, in part, by research on the counterfactual experiences of elementary school-aged children. For example, Lareau's (2011) ethnographic research suggests that higher SES families invest in the "concerted cultivation" of their

children's cognitive and social skills through a range of enrichment opportunities (e.g.,

academics, sports, music) whereas lower SES families give children considerable freedom in

how they spend their time outside school. Put simply, the counterfactual home literacy

experiences of elementary school children vary dramatically by family income (Entwisle, et al.,

2000). READS may have created a stronger treatment–control contrast in the out-of-school

literacy experiences and ultimately the reading comprehension outcomes of children in high

poverty than moderate poverty schools.

Practical Significance

Descriptively, the intention-to-treat effect of READS represents a 1.8 percentage point

increase overall in the number of students scoring on grade level by EOG proficiency standards

(from 44.7% proficient to 46.5%), or 2.39 percentage point increase in high poverty schools.

Future research should explore whether a multi-year implementation of READS could produce

cumulative effects over time.

An important practical motivation for this study was to assess the effects, costs, and cost-

effectiveness of READS relative to other literacy interventions that target high poverty schools

and struggling readers. Table 5 in the online supplementary materials provides comparative

information on average intent-to-treat effects, costs per child, and cost-effectiveness ratios for

READS and tutoring interventions that target high poverty schools. To compare the cost

effectiveness of READS with other literacy interventions, we compared the effect sizes per

$1,000 cost per-pupil. This metric is useful for evaluating how a fixed amount of money might

be spent since lower cost programs could be delivered to more students. We estimated that the

cost to provide the intervention described in this study was between $250 and $480 per student,

including the cost of the pretest survey and test, books, mailings, stipends to teachers, and

salaries of staff who managed the implementation (Levin & McEwan, 2001). The second to fourth columns of Table 5 show effects, costs, and cost-effectiveness ratios for supplemental educational services (SES) tutoring provision (assuming a 10-to-1 and 2-to-1 pupil-tutor ratio) funded by federal Title I dollars (Heinrich et al., 2014) and a recent randomized controlled trial of Reading Partners, a tutoring program involving volunteer tutors (Jacob et al., 2015). READS was generally more cost-effective than other out-of-school interventions like SES tutoring; furthermore, the cost-effective ratio of READS and Reading Partners was comparable.

Limitations

A number of assumptions underlie the use of IV analyses with single and multiple mediators, and many of them cannot be empirically tested and must be justified by strong theory (Reardon & Raudenbush, 2013). Thus, the robustness of these results needs to be tested with other analytic strategies and research designs (Frangakis & Rubin, 2002; Ludwig, Kling, & Mullainathan, 2011).

Relatedly, while our key "trifold returns" mediator was intended to capture the extent to which children participated in home-based summer book reading routines, we could not isolate the extent to which children read their books independently or the type of parental assistance that supported home book reading. In the READS afterschool family literacy event, parents learned to encourage their children to read their summer books and to complete the comprehension routines in the trifolds. These types of informal parent literacy activities may increase children's exposure to books (Senechal & Young, 2008), and enhance future literacy skills (Mol & Bus, 2011). However, more research is needed to understand the specific types of parental support and home-based book reading routines that foster children's reading skills.

Conclusion

Large-scale and low-cost reading interventions may be particularly important in an era when researchers and policymakers are seeking to scale-up evidence-based literacy reforms in high poverty elementary schools. The results from this study provide the best evidence to date that a large-scale and low-cost summer reading intervention can improve reading comprehension outcomes for children in the most economically disadvantaged schools.

References

Alexander, K. L., Entwisle, D. R., & Olson, L. S. (2007). Lasting consequences of the summer
    learning gap. *American Sociological Review, 72*, 167–180. doi:
    10.1177/000312240707200202

Allington, R.L., McGill-Franzen, A., Camilli, G., Williams, L., Graff, J., Zeig, J., & Nowak, R.
    (2010). Addressing summer reading setback among economically disadvantaged
    elementary students. *Reading Psychology*, 31(5), 411–427. doi:
    10.1080/02702711.2010.505165

Bronfenbrenner, U., & Morris, P. A. (2006). The bioecological model of human development. In
    W. Damon & R. M. Lerner (Eds.), *Handbook of child psychology: Theoretical models of*
    human development (5th ed., Vol. 1, pp. 795-825). Hoboken, NJ: John Wiley & Sons Inc.

Chaplin, D., & Capizzano, J. (2006). Impacts of a summer learning program: A random
    assignment study of Building Educated Leaders for Life (BELL).

Connor, C. M., Morrison, F. J., & Katch, L. E. (2004). Beyond the reading wars: Exploring the
    effect of child-instruction interactions on growth in early reading. *Scientific Studies of*
    Reading, 8, 305-336. doi: 10.1207/s1532799xssr0804_1

Cunningham, A. E., & Stanovich, K. E. (1991). Tracking the unique effects of print exposure in
    children: Association with vocabulary, general knowledge, and spelling. *Journal of*
    Educational Psychology, 83, 264–274. doi: 10.1037/0022-0663.83.2.264

Cutting, L. E. & Scarborough, H. S. (2006) Prediction of reading comprehension: Relative
    contributions of word recognition, language proficiency, and other cognitive skills can
    depend on how comprehension is measured. *Scientific Studies of Reading, 10*, 277-299.
    doi: 10.1207/s1532799xssr1003_5

Donahue, P. L., Finnegan, R. J., Lutkus, A. D., Allen, N. L., & Campbell, J. R. (2001). *The*

nation's report card: Fourth-grade reading 2000 (NCES 2001–499). Washington, DC:

U.S. Department of Education, Office of Educational Research and Improvement,

National Center for Education Statistics.

Downey, D.B., von Hippel, P.T., & Broh, B.A. (2004). Are schools the great equalizer?

Cognitive inequality during the summer months and the school year. *American*

Sociological Review, *69*, 613–635.

Duke, N. K., & Billman, A. K. (2009). Informational text difficulty for beginning readers. In

E. H. Hiebert & M. Sailors (Eds.), *Finding the right texts, what works for beginning and*

struggling readers (pp. 109-128). New York: Guilford Press.

Duke, N., & Pearson, P. D. (2002). Effective practices for developing reading comprehension. In

A.E. Farstrup & S.J. Samuels (Eds.), *What research has to say about reading instruction*

(3rd ed.; pp. 205–242). Newark, DE: International Reading Association.

Entwisle, D. R., Alexander, K. L., & Olson, L. S. (2000). Summer learning and home

environment. In R. D. Kahlenberg (Ed.), *A notion at risk* (pp. 9–30). New York: The

Century Foundation Press.

Frangakis, C. E., & Rubin, D. B. (2002). Principal stratification in causal inference. *Biometrics*,

58, 21–29. doi: 10.1111/j.0006-341X.2002.00021.x

Fredericks, J. A., Blumenfeld, P. C., & Paris, A. H. (2004). School engagement: Potential of the

concept, state of the evidence. *Review of Educational Research, 74*, 59 – 109. doi:

10.3102/00346543074001059

Fuchs, L. S., Fuchs, D., Hosp, M. K., & Jenkins, J. R. (2001). Oral reading fluency as an

indicator of reading competence: A theoretical, empirical, and historical analysis. *Scientific Studies of Reading, 5*, 239-256.

Good, R. J., Kaminski, R. A., Dewey, E. N., Wallin, J., Powell-Smith, K. A., & Latimer, R. J. (2013). *Dynamic Indicators of Basic Early Literacy Skills (DIBELS) Technical Manual*. Eugene, OR: Dynamic Measurement Group, Inc.

Graves, M., & Graves, B. (2003). *Scaffolding reading experiences: Designs for student success*. Norwood, MA: Christopher-Gordon.

Guryan, J., Hurst, E., & Kearney, M. (2008). Parental education and parental time with children. *Journal of Economic Perspectives, 22*, 23–46. doi: 10.1257/jep.22.3.23

Guryan, J., Kim, J. S., & Quinn, D. M. (2014). Does Reading During the Summer Build Reading Skills? Evidence from a Randomized Experiment in 463 Classrooms. National Bureau of Economic Research, Cambridge, MA. Working Paper 20689. http://www.nber.org/papers/w20689

Guthrie, J. T., & Humenick, N. M. (2004). Motivating students to read: Evidence for classroom practices that increase reading motivation and achievement. In P. McCardle & V. Chhabra (Eds.), *The voice of evidence in reading research* (pp. 329–354). Baltimore, MD: Brookes Publishing.

Guthrie, J. T., Wigfield, A., & You, W. (2012). Instructional contexts for engagement and achievement in reading. In S. L. Christenson, A. L. Reschly, & C. Wylie (Eds.), *Handbook of research on student engagement* (pp. 601-634). New York, NY: Springer.

Heinrich, C. J., Burch, P., Good, A., Acosta, R., Cheng, H., Dillender, M., Kirshbaum, C., Nisar

H., & Stewart, M. S. (2014). Improving the implementation and effectiveness of out-of-

school-time tutoring. *Journal of Policy Analysis and Management, 33*, 471–494. doi:

10.1002/pam.21745

Hiebert, E. H. (2015). *Teaching stamina and silent reading in the digital-global age*. Santa

Cruz: TextProject, Inc.

Hoover, H.D., Dunbar, S.B., Frisbie, D.A., Oberley, K.R., Ordman, V.L., Naylor, R.J., & …

Shannon, G. P. (2003). *The Iowa Tests of Basic Skills Forms A and B: Guide to research

and development*. Itasca, IL : Riverside.

Jacob, R. T., Armstrong, C. Willard, J. A., Bowden, A. B., & Pan, Y. (2015). *Mobilizing

volunteer tutors to improve student literacy*. New York: MDRC.

Keenan, J. M. (2012). Measure for measure: Challenges in assessing reading comprehension. In

J. P. Sabatini, E. Albro, & T. O'Reilly (Eds.), *Measuring up: Advances in how we assess

reading ability* (pp. 77-87). Lanham, MD: Rowman & Littlefield.

Kim, J. S. (2006). Effects of a voluntary summer reading intervention on reading achievement:

Results from a randomized field trial. *Educational Evaluation and Policy Analysis*, *28*(4),

335–355. doi: 10.3102/01623737028004335

Kim, J. S. (2007). The effects of a voluntary summer reading intervention on reading activities

and reading achievement. *Journal of Educational Psychology*, *99*(3), 505-515. doi:

10.1037/0022-0663.99.3.505

Kim, J. S., & Guryan, J. (2010). The efficacy of a voluntary summer book reading intervention

for low-income Latino children from language minority families. *Journal of Educational

Psychology, 102*(1), 20-31. doi: 10.1037/a0017270

Kim, J. S. & Quinn, D. M. (2013). The effects of summer reading on low-income children's

literacy achievement from kindergarten to grade 8: A meta-analysis of classroom and

home interventions. *Review of Educational Research, 83*(3), 386-431. doi:

10.3102/0034654313483906

Kim, J. S., & White, T. G. (2008). Scaffolding voluntary summer reading for children in grades 3

to 5: An experimental study. *Scientific Studies of Reading*, *12*(1), 1–23. doi:

10.1080/10888430701746849

Kintsch, W. (1994). Text comprehension, memory, and learning. *American Psychologist, 49*,

294-303. doi: 10.1037/0003-066X.49.4.294

Kling, J. R., Liebman, J. B., & Katz, L. F. (2007). Experimental analysis of neighborhood

effects. *Econometrica, 75*, 83–119

Lareau, A. (1989). *Home advantage: Social class and parental intervention in elementary

education*. New York: The Falmer Press.

Lareau, A. (2011). *Unequal childhoods: Class, race, and family life*. Berkeley: University of

California Press.

Levin, H. M., & McEwan, P. (2001). *Cost-effectiveness analysis: Methods and applications*.

Monterey Park, CA: Sage.

Lipsey, M.W., Puzio, K., Yun, C., Hebert, M.A., Steinka-Fry, K., Cole, M.W., Roberts, M.,

Anthony, K.S., Busick, M.D. (2012). Translating the statistical representation of the

effects of education interventions into more readily interpretable forms. (NCSER 2013-

3000). Washington, DC: National Center for Special Education Research, Institute of

Education Sciences, U.S. Department of Education.

Ludwig, J., Kling, J. R., & Mullainathan, S. (2011). Mechanism experiments and policy

Evaluations. *Journal of Economic Perspectives, 25*, 17–38. doi: 10.1257/jep.25.3.17

May, H., Gray, A., Gillespie, J.N., Sirinides, P., Sam, C., Goldsworthy, H., … & Tognatta, N. (2014). Evaluation of the i3 Scale-up of Reading Recovery: Year One Report, 2011-12. RR-76. Philadelphia, PA: Consortium for Policy Research in Education. Princeton, NJ: Mathematica Policy Research Reports.

McCoach, D.B., O'Connell, A.A., Reis, S.M., & Levitt, H.A. (2006). Growing readers: A hierarchical linear model of children's reading growth during the first 2 years of school. *Journal of Educational Psychology*, *98*, 14–28. doi: 10.1037/0022-0663.98.1.14

McCombs, J.S., Augustine, C.H., Schwartz, H.L., Bodily, S.J., McInnis, B., Lichter, D.S., & Cross, A.B. (2011). *Making summer count: How summer programs can boost children's learning.* Santa Monica, CA: RAND Corporation.

McCombs, J. S., Pane, J. F., Augustine, C. H., Schwartz, H. L., Martorell, P., & Zakaras, L. (2014). *Ready for fall? Near-term effects of voluntary summer learning programs on low-income students' learning opportunities and outcomes*. RR-815-WF. Santa Monica, CA: RAND Corporation.

McGinley, W.J., & Denner, P.R. (1987). Story impressions: A prereading/writing activity. *Journal of Reading*, *31*, 248–253.

McNamara, D. S., Kintsch, E., Songer, N. B., & Kintsch, W. (1996). Are good texts always better? Interactions of text coherence, background knowledge, and levels of understanding in learning from text. *Cognition and Instruction, 14,* 1–43. doi: 10.1207/s1532690xci1401_1

MetaMetrics (2011). *Linking the R-CMB and MAZE with the Lexile framework for reading.* Durham, NC: MetaMetrics, Inc.

Meyer, B. J. F. (1987). Following the author's top-level organization: An important skill for

reading comprehension. In R. J. Tierney, P. L. Anders, & J. Nichols Mitchell (Eds.),
Understanding readers' understanding: Theory and practice (pp. 59-76). Hillsdale, NJ:
Erlbaum.

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

North Carolina Department of Public Instruction. (2014). Technical Brief-EOG and EOC
Assessments. Downloaded from:
http://www.ncpublicschools.org/docs/accountability/testing/eoceogtechbrief14.pdf

Quint, J. S., Balu, R., DeLaurentis, M., Rappaport, S., Smith, T. J., Zhu, P., Alterman, E., &
Pramik, E. (2015). The Success for All Model of school reform interim findings from the
Investing in Innovation (i3) Scale-Up. New York, NY: MDRC.

Rapp, D. N., van den Broek, P., McMaster, K. L., Kendeou, P. & Espin, C. A. (2007). Higher-
order comprehension processes in struggling readers: A perspective for research and
intervention. *Scientific Studies of Reading, 11*, 289-312. doi: 0.1080/10888430701530417

Raudenbush, S. W., & Eschmann, R. D. (2015). Does schooling increase or reduce social
inequality? *Annual Review of Sociology, 41*, 1-28. doi: 10.1146/annurev-soc-071913-
043406

Reardon, S. F., & Raudenbush, S. W. (2013). Under what assumptions do site-by-treatment
Instruments identify average causal effects? *Sociological Methods and Research, 42*,
143–163. doi: 10.1177/0049124113494575

Senechal, M., & Young, L. (2008). The effect of family literacy interventions on children's
acquisition of reading from Kindergarten to Grade 3: A meta-analytic review.
Review of Educational Research, 78, 880-907.

Shanahan, T., Callison, K., Carriere, C., Duke, N.K., Pearson, P.D., Schatschneider, C., &
    Torgesen, J. (2010). *Improving reading comprehension in kindergarten through 3rd*
    grade: A practice guide (NCEE 2010–4038). Washington, DC: Institute of Education
    Sciences, U.S. Department of Education.

Stein, N. L., & Glenn, C. G. (1979). An analysis of story comprehension in elementary school
    children. In R. D. Freedle (Ed.), *Advances in discourse processes: Vol. 2. New*
    directions in discourse processing (pp. 53-120). Norwood, NJ: Ablex.

Tindal, G. (2005). Oral reading fluency, 90 years of measurement. *Behavioral Research and*
    Teaching. University of Oregon.

Tseng, V., & Seidman, E. (2007). A systems framework for understanding social settings.
    American Journal of Community Psychology, 39, 217–228. doi: 10.1007/s10464-007-
    9101-8

U.S. Department of Education (2012). Reading 2011: National Assessment of Educational
    Progress at Grades 4 and 8, Washington, DC: U. S. Department of Education: National
    Center for Education Statistics, Institute of Education Sciences.

White, T. G., Kim, J. S., Kingston, H. C., & Foster, L. F. (2014). Replicating the effects of a
    teacher-scaffolded voluntary summer reading program: The role of poverty." *Reading*
    Research Quarterly, 49, 5-30.

Wigfield, A., Guthrie, J.T., Perencevich, K. C., Taboada, A., Klauda, S.L., McRae, A., &
    Barbosa, P. (2008). Role of reading engagement in mediating the effects of reading
    comprehension instruction on reading outcomes. *Psychology in the Schools*, *45,* 432-445.
    doi: 10.1002/pits.20307

Williams, J. P., Hall, K. M., Lauer, K. D., Stafford, B., DeSisto, L. A., & deCani, J. S. (2005).

Expository text comprehension in the primary grade classroom. *Journal of Educational*

Psychology, 97, 538-550. doi:10.1037/0022-0663.97.4.538

Wright, B. D., & Stenner, A. J. (2000). Lexile perspectives. *Popular Measurement, 3*, 14-17.

Figure 1. READS Program theory for improving reading comprehension by fostering children's participation in home-based summer book reading routines with trifolds
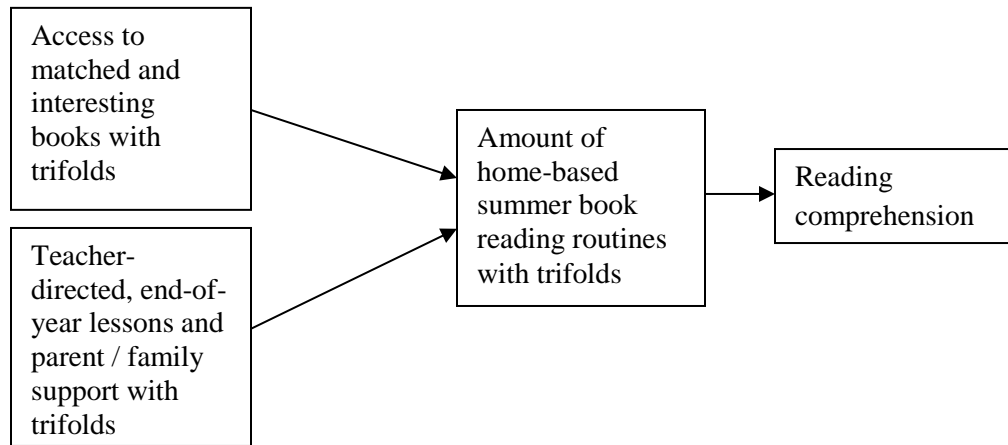
Table 1. Characteristics of children in the treatment and control condition for the full sample and analysis sample

Variable	Full Sample					Analysis Sample				
	Grades 2 and 3									
	Control	READS	*sd*	*p*-value	*n*	Control	READS	*sd*	*p*-value	*n*
% Female	51.79	50.54		0.44	6383	53.82	51.18		0.12	5569
% FRL	77.55	77.34		0.61	6326	77.28	76.83		0.39	5531
% LEP	16.62	16.37		0.95	6304	16.64	16.09		0.73	5529
% Hispanic	21.33	21.33		0.92	6287	21.55	21.60		0.82	5525
% Black	40.03	39.14		0.43	6287	39.46	38.14		0.25	5525
% White	22.63	23.15		0.65	6287	22.48	23.39		0.56	5525
% Other Race	16.01	16.38		0.64	6287	16.51	16.87		0.39	5525
Spr. 2013 ITBS Comp. Z-Score	0.00	0.01	1.01	0.71	6088	0.05	0.07	0.99	0.79	5569
Spr. 2013 ITBS Comp Std. Score	174.31	174.77	22.11	0.69	6088	175.49	175.86	21.80	0.75	5569
# Trifolds Returned Summer 2013	0.04	3.62	3.13	0.00	6383	0.04	3.81	3.20	0.00	5569
Total # Comp. Qs Correct across Tri-folds	0.07	7.50	6.74	0.00	6383	0.07	7.94	6.91	0.00	5569
Spring 2014 EOG Reading Score	439.07	439.44	10.65	0.16	5810	439.17	439.67	10.60	0.12	5569
Spring 2014 EOG Reading Z-Score	-0.02	0.02	1.00	0.16	5810	-0.01	0.04	1.00	0.12	5569
Spring 2014 EOG Reading % Proficient	44.58	45.89		0.36	5810	44.72	46.51		0.27	5569
	Grade 2									
% Female	50.43	50.18		0.98	3433	52.10	50.54		0.48	2971
% FRL	77.87	76.78		0.25	3401	77.45	76.30		0.16	2950
% LEP	17.49	16.23		0.32	3392	17.35	15.67		0.25	2948
% Hispanic	22.51	23.04		0.66	3381	23.08	23.08		0.82	2946
% Black	39.68	38.69		0.54	3381	38.69	37.60		0.38	2946
% White	22.04	23.16		0.47	3381	21.87	23.84		0.30	2946
% Other Race	15.77	15.10		0.39	3381	16.35	15.48		0.60	2946
Spr. 2013 ITBS Comp. Z-Score	0.00	-0.01	1.00	0.95	3255	0.06	0.05	0.98	0.88	2971

Spr. 2013 ITBS Comp Std. Score	167.60	167.46	18.84	0.95	3255	168.66	168.48	18.56	0.88	2971
# Trifolds Returned Summer 2013	0.03	3.74	3.19	0.00	3433	0.03	3.90	3.24	0.00	2971
Total # Comp. Qs Correct across Tri-folds	0.05	7.40	6.66	0.00	3433	0.05	7.79	6.81	0.00	2971
Spring 2014 EOG Reading Score	436.63	436.88	10.58	0.34	3113	436.78	437.06	10.54	0.35	2971
Spring 2014 EOG Reading Z-Score	-0.01	0.01	1.00	0.34	3113	0.00	0.03	1.00	0.35	2971
Spring 2014 EOG Reading % Proficient	46.70	47.43		0.66	3113	47.03	48.17		0.54	2971
Grade 3										
% Female	53.38	50.95		0.23	2950	55.79	51.89		0.07	2598
% FRL	77.17	77.98		0.61	2925	77.08	77.43		0.83	2581
% LEP	15.61	16.53		0.30	2912	15.83	16.56		0.40	2581
% Hispanic	19.94	19.35		0.78	2906	19.78	19.92		0.92	2579
% Black	40.44	39.67		0.57	2906	40.34	38.75		0.36	2579
% White	23.33	23.13		0.86	2906	23.18	22.88		0.69	2579
% Other Race	16.29	17.84		0.19	2906	16.69	18.44		0.16	2579
Spr. 2013 ITBS Comp. Z-Score	0.00	0.04	1.02	0.55	2833	0.05	0.09	0.99	0.59	2598
Spr. 2013 ITBS Comp Std. Score	182.16	183.00	22.85	0.55	2833	183.34	184.25	22.38	0.59	2598
# Trifolds Returned Summer 2013	0.04	3.48	3.07	0.00	2950	0.05	3.70	3.15	0.00	2598
Total # Comp. Qs Correct across Tri-folds	0.09	7.62	6.84	0.00	2950	0.09	8.12	7.02	0.00	2598
Spring 2014 EOG Reading Score	441.87	442.40	9.98	0.24	2697	441.91	442.63	9.93	0.15	2598
Spring 2014 EOG Reading Z-Score	-0.03	0.03	1.00	0.24	2697	-0.02	0.05	1.00	0.15	2598
Spring 2014 EOG Reading % Proficient	42.15	44.10		0.40	2697	42.07	44.62		0.33	2598

Note. Analysis sample is comprised of students with non-missing pre-test and posttest measures. *p*-values are for test of equality across conditions, from a regression model that controls for fixed effects of homeroom and clusters standard errors at the school level.

Table 2. Delayed intent-to-treat effects on spring 2014 EOG posttest reading scores

Variable	(1) All Schools	(2) High Poverty Schools	(3) Moderate Poverty Schools
READS	0.039*	0.052*	0.012
	(0.019)	(0.022)	(0.035)
ITBS Pretest (std)	0.741***	0.752***	0.722***
	(0.012)	(0.016)	(0.018)
N	5569	3815	1754
R^2	0.626	0.618	0.617

Note. EOG and ITBS are standardized within grade. Standard errors clustered at school-level in parentheses. All models control for fixed effects of spring 2013 homeroom.

$\sim p < 0.10$. $^*p < .05$. $^{**}p < .01$. $^{***}p < .001$.

Table 3. Instrumental variables estimates of effect of measures of summer book reading on spring 2014 EOG reading test

Variable	(1) All Schools	(2) High Poverty Schools	(3) Moderate Poverty Schools
Number of trifolds returned	0.010* (0.005)	0.014* (0.006)	0.002 (0.007)
N	5569	3815	1754
High number of trifolds returned	0.086* (0.035)	0.114* (0.045)	0.040 (0.057)
N	5569	3815	1754
Num. trifolds >=1 Correct	0.012* (0.005)	0.017* (0.007)	0.004 (0.008)
N	5569	3815	1754
Self-reported books read	0.004 (0.004)	0.008~ (0.004)	-0.008 (0.007)
N	5569	3815	1754
High self-reported books read	-0.003 (0.055)	0.040 (0.064)	-0.101 (0.100)
N	5569	3815	1754

Note. Standard errors clustered at the school level are in parentheses. All models control for fixed effects of spring 2013 homeroom and baseline test scores. Treatment*Homeroom interactions used as instruments.
~$p < 0.10$. *$p < .05$. **$p < .01$. ***$p < .001$.

Online Supplementary Materials

Fidelity of Lesson Implementation Analyses

In a previous study (Authors, 2015) testing the feasibility of the lessons with 29 third-grade teachers in 14 schools, teachers' mean adherence rate was 73%, and consistent with other experimental studies of comprehension instruction in the elementary grades (Williams et al., 2005). Teacher adherence was unrelated to school poverty and student pretest reading scores or posttest reading scores, underscoring the feasibility of scaling up the lessons in a larger experimental study. Given the acceptable levels of fidelity observed previously and the costs associated with video-taping lessons, we chose to use teacher logs to check fidelity in this study. All participating teachers completed a log after each of the six lessons and reported the number of lesson components that were implemented. Overall, 95% of teachers completed the six lesson logs, and teachers reported implementing over 90% of the lesson components. Lessons 1, 2, and 3 had 30, 43, and 41 lesson components and lessons 4, 5, and 6 had 41, 47, and 38 lesson components. On average, the narrative text lessons (1-3) had 38 components and the informational text lessons (4-6) had 42 components. Teachers completed 6 lesson logs and reported whether they implemented of lessons component. In addition, 94% of the children attended all six comprehension lessons, which were implemented during the regular school day.

Sensitivity Analyses

First, the results reported in Table 4 suggest that there were no statistically significant interactions involving the treatment variable and gender, pretest readings, grade level, free lunch status, and limited English proficiency status. There was a negative treatment by White student interaction suggesting that treatment effects were significantly lower for white than non-white students. With the exception of this interaction effect, the treatment effects appear consistent

across baseline characteristics and support the conclusion that there was an overall positive main

effect in EOG reading scores. In addition, we conducted several sensitivity analyses related to

missing data. First, we compared intent-to-treat estimates based on whether children did or did

not take the fall ITBS reading test. There was a positive and statistically significant treatment

effect of .042 for children with fall reading scores was similar to the estimate for the full analytic

sample. Among children with missing fall scores who moved to another non-study school and

took the NC EOG posttest, the treatment effect was .02 and not statistically significant. Second,

for children with non-missing spring ITBS reading scores, we also imputed missing values for

other variables using the mean value within the same classroom and experimental group (in cases

where all students in a classroom/treatment cell were missing on the variable in question, we

imputed the grade level/treatment group-specific mean). All regressions include as controls

indicators for observations with imputed values. Imputation allows us to analyze a consistent

sample across the different analyses we perform. Analyses not reported here show qualitatively

similar results when we do not impute missing values and allow the samples to vary across

analyses. These results are available from authors upon request.

Second, we used an alternative modeling strategy to estimate intent-to-treat estimates.

When we estimate intent-to-treat estimates using a HLM model with school-level random

effects, the main effect of .036 remained positive and statistically significant. The results were

also replicated in models that include both school-level random effects and controls for the fixed

effect of the homeroom classroom block. Results are available from authors.

Third, we used a single instrumental variable (randomization to condition) to identify the

effects of the quantity and quality of summer book reading as measured by tri-fold data.

Specifically, we replicated the results in Table 3 using only a single IV rather than multiple site

by treatment IVs, suggesting that the results in Table 3 are robust across model specifications.

Results are available from authors. Finally, for the grade 2-3 cohort, students were administered

the oral reading fluency assessment in spring 2014. There was no treatment effect overall or by

school poverty on this outcome, consistent with findings from two earlier studies (Authors 2006;

2008). Results are available from authors.

Fourth, we assessed treatment effects on oral reading fluency outcomes for the cohort of

children who started the study as second-graders. In spring 2014, the subsample of children who

started as second-graders was administered an oral reading fluency assessment at the end of year,

in third grade. Oral reading fluency was assessed with the Dynamic Indicators of Basic Early

Literacy Skills (DIBELS; Good et al., 2013). Children were asked to read 3 different grade-level

passages for 1 minute and the number of words correctly read was used in the analysis. The

descriptive statistics for our sample ($M = 104$, $SD = 41$) were close to those reported in studies

involving a national sample of third-grade children (Tindal, 2005). For the subsample of grade 2

children, there was no significant effect on oral reading fluency. Results are available from

authors.

Table 4. Intent-to-treat estimates involving interactions between treatment and baseline demographic characteristics

Variable	(1)	(2)	(3)	(4)	(5)	(6)
READS	0.0438~	0.0609**	0.0405*	0.0355	0.0509	0.0389~
	(0.0237)	(0.0190)	(0.0185)	(0.0242)	(0.0367)	(0.0209)
Female	0.0386~					
	(0.0222)					
READS*Female	-0.00692					
	(0.0333)					
ITBS Pretest (std)	0.740***	0.735***	0.751***	0.741***	0.725***	0.734***
	(0.0123)	(0.0128)	(0.0176)	(0.0123)	(0.0127)	(0.0129)
White		0.141***				
		(0.0352)				
READS*White		-0.104**				
		(0.0386)				
READS*Pretest			-0.0198			
			(0.0196)			
Grade 3				-0.353*		
				(0.145)		
READS*Grade 3				0.00823		
				(0.0343)		
FRL					-0.175***	
					(0.0300)	

READS*FRL						-0.0173
						(0.0407)
LEP						-0.127***
						(0.0316)
READS*LEP						-0.00529
						(0.0416)
Constant	-0.0708***	-0.0800***	-0.0507***	0.115~	0.0875***	-0.0273*
	(0.0136)	(0.0121)	(0.00921)	(0.0677)	(0.0250)	(0.0107)
N	5569	5525	5569	5569	5531	5529
R^2	0.626	0.627	0.626	0.626	0.631	0.627

Note. EOG and ITBS are standardized within grade. Standard errors clustered at the school level are in parentheses. All Models control for fixed effects of spring 2013 homeroom. Results that use imputed data for missing demographic variables support the same conclusions.

~$p < 0.10$. *$p < .05$. **$p < .01$. ***$p < .001$.

Table 5. Cost-effectiveness estimates for READS, supplemental educational services (SES) tutoring, and Reading Partners

	READS[a]	SES (Chicago)[b] Assuming 10:1 student tutor ratio	SES (Chicago)[c] Assuming 2:1 student tutor ratio	Reading Partners[e]
Intent-to-treat effect (ITT)	0.04	0.06	0.06[d]	0.10
Cost/child				
Lower bound	$250	$1,580	$790	$480
Upper bound	$480	$1,700	$850	$1,270
Cost-effectiveness ratio				
Lower bound	.16 SD	.038 SD	.08 SD	.21 SD
Upper bound	.08 SD	.035 SD	.07 SD	.08 SD

Notes. Cost-effectiveness estimates presented in this table reflect only direct monetary expenditures for program delivery (C. Heinrich, personal communication, June 10, 2015; McCombs et al., 2011, p. 77). While "hidden costs" such as volunteer hours and utilization of school facilities contribute to the overall cost of an intervention, this category was excluded uniformly for various reasons. The READS program did not rely on volunteers or additional use of school facilities. Heinrich et al. (2014), did not set out to conduct an exhaustive cost analysis in their study of SES providers in Chicago Public Schools (along with three other districts), instead focusing broadly on providers' average hourly charge rate and the average number of SES hours each student received.

[a] Cost estimates for READS were derived from the project's realized implementation expenditures in the study year. To account for uncertainty and cost variation depending on district and school size, the authors generated a range of plausible values by calculating costs per pupil in several district-level scenarios representing varying scales of implementation. Because this comparison addresses the costs of implementing READS as a program rather than as an experimental study, implementation costs were adjusted to exclude "research-only" costs such as ITBS post-testing and indirect costs charged by colleges and universities. Major cost ingredients reflected in these estimates include salaries for program support staff and contracted service providers, books and READS lesson materials, teacher and school stipends, ITBS pre-testing, and parent communication through READS family night.

[b] Cost estimates for SES in Chicago Public Schools (CPS) were calculated using the average hourly rate paid by the district to SES providers in 2011-12 ($44/hour). Heinrich et al. (2014) present the average number of hours of SES received by CPS students in each of four school years from 2008-2012. Year-to-year variation in the number of hours per student allowed the authors to estimate a range of values for the average cost of SES per student.

[c] Heinrich et al. (2014) note that "approximately 80 percent of all sessions had a student–tutor ratio of less than 4:1" (p. 19).

[d] Pooled across four study years, 2008-2012.

[e] Jacob et al. (2015) report an average effect size of .10 in reading comprehension (p. 53) and average per pupil cost of $710, ranging from a minimum of $480 to a maximum of $1,270 (p. 75).

Figure 2. Informational text trifold
**Information Impressions for Barack
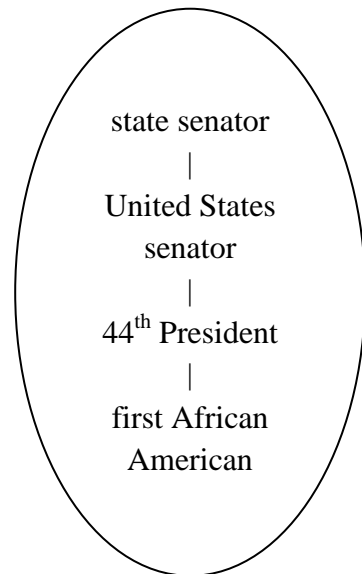Obama**
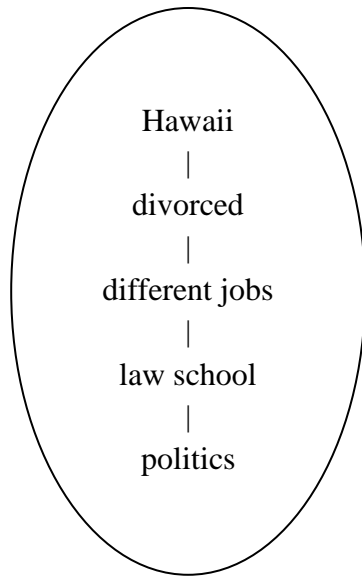
Main Idea Predictions

For Students: Please circle the best answer. Look
back at the passage if you need to.

Hawaii
|
divorced
|
different jobs
|
law school
|
politics

⟹

a. Obama's childhood and
background

b. Different places Obama
lived

c. Obama's childhood friends

1. **What is a job that Barack Obama had before
becoming the President?**
    a. Doctor
    b. Senator
    c. Teacher
    d. Surfer

2. **Where did Barack Obama first go to
college?**
    a. New York City
    b. Hawaii
    c. Indonesia
    d. California

3. **What would be a good title for this
passage?**
    a. Barack Obama: Our President
    b. Different Jobs for People
    c. From California to New York
    d. United States Presidents

state senator
|
United States
senator
|
44th President
|
first African
American

⟹

a. Politics in Hawaii

b. Important people in
Obama's life

c. Why Obama is famous

4. **This story was**
    a. Too easy
    b. Just right
    c. Too hard

5. **How much did you like this story?**
    a. I love it
    b. I liked it
    c. It was okay
    d. I didn't like it

Figure 3. Sample narrative trifold
Story Impression for "The Closet Creature"

<div align="center">

Story Prediction

</div>

Adam

|

scary sounds

|

darkness

|

closet door

|

David

|

bad dream

|

flashlight

|

pillowcase

|

orange cat

_____

_____

_____

_____

_____

_____

_____

_____

_____

_____

_____

_____

_____

_____

_____

_____

_____

_____

For Students: Please circle the best answer. Look back at the story if you need to.

1. **Who was Apricot?**
   a. Adam's sister
   b. Adam's snack
   c. Adam's cat
   d. Adam's stuffed animal

2. **Why did Adam run to David's room?**
   a. He wanted to play
   b. He heard a noise in his closet
   c. He couldn't sleep
   d. He had a nightmare

3. **Based on the story, what can you guess about Adam and David?**
   a. They live in a big city
   b. They are neighbors
   c. They like dogs
   d. They help each other out

4. **This story was**
   a. Too easy
   b. Just right
   c. Too hard

5. **How much did you like this story?**
   a. I loved it
   b. I liked it
   c. It was okay
   d. I didn't like it