Impact of the UPSTART Program on Forestalling Summer Learning Loss

L. Jon Hobbs
Evaluation and Training Institute

Melanie Overby
Evaluation and Training Institute

July 1, 2019

Authors Note

This study is supported by a grant from the United States Department of Education, Investing in Innovation (i3) Program under award number U411B130020. The content is solely the responsibility of the authors and does not necessarily represent the views or policy of the U.S. Department of Education.

Correspondence concerning this paper should be addressed to L. Jon Hobbs, Evaluation and Training Institute, 100 Corporate Pointe, Suite 387, Culver City, CA 90230. Contact: jhobbs@eticonsulting.org.

Abstract

The UPSTART Summer program is a federally funded i3 validation project that uses a computer-based program to maintain and develop the literacy skills of elementary school students in rural Utah during the summer months when school is out of session. Researchers used a quasi-experimental design to evaluate the impact of the program in forestalling literacy learning loss during several summer periods. Students in the treatment group participated in the UPSTART Summer program, in the summer periods after kindergarten, first grade, and/or second grade. A second group of children, who were not enrolled in the program served as a comparison. Statistical matching procedures were used to create separate treatment and comparison analytic samples for each outcome measure that were equivalent on baseline scores and demographic variables (e.g., school, gender, race, language learner status, household income, Title 1 school enrollment, etc.). Standardized literacy assessments of letter knowledge, phonics, and reading fluency were administered prior to program commencement at the end of the academic school year and upon program completion at the beginning of the following school year. Results revealed that the UPSTART Summer program had a significant impact in reducing literacy learning loss in rising first graders on assessments of letter naming fluency, nonsense word fluency (correct letter sounds), and a reading composite score when compared to a matched comparison group. There were no differences in learning loss rates between rising first graders and comparison students on assessments measuring phoneme segmentation fluency or nonsense word reading (whole words read). Additionally, the UPSTART Summer program did not have an impact on literacy learning loss prevention in rising second or third grade students as measured by assessments of nonsense word reading, oral reading fluency, or overall reading composites. Taken together these results suggest that the UPSTART program helps to maintain early literacy skills in the summer months between Kindergarten and first grade.

Impact of UPSTART Program on Forestalling Summer Learning Loss

Research suggests that students' achievement test scores often decline or plateau in the summer months, a phenomenon known as "summer learning loss" (Borman, Overman, Fairchild, Boulay, & Kaplan, 2004; Cooper, Nye, Charlton, Lindsay, & Greathouse, 1996). In professional practice the phenomenon is also known as the "summer learning slump" or "learning decay," and it requires educators to develop strategies to help students re-learn subjects from the previous school year. For example, Cooper et al. (1996) synthesized 39 studies measuring summer learning loss and quantified the effect as equivalent to a minimum of one month of loss on a grade equivalent scale. In other words, after summer vacation, students entered school in the fall one month behind where they left in the spring.

Cooper et al. (1996) also noted differences in learning loss based on subject area and socioeconomic status. While middle-class and low-income students experienced similar rates of decline in math, the gap between middle-class and low-income students widened significantly in reading. Additional research goes on to show middle-class students experienced growth in reading over the summer and low-income students experienced learning decay (as cited in Paris, Pearson, Cervetti, et al., 2012). One explanation for this disparity has been proposed through the "Faucet Theory." This theory explains that when school is in session, the resource faucet is on for children of all socio-economic backgrounds, but when school is not in session, the faucet is turned off. Middle-income children, however, have the educational resources to keep the faucet on during summer months (Borman, Overman, Fairchild, Boulay & Kaplan, 2004).

In response to the growing body of research on the problems associated with summer learning loss and students' academic achievement, policymakers and school districts often

implement summer learning programs as interventions. While summer learning programs may also help narrow the gap between students of different socio-economic statuses. While not all summer programs succeed in preventing students' learning decline, some evidence suggests that a successful program can help to improve the academic achievement of students. For example, a study of students in Grades 2-5 found that 41-60% of students experienced improvement in language usage, math, or reading (Bakle, 2010). This is a prime example of how a summer program may alleviate the summer learning loss through summer instruction.

While a traditional summer learning program may help address summer learning loss, it may not be the most feasible solution in certain situations. Rural communities, for example, often face additional challenges implementing summer programming, including resource constraints, a lack of human capital and accessibility, all of which limit the level of services that may be provided to children within these communities (Phillips, Harper, Gamble, 2007). These rural communities may also have the greatest need for services. For example, children living in rural communities are more likely to live in lower income families compared to children living in urban areas (Addy, Englehardt, & Skinner, 2013) and may have limited access to high-quality preschools (Khan, Justice, & Jiang, 2016). Moreover, findings from a recent Early Childhood Longitudinal Study show that kindergarteners in rural communities perform more poorly on math and reading measures than urban and suburban areas (Miller & Votruba, Drzal, 2013). Funding is also an issue for providing educational resources and high-quality pre-kindergarten programs to low-income children, as 15% of children in rural communities attend a high-quality preschool as opposed to the 30% of children in rural and suburban areas (Nores & Barnett, 2014). Providing these resources to rural, often low-income areas is essential for the success of a summer program.

Educational Technology as a Solution for Rural Communities

With the increasing adoption of wide-spread broadband connectivity, educational technologies can help increase access to educational opportunities in rural communities where access to is hindered by a lack of resources and geographical barriers (Foster, Anthony, Clements, Sarama, & Williams, 2016). Although the research is limited on the use of technology programs for mitigating the summer slide, there is evidence of the effectiveness of computer-based instructional programs on improving students' reading outcomes. For example, a review of forty-two studies of computer-based instructional programs with beginning readers found an overall positive effect size of .19 (Blok, Oostsdam, Otter, & Overmat, 2002). Studies of programs such as Building Blocks Software, Head Start, and E-Spark Learning have also produced promising results. Building Blocks Software aims to provide math practice online year-round for children in pre-k to 8th grade. Program results were consistent with prior research in that the judicious use of research-based educational software was shown to have a moderate effect on students' math development, with an effect sizes of 0.43 and 0.37 for children's numeracy and applied mathematics skills, respectively (Foster et al., 2016). Similarly, a study of computer-based instruction conducted with 122 Head Start children revealed that the experimental group had significantly higher post-test scores compared to the control group (54.66 vs. 40.23, $p < .05$) (Li, Atkins & Stanton, 2006).

There is limited research on how computer-based learning programs may help reduce the summer slide affect. E-Spark learning was one of the only computer-based instructional programs used for the purposes of reducing the summer slide effect that has also been studied. The study used the nationally normed NWEA Measures of Academic Progress (MAP) assessment data to measure student outcomes and found that participating students grew 1.5

percentile points in reading, whereas comparison students lost 5 percentile points over the summer (Tollerson & Guckert, 2016).

Noting these trends in educational technology and the need for rigorous research to measure its impact, the U.S. Department of Education identified educational programs that serve rural communities and instructional solutions that effectively use technology as two of the most pressing priorities for education researchers in the 2013 Investing in Innovation (i3) Notice of Final Priorities, Requirements, Definitions, and Selection Criteria.

The Waterford Institute was awarded an i3 grant from the United States Department of Education to fund the validation of "Utah Preparing Students Today for a Rewarding Tomorrow (UPSTART)", a pilot project that uses a home-based education technology approach to develop the school readiness skills of preschoolers in rural Utah and forestall learning decay during the summer months when children are enrolled in elementary school. The Evaluation and Training Institute (ETI) was contracted as an external evaluator to conduct an independent evaluation of the UPSTART Summer program's effectiveness in rural communities in order to help expand our understanding of what works in education and for whom and what contexts specific interventions are effective.

In this impact evaluation report, we first present an overview of the UPSTART Summer program, which was implemented in Utah rural districts from 2016 to 2018. We then list the research questions that guide the evaluation of the program on the literacy achievement of participating students. We describe the study design, outline the operationalization of outcomes, characteristics of participants in the intervention and comparison conditions, matching process, methods for analysis, and a description of the baseline equivalence of our treatment and control

samples by cohort. We then provide our results and a discussion of our interpretations of the findings.

Program Description

The chief objective of the UPSTART Summer Program was to provide educational enrichment to elementary school children in order to maintain their school learning during the summer months. The primary component of the UPSTART Summer program is the Waterford Early Learning (WEL) software, an in-home computer-based program that provided children with literacy instruction. The software provided an individualized learning experience that adjusts to children's skill level and content is delivered online through adaptive lessons, digital books, and activities. The program was designed to promote mastery of literacy skills that prepared students over the summer for entry into the next grade level through an individualized learning sequencer that adapted to each child's skill level. Based on student performance, the sequencer ran remedial activities to reteach and practice skills again or advanced to another objective if students mastered concepts. Required usage of the program software was 15 minutes of day, five days a week from June through August.

Based on reading instruction guidelines outlined by the National Reading Panel (2008) that emphasize phonemic awareness, phonics, fluency, comprehension, and vocabulary, the UPSTART Summer Reading program was designed to use research-based best practices for literacy instruction. Table 1 showcases the reading domains and skills in phonics, comprehension/vocabulary, language concepts, phonological awareness, and fluency taught by the UPSTART Summer at the pre-reading, basic reading, and fluent reading levels.

Table 1

Waterford Reading Curriculum

Overview	**Pre-Reading**	**Basic Reading**	**Fluent Reading**
PHONICS Systematically presents letter-sound relationships and decoding skills to help students break the reading code.	• Spell child's name • Recognize A through Z • Recognize a through z • Learn all letter sounds and 20 power words to read 10 leveled readers	• Learn common spelling pattern(s) for all 44 sounds in English • Learn 160 power words • Read leveled readers • Begin reading with fluency	• Learn more complex spelling patterns • Learn 94 power words • Read leveled readers • Practice automatic word recognition
COMPREHENSION & VOCABULARY Develops vocabulary and critical thinking skills through rich reading experiences.	• Read along and understand nursery rhymes • Read along and understand alliterative and Read-along books • Learn 308 target vocabulary words	• Read along and understand traditional tales • Learn 78 target vocabulary words • Learn common word structure as clues to the meaning of words	• Read along and understand Read-along books • Learn 262 target vocabulary words
LANGUAGE CONCEPTS Builds knowledge of written language (from print concepts to basic grammar).	• Understand print • Develop understanding of parts of speech and sentence structure • Learn about the writing process through drawing and writing	• Learn basic grammar concepts such as sentences, punctuation, and capitalization. • Learn about the writing process through different types of text	• Learn parts of speech (nouns, verbs, and adjectives) and parts of words (prefixes, suffixes) • Lear about the writing process (prewriting, drafting, revising, and editing)
PHONOLOGICAL AWARENESS Develops awareness of the sounds of language including syllables, rhyming and the individual sounds in words.	• Listening skills • Match rhyming words • Identify the number of syllables in words • Identify initial and final sounds in words • Break words into individual sounds and blend individual sounds into words. • Change a sound in a word to make a new word	• Identify initial and final sounds in words • Break words into individual sounds • Change a sound in a word to make a new word	
FLUENCY Develops the ability to read text accurately, automatically, and with expression and correct phrasing.		• Build oral reading expression • Build oral reading speed	• Build oral reading expression • Build reading speed to 90 words per minute

The UPSTART Summer program had several resources available to families to assist children in meeting the prescribed usage requirements. Before beginning the program, participating parents attended a comprehensive orientation where they reviewed state grade level guidelines alongside UPSTART curricular content, discussed strategies for motivating children to use the program consistently, and learned about software features and resources. A password-protected Parent Manager portal in the UPSTART software program allowed parents to monitor children's usage on a daily basis, as well as review children's unit lesson scores, placement results, and specific activity recommendations for enrichment. During the summer when children were actively engaged with the program, weekly emails were sent to parents that contained program news and updates, described usage contests and incentives, and included a graph that tracks students' weekly use. Parents were also paired with a program staff member that provided technical and motivational support during the summer.

Lastly, computer hardware and/or high-speed internet access were provided to families in need for the duration of the UPSTART Summer program to ensure equitable access and allow families connect to the software online.

Research Questions

We designed a quasi-experiment study to test our primary research questions that investigate the impact of the UPSTART Summer program on forestalling elementary students' learning loss over the summer months:

> *RQ1*: Do children who use the UPSTART Summer program over the summer between K and first grade (Cohort 1) have less learning loss during the summer break (measured as the difference between End of Year kindergarten and Beginning of Year first grade

Acadience Reading test scores), compared to matched children who did not use the UPSTART program?

RQ2: Do children who use the UPSTART Summer program over the summer between first grade and second grade (Cohort 2) have less learning decay during the summer break (measured as the difference between End of Year first grade and Beginning of Year second grade Acadience Reading test scores), compared to matched children who did not use the UPSTART program?

RQ3: Do children who use the UPSTART Summer program over the summer between second and third grade (Cohort 3) have less learning loss during the summer break (measured as the difference between End of Year second grade and Beginning of Year third grade Acadience Reading oral reading fluency subscale scores), compared to matched children who did not use the UPSTART program?

Methods

Research Design

The evaluation investigates the impact of the UPSTART Summer program on minimizing participating students' literacy learning loss upon reentry into school. Outcome and student characteristic data were collected from state student demographic and standardized test data. The study utilized a quasi-experimental design (QED). A matching procedure was used to ensure equivalence of treatment and comparison groups in order to draw causal inferences from the intervention to outcomes. Unlike a randomized control design where the treatment and control groups are considered equivalent because participants have an equal chance of being assigned to a given condition, quasi-experimental designs rely on procedures other than randomization to create balanced treatment and comparison groups.

In the current study, we compare students attending elementary schools in Utah who participated in the UPSTART Summer program (treatment group) to students who did not enroll in the program but are similar to the treatment group on observable characteristics and who were enrolled in the same rural elementary schools (comparison group). To maximize the equivalence of the two groups, we matched at the school level so that if a treatment student attended a particular school, and their comparison counterpart was also enrolled in the same school. Within schools, students were matched on the following variables: baseline literacy measures and demographic variables, including students' gender, race, household income, attendance at a Title 1 school, English Language Learner status, and participation in the special education program.

In addition to the matching process, we utilized statistical analyses techniques to mitigate any remaining differences between the treatment and comparison groups, such as adding pre-test literacy scores as a covariate in linear regression models that estimate the impact of the intervention on preventing learning loss over the summer. We assess the relationship between the outcome variable (literacy learning loss), treatment/comparison group membership, and relevant covariates (race, low income status).

Intervention and Evaluation Participants

In keeping with an intent-to-treat sample that includes all participants regardless of treatment received, deviations from protocols, or withdrawal from the treatment group (Gupta, 2011), all students who participated in the UPSTART Summer program were eligible for the evaluation, irrespective of the amount of UPSTART software use. To participate in the UPSTART Summer program, students had to be matriculating into the correct grade (first grade for Cohort 1, second grade for Cohort 2, and third grade for Cohort 3), residing in one of eighteen rural school districts in Utah, and have parents who completed parent orientation and

training. Children enrolled in the UPSTART program were not prevented from engaging in other summer educational and/or enrichment activities.

Students in the comparison group did not receive any aspect of the UPSTART Summer program or any alternative interventions and continued to engage in their usual summer activities (i.e., a business-as-usual comparison group). No services were withheld from the comparison group.

Outcome Measures

Acadience Reading (previously published as DIBELS Next), a standardized tool designed to measure the acquisition of early literacy skills from kindergarten through sixth grade, was used as the baseline and outcome measure (Good & Kaminski, 2011). Measures of reliability for Acadience Reading, including alternate-form (.95 to .97) and test-retest reliability (.81 to .94) are strong and each individual subtest assesses key skills determined to be essential for reading: phonemic awareness, phonics, accuracy, and fluency (National Reading Panel, 2000). The Acadience Reading subscales are presented as age-appropriate literacy indicators as outlined in Table 2 and were administered without modifications. The Acadience Reading Composite Score is a combination of multiple grade-level appropriate Acadience Reading subscale scales and provides the best overall estimate of a students' reading proficiency.

Table 2

Acadience Reading Subscales

Subscale Name (*Literacy Construct*)	**Grade Level**	**Description**
First Sound Fluency (FSF) (*Phonemic Awareness*)	K	Child provides the initial phoneme (2 points) or the initial consonant blend, consonant plus vowel, or consonant blend plus vowel (1 point) of as many target words as possible in one minute.
Letter Naming Fluency (LNF) (*Alphabet Knowledge*)	K-1st Grade	Child names as many letters as possible in 1 minute.
Phoneme Segmentation Fluency (PSF) (*Phonemic Awareness*)	K-1st Grade	Child segments as many target words into its component parts as possible in one minute. Partial credit is given for partial segmentation – for example (/s/ /un/) for sun receives 2 points.
Nonsense Word Fluency (NWF) (*Phonics*)	K-2nd Grade	Child receives credit for as many nonsense or pseudo-words read correctly (WWR) or reading correct letter sounds (CLS) within each word in 1 minute.
Oral Reading Fluency (ORF) (*Accuracy/Fluency*)	1st - 6th Grade	Child reads an unfamiliar grade-level passage for 1 minute and scored by the number of words read correctly.

Where possible the Acadience Reading Composite Score served as the overall measure for confirmatory analyses of the UPSTART Summer program intervention, while the Acadience Reading subscales provided exploratory analyses of how the program impacts specific areas of early literacy.

Data Collection

Utah public schools administer the Acadience Reading assessment during three testing windows: at the beginning of the school year (Beginning of the Year, or BOY); in December or January (Middle of the Year, or MOY); and at in mid-April through June (End of Year, or EOY).

Acadience Reading is individually administered to all students by teachers during the course of the school day as part of the Utah statewide standardized testing program. In the current study, assessments measured in the spring during the End of Year (EOY) testing administration are used as the baseline measure of literacy achievement before students leave school for summer vacation. Treatment students participate in the UPSTART Summer program outside of school during the summer break while comparison students engage in "business as usual" summer activities, and the assessments measured upon return to school in the fall during Beginning of Year (BOY) testing serve as the outcome measure. Figure 1 displays the administration timeline of the baseline and outcome measures for Cohorts 1, 2, and 3.

Figure 1

Administration of Baseline and Outcome Measures

2016 EOY Spring	UP Sum	**2016** BOY Fall	**2017** EOY Spring	UP Sum	**2017** BOY Fall	**2018** EOY Spring	UP Sum	**2018** BOY Fall
Kindergarten		1st Grade			2nd Grade			3rd Grade
Cohort 1			**Cohort 2**			**Cohort 3**		
Baseline Measures	Outcome Measures		Baseline Measures	Outcome Measures		Baseline Measures	Outcome Measures	
COMP (KEOY)	COMP (1BOY)		COMP (1EOY)	COMP (2BOY)		ORF (2EOY)	ORF (3BOY)	
LNF (KEOY)	LNF (1BOY)		NWF (1EOY)	NWF (2BOY)				
PSF (KEOY)	PSF (1BOY)		ORF (1EOY)	ORF (2BOY)				
NWF (KEOY)	NWF (1BOY)							

Note: Measure names are the following: COMP (Acadience Reading Composite Score); LNF (Letter Naming Fluency); PSF (Phoneme Segmentation Fluency); NWF (Nonsense Word Fluency); ORF (Oral Reading Fluency). Time periods are the following: KEOY (Kindergarten End of Year); 1BOY (1st grade Beginning of the Year); 1EOY (1st grade End of Year); 2BOY (2nd grade Beginning of Year); 2EOY (2nd grade End of Year); 3BOY (3rd grade Beginning of Year).

Data Management

All data used for the impact analyses were obtained from program records from the Waterford Institute, makers of the UPSTART Program, and from student demographic and academic data that participating UPSTART school districts elected to share with the evaluator. Annual demographic files contained information about student race, gender, English Language Learner status, free or reduced-price lunch eligibility, and school enrollment for all public school students. Provided yearly academic data consisted of separate data files that contained end of year (EOY) literacy scores taken before the UPSTART Summer program began for intervention students and beginning of year (BOY) literacy scores that were collected after the program was completed. All students who participated in the UPSTART summer program were eligible for selection into the treatment group and all remaining students who did not use the UPSTART program were eligible for the comparison group. Student demographic files were merged with literacy academic data and UPSTART program usage files to create a complete master dataset.

Selection of a Matched Comparison Group of Students

Separate analytic samples were constructed from the master dataset for each outcome measure using a statistical procedure called Coarsened Exact Matching (CEM) to match control students to treatment students. CEM has been shown to be superior to methods, such as propensity score and Mahalanobis matching, in its ability to reduce imbalance, model dependence, estimation error, bias, and other criteria (Iacus, King, and Porro). Additionally, CEM is considerably faster than other matching methods with data that has continuous, discrete, and mixed variables (King & Nielsen, 2016).

During the CEM procedure, each treatment child is statistically matched with a control child who is most similar to them and if no matches can be made, children are removed from the

sample. Additional tests are preformed to assess the balance between the treatment and control group to ensure that the groups are as similar as possible and the resulting matched treatment-control sample consists of treatment children who have a statistical control "twin" on specific observable features. We selected several relevant characteristics by which to match comparison students to children enrolled in the UPSTART Summer program intervention. UPSTART students were matched one-to-one to a group of non-software-using students by:

- School
- End-of-prior-school year score on the Acadience Reading subscale that corresponded to the outcome subscale (i.e., pre-intervention baseline measure)

and a set of binary covariates for:

- Gender
- White indicator
- Hispanic indicator
- Low-income status
- Title I status
- English Language Learner status
- Participation in a special education program.

The CEM procedure produced separate analytic samples by program cohort for the Acadience Reading composite score and each of the appropriate subscale scores, depending on the grade level tested. Each sample contained equal numbers of treatment and control subjects exactly balanced in terms of the categorical matching variables and statistically non-significantly distinct in terms of baseline Acadience Reading scores. Specifically, for each school, treatment and control observations were selected for the same number of boys and girls, the same number

of whites and non-whites, the same number of low-income and non-low-income students, and so on. This creation of a matched control group will improve our precision in estimating treatment effects. Appendix A displays the demographic breakdown of the matched treatment and control groups for each of the analytic samples.

Analysis

The evaluation investigated the impact of the UPSTART Summer program on participating students' summer learning loss. Outcome and student demographic data were collected from standardized test scores administered as part of the Utah state testing battery and statewide demographic data. The study used a non-experimental or quasi-experimental design that relies on procedures other than randomization to create a treatment-control contrast, such as statistical techniques to improve the equivalence of treatment and comparison groups so that we may make valid inferences of program impact. Appendix B provides the correlation between pre-program baseline measures and post-program outcome measures.

We used a series of linear regressions to estimate the impact of the UPSTART Summer program on students' summer literacy learning loss, calculated as the literacy outcome score at the beginning of the school year from the baseline score measured at the end of the previous school year. We regress the outcome variable (learning loss) on the predictor of interest (UPSTART Summer participation/comparison group status), along with relevant covariates (baseline literacy scores of the outcome variable, socioeconomic status, and a set of dummy variables for forty-four schools to account for naturally occurring blocking and stratification).

We define the following variables for each elementary student in linear regressions to estimate the impact of the UPSTART Summer program on our outcome variables of interest: Y_{ij} is the preschooler's score on post-test composite literacy measures; Treatment (T_{ij}) is an

indicator for whether the student received the intervention (Treatment = 1 if the student voluntarily participated in the UPSTART Summer program treatment intervention during a particular cohort year, Treatment = 0 if the student did not participate in the UPSTART Summer program during a particular cohort year); $Y^{Pre}{}_{ij}$ is the preschooler's score on pre-test literacy measures (pre-test covariate); $Y^{White}{}_{ij}$ is an indicator for students who are non-Hispanic (White = 1) compared to those who are not (White = 0); $Y^{Lowincome}{}_{ij}$ specifies students who live in a low income household (Low income= 1) compared to those who are not (Low income = 0); and $Block_j$ is a school block dummy covariate and indicates whether or not a preschooler resided in one of the forty-four specific schools while participating in the UPSTART Summer program. One possible linear regression model that uses these variables is the following:

$$Y_{ij} = \beta_0 + \beta_1(T_{ij}) + \beta_2(Y^{Pre}{}_{ij}) + \beta_3(Y^{White}{}_{ij}) + \beta_4(Y^{Lowincome}{}_{ij}) + \sum_{j=1}^{J-1} \gamma_j Block_j + \varepsilon_{ij}$$

The βs in Eq. 1 are regression coefficients that describe the relationship between each variable and the elementary student's post-test score:

- β_0 is the intercept;
- β_1 is the expected increase in the post-test score for students who participated in the UPSTART Summer program intervention relative to students who did not receive the intervention.
- β_2 is the effect of pre-test data;
- β_3 is the effect of being White, non-Hispanic;
- β_4 is the effect of living in a low-income household;
- β_5 through β_{49} are fixed block dummy variables to measure school effects.

Separate linear regressions were run to estimate the effects of our outcomes of interest.

Adhering to WWC standards, we did not impute missing outcome data, nor did we impute missing covariate data. All impact analyses samples include complete observations that have non-missing post-intervention and baseline data.

Establishing Baseline Equivalence

The use of both a pre-test and a comparison group allow for the examination of potential threats to validity, which could jeopardize a clear interpretation of the results (Shadish, Cook, & Campbell, 2002). Since students could not be randomly assigned to treatment or control groups, the groups begin as nonequivalent by definition, and consequently selection bias can be assumed to operate to some degree in some manner. The pre-test allows us to examine the potential for selection bias by determining the nature of the bias as well as its size and direction (i.e., which group is favored over the other by a particular inequality).

Because all quasi-experimental designs possess a degree of "uncertainty about intervention and comparison group similarity prior to the introduction of the intervention (What Works Clearinghouse, 4.0), we followed the recommendations set by WWC to assess baseline equivalence in the analytic sample. To determine the standardized difference in baseline means, we calculated the mean difference of the baseline measures and divided it by the pooled standard deviation. Determination of the magnitude of the baseline differences between the intervention and comparison group is based on the following standards:

- Less than or equal to .05 standard deviations: Baseline equivalence has been established
- Between .05 and .25 standard deviations: Analysis requires statistical adjustment to meet baseline equivalence requirement (e.g., regression adjustment, analyses of covariance, etc.)
- Greater than .25 standard deviations: Baseline equivalence requirement has not been met.

Effect Sizes

Effect sizes were calculated based on the adjusted mean difference between the treatment and control groups divided by the unadjusted pooled standard deviation, an estimate known as Hedges' g. The adjusted mean difference between the two groups was derived from the linear regression analysis and controlled for pre-test scores, low income status, race, and school enrollment.

Results

We begin by providing the evidence used to establish baseline equivalence for the treatment and comparison groups used in each analytic sample by cohort. Table 3 displays the sample size for each analytic sample by experimental group as well as the baseline means, standard deviations, unstandardized mean difference, and standardized mean difference. Treatment and comparison groups do not differ by more than .05 standardized units on any baseline literacy measure in any of the three cohorts. Consequently, the treatment and comparison groups can be assumed to be equivalent and well balanced across baseline literacy achievement measures prior to the Summer UPSTART intervention in each of the three cohorts under investigation.

Table 3

Baseline Equivalence Assessment for UPSTART Summer Program Participants

		Treatment Group			Comparison Group			T-C difference	Standardized difference
Cohort	Measure	N	Mean	(SD)	N	Mean	(SD)		
1	LNF	297	53.82	12.51	297	53.65	12.36	.17	.01
	PSF	318	59.01	10.52	318	58.87	10.47	.14	.01
	NWF-CLS	300	39.97	15.25	300	40.22	15.34	-.25	-.02
	NWF-WWR	363	3.47	5.96	363	3.39	5.97	.08	.01
	COMP	333	152.22	31.44	333	152.81	31.34	-.59	-.02
2	NWF-CLS	222	90.86	35.29	222	90.68	35.43	.17	.00
	NWF-WWR	190	28.00	14.19	190	28.17	14.12	-.17	-.01
	ORF	251	64.58	27.43	251	64.18	27.29	.4	.01
	COMP	235	215.69	68.23	235	216.09	67.86	-.41	-.01
3	ORF	222	104.54	32.96	222	104.01	32.84	.54	.02
	COMP	207	313.86	69.80	207	314.37	70.02	-.52	.01

Note: Measure names are the following: LNF (Letter Naming Fluency); PSF (Phoneme Segmentation Fluency); NWF-CLS (Nonsense Word Fluency - Correct Letter Sounds); NWF-WWR (Nonsense Word Fluency - Whole Words Read); ORF (Oral Reading Fluency); COMP (Acadience Reading Composite).

Literacy outcomes

Our primary research questions investigate the observed impact of the UPSTART Summer program on summer literacy learning loss. Literacy learning loss is operationalized as the change in reading test scores from the end of the one school year (pre-summer program) to the beginning of the next school year (post-summer program). To estimate the impact of the UPSTART Summer program on each cohort of students, we compare estimates of change in literacy achievement for UPSTART participants to the change demonstrated by similar comparison students who did not elect to use the UPSTART Summer program.

Students who used the UPSTART Summer program over the summer between kindergarten and first grade (Cohort 1) were compared to matched counterparts on literacy measures of alphabet knowledge, phonemic awareness, and phonics. The difference between literacy scales administered at the end of kindergarten (pre-program) and literacy scales administered at the beginning of first grade (post-program) yielded measures of literacy learning loss.

Cohort 1 Analyses

Results of the regression analysis displayed in Table 4 predicted 32% of the variability in learning loss and provide confirmation that participation in the UPSTART Cohort 1 Summer program was a significant predictor in reducing overall literacy learning loss (β = .11, p = .003) as measured by the Acadience Reading Composite score, a global metric that measures letter knowledge, phonemic awareness, and phonics. Low income status was also a significant predictor of learning loss (β = -.09, p = .031), in that Cohort 1 UPSTART students classified as low income had a predicted Reading Composite loss of 21 points compared to the predicted learning loss of 25 points in low income students who did not enroll in the UPSTART Summer program. Students who enrolled in UPSTART but were not classified as low income had a predicted Reading Composite loss of 20 points, while comparison students who were not low income had a predicted loss of 24 points.

Table 4

Regression Analysis of Predictors of Cohort 1 Composite Scale Learning Loss (N=576)

Variable	*B*	*SE B*	*β*
(Constant)	9.76	7.80	
Treatment	5.17	1.71	.11**

Variable	*B*	*SE B*	*β*
Baseline	-.27	.03	-.35**
Low-income status	-4.51	2.08	-.09*
White racial group	5.95	4.97	.05
R^2	0.32		
F	5.39**		

* p<0.05; ** p<0.01

Results from a linear regression model displayed in Table 5 that predicted letter naming fluency learning loss for Cohort 1 students and included income, racial group membership, and baseline pre-program performance as covariates predicted 30% of the variability in the outcome measure. Participation in the UPSTART Summer program in Cohort 1 had a significant impact on reducing letter naming fluency learning loss ($\beta = .09$, $p = .021$).

Table 5

Regression Analysis of Predictors of Cohort 1 Letter Naming Fluency Scale Learning Loss (N=504)

Variable	*B*	*SE B*	*β*
(Constant)	5.92	4.09	
Treatment	1.94	.84	.09*
Baseline	-.35	.04	-.41**
Low-income status	-.06	1.04	-.00
White racial group	5.40	2.63	.10*
R^2	.30		
F	4.52**		

* p<0.05; ** p<0.01

As seen in Table 6, after controlling for pre-test scores, school, income, and racial group membership, participation in the UPSTART Summer program had a significant impact on

preventing learning loss on the phonics indicator Nonsense Words - Correct Letter Sounds (β = .12, p = .003). Household income status also had a significant impact on learning loss for Nonsense Word Fluency - Correct Letter Sounds, β = -.16, p = .002, with children residing in low income households more likely to benefit from participating in the program compared to children who were not living in low income households.

Table 6
Regression Analysis of Predictors of Cohort 1 Nonsense Word Fluency-Correct Letter Sounds Scale Learning Loss (N=512)

Variable	*B*	*SE B*	*β*
(Constant)	3.93	4.42	
Treatment	3.03	1.02	.12**
Baseline	-.29	.04	-.35**
Low-income status	-4.03	1.27	-.16**
White racial group	-.04	2.76	-.00
R^2	.26		
F	3.46**		

* p<0.05; ** p<0.01

There was not a significant program impact on the other phonics indicator Nonsense Word - Whole Words Read (β = .04, p = .217), nor was participation in the UPSTART Summer program a significant predictor of learning loss as measured by the Phonemic Segmentation Fluency subscale (β = .05, p = .142). The regression models for the Nonsense Word - Whole Words Read and Phonemic Segmentation Fluency outcomes can be found in Appendix B.

Table 7 displays the predicted mean score change for the Cohort 1 UPSTART treatment group and comparison groups by outcome, along with standard deviations, raw mean and standardized mean differences, and the resulting p value. The program had a positive impact on

three outcome measures: Reading Composite Scale (a composite measure combining scores from all scales administered at the grade level), Nonsense Word Fluency - Correct Letter Sounds (NWF-CLS) and Letter Naming Fluency (LNF). Treatment students had an average increase of 5.17 points on their Reading Composite Scale test scores over their matched control group peers (Hedges's g = .22). Treatment students also had an average advantage of 3.03 points on the NWF-CLS (Hedges's g = .32), and 1.94 points on the LNF (Hedges's g = .17).

Table 7

Impact Analysis Results for UPSTART Summer Program Cohort 1 Participants

	Treatment Group			Comparison Group					
Measure	Unadj N	Model Adjusted Mean	Unadjusted Post-Test SD	Unadj N	Unadjusted Post-Test Mean	Unadjusted Post-Test SD	Impact estimate	Standardized difference	P value
COMP Change	333	-22.80	26.14	333	-27.45	23.42	5.17	.21	.003
LNF Change	297	-7.88	10.64	297	-10.70	10.98	1.94	.18	.021
PSF Change	318	-8.76	13.27	317	-11.18	13.26	1.26	.09	.142
NWF-CLS Change	300	-4.48	12.57	300	-7.43	12.75	3.03	.24	.003
NWF-WWR Change	363	0.66	5.61	363	0.18	5.07	0.47	.09	.217

Note: Ns are unadjusted Ns for Treatment and Comparison groups
Measure names are the following: COMP (Acadience Reading Composite Score); LNF (Letter Naming Fluency); PSF (Phoneme Segmentation Fluency); NWF-CLS (Nonsense Word Fluency - Correct Letter Sounds); NWF-WWR (Nonsense Word Fluency - Whole Words Read).

Cohort 2 Analyses

The second cohort of the UPSTART Summer program participated over the summer between first and second grade. Participating students along with a group of matched comparison students who elected not to enroll in UPSTART were assessed on literacy measures of overall

reading ability (Acadience Reading Composite Scale) at the end of first grade prior to the program and at the beginning of second grade after program completion. The difference between the two Reading Composite scores resulted in the Reading Composite learning loss outcome. A regression model predicting Cohort 2 reading learning loss that included baseline scores, low-income status, racial group membership, and schools resulted in a significant model for the Acadience Reading Composite score ($R^2 = 0.15$). However, participation in the Cohort 2 UPSTART Summer program was a not a significant predictor of forestalling learning loss. A summary of the Cohort 2 regression model for the Acadience Reading Composite score can be found in Appendix B.

Students who participated in the UPSTART Summer program and those who did not were also assessed on measures of phonics (Nonsense Word Fluency - Correct Letter Sounds, Nonsense Word Fluency - Whole Words Read) and reading fluency (Oral Reading Fluency) at the end of first grade prior to the program and at the beginning of second grade after program completion. The difference between the two scores resulted in the Nonsense Word Fluency and Oral Reading Fluency learning loss outcome. A series of regression models predicting Cohort 2 learning loss that included baseline scores, low-income status, racial group membership, and schools resulted in significant models for Nonsense Word Fluency - Correct Letter Sounds ($R^2 = 0.198$), Nonsense Word Fluency - Whole Words Read ($R^2 = 0.241$), and Oral Reading Fluency ($R^2 = 0.163$); however, UPSTART Summer program participation was a not a significant predictor of learning loss for any of these Cohort 2 outcomes. A summary of each Cohort 2 regression model for the outcomes can be found in Appendix C.

Model adjusted means of each outcome change score is presented in Table 8 for the treatment and comparison groups, along with raw and standardized mean differences. As seen in

Table 8, there was no significant impact of the UPSTART Summer Cohort 2 Program in preventing summer literacy learning loss for rising second graders.

Table 8

Impact Analysis Results for UPSTART Summer Program Cohort 2 Participants

	Treatment Group			Comparison Group					
Measure	Unadj N	Model Adjusted Mean	Unadjusted Post-Test SD	Unadj N	Unadjusted Post-Test Mean	Unadjusted Post-Test SD	Impact estimate	Standardized difference	P value
COMP Change	235	17.35	33.77	235	-21.81	32.62	4.46	.13	.137
NWF-CLS Change	222	-11.75	22.21	222	-13.52	20.83	1.80	.08	.354
NWF-WWR Change	190	-3.96	9.29	190	-4.87	8.37	.91	.10	.297
ORF Change	251	-1.52	13.70	251	-2.65	14.07	1.13	.08	.337

Note: Measure names are the following: COMP (Acadience Reading Composite); LNF (Letter Naming Fluency); PSF (Phoneme Segmentation Fluency); NWF-CLS (Nonsense Word Fluency - Correct Letter Sounds); NWF-WWR (Nonsense Word Fluency - Whole Words Read); ORF (Oral Reading Fluency);

Cohort 3 Analyses

Students enrolled in Cohort 3 participated in the UPSTART Summer program during the summer after second grade, prior to third grade. Cohort 3 students and their comparison counterparts who did not use the UPSTART program were assessed at the end of the second grade and beginning of third grade on the Acadience Reading Oral Reading Fluency measure to determine the extent of students' oral reading fluency learning loss. While a regression model that included baseline scores, low-income status, racial group membership, and schools as covariates was significant and predicted 29% of the variability in oral reading fluency learning

loss (R^2=0.289), UPSTART Summer program use was not a significant predictor of learning loss prevention for Cohort 3 students. A summary of the regression analytic model is presented in Appendix B.

Table 9 presents the regression model adjusted means from the Oral Reading Fluency (ORF) change scores derived from the linear regression model for the treatment and control group, as well as the difference between the two groups, the standardized effect size, and the resulting p-value. The UPSTART Summer program did not have a significant impact in ameliorating rising third graders' summer learning loss as measured by oral reading fluency (Hedges's g = .03).

Table 9

Impact Analysis Results for UPSTART Summer Program Cohort 3 Participants

	Treatment Group			Comparison Group					
Measure	Unadj N	Model Adjusted Mean	Unadjusted Post-Test SD	Unadj N	Unadjusted Post-Test Mean	Unadjusted Post-Test SD	Impact estimate	Standardized difference	P value
ORF Change	222	-17.48	16.92	222	-18.05	16.24	.62	.04	.660

Note: Measure name is the following: ORF (Oral Reading Fluency)

Discussion

Adding to the body of research on educational technology, literacy instruction, and interventions to combat summer learning loss, results from this quasi-experimental study show that the UPSTART Summer program had a significant effect on combatting rural elementary students' learning loss over the summer between leaving kindergarten and rising to first grade. Specifically, rising UPSTART first graders had lower levels of summer literacy learning decay as measured by letter naming fluency (i.e., alphabet knowledge), the ability to read nonsense

word sounds (i.e., phonics), and an overall reading composite score compared to similar students who did not participate in the UPSTART Summer program. Skills in phonics, in particular, have been shown to help students' emerging reading ability better than all forms of other instruction, including whole language, and these effects are larger when phonics instruction begins prior to first grade (Ehri, Nunes, Stahl, & Willows, 2001).

There were, however, no significant differences between older UPSTART students who were rising into second and third grade and their comparison student counterparts after controlling for prior knowledge, school enrollment, household income, and race on measures of literacy learning loss over the course of a summer. It appears as though the program's benefits are more pronounced with the fundamental concepts of reading that require direct instruction such as recognizing letters or phonics, as opposed to complex skills like oral reading fluency measured by Acadience Reading in first through third grades that are hallmarks of more proficient reading and require multifaceted coordination of talents such as the ability to translate letters into sounds, recognize whole words and connections within and between sentences, and relate text to prior information (Fuchs, Fuchs, Hosp, & Jenkins, 2001).

Acknowledging calls from researchers to conduct systemic evaluation of the effectiveness of educational reading software (Grant et al., 2012), the current study provides evidence that the UPSTART Summer program has some merit in preparing young students for entry into first grade. Previous research of computer-based instructional programs has shown that benefits are linked with sufficient program use (Macaruso & Rodman, 2011) and additional analysis is needed to determine if outcomes are stronger for students who meet the requirements for minimum program use of the UPSTART Summer program compared to children who do not fulfill the recommended program use. If, for example, older children are using the UPSTART

program less as they encounter increasing demands of competing outside-the-home summer activities, one might expect fewer positive impacts on literacy outcomes. Further research of the UPSTART Reading program will explore more nuanced questions such as the potential differential effects of program usage and the additive effect of participating in the UPSTART program over multiple cohorts that may influence relationship between UPSTART Reading participation and the prevention of summer learning loss.

References

Addy, S., Englehardt, W., & Skinner, C. (2013). *Basic facts about low-income children: Children under 6 years, 2011*. New York, NY: National Center for Children in Poverty.

Bakle, B. R. (2015). *Summer learning loss: The influence of summer school programs on student achievement in language usage, math, and reading*. Doctoral dissertation abstract retrieved from ERIC database. (ED517121)

Blok, H., Oostdam, R., Otter, M. E., & Overmaat, M. (2002). Computer-assisted instruction in support of beginning reading instruction: A review. *Review of educational research*, *72*(1), 101-130.

Borman, G. D., Overman, L. T., Fairchild, R., Boulay, M., & Kaplan, J. (2004). Can a multiyear summer program prevent the accumulation of summer learning losses? In G. D. Borman & M. Boulay (Eds.), *Summer Learning: Research, Policies, and Programs* (pp. 233-253). New York, NY: Routledge.

Cooper, H., Nye, B., Charlton, K., Lindsay, J., & Greathouse, S. (1996). The Effects of Summer Vacation on Achievement Test Scores: A Narrative and Meta-Analytic Review. *Review of Educational Research*, *66*(3), 227–268.

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.

Foster, M. E., Anthony, J. L., Clements, D. H., Sarama, J., & Williams, J. M. (2016). Improving mathematics learning of kindergarten students through computer-assisted instruction. *Journal for Research in Mathematics Education*, *47*(3), 206-232.

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*(3), 239-256.

Good, R. H., & Kaminski, R. A. (2011). *Acadience Reading: Assessment manual*. Dynamic Measurement Group, Inc: Eugene, OR.

Gupta, S. K. (2011). Intention-to-treat concept: a review. *Perspectives in clinical research*, *2*(3), 109.

Grant, A., Wood, E., Gottardo, A., Evans, M., Phillips, L., & Savage, R. (2012). Assessing the content and quality of commercially available reading software programs: Do they have the fundamental structures to promote the development of early reading skills in children? *NHSA Dialog*, *15*(4), 319-342.

King, G., & Nielsen, R. (2019). Why propensity scores should not be used for matching. *Political Analysis.* Retrieved from http://j. mp/2ovYGsW

Khan, K., Justice, L., & Jiang, H. (2016). *Profiles of school readiness among rural Appalachian children from low-income homes*. Columbus, OH: Crane Center for Early Childhood Research and Policy, The Ohio State University.

Li, X. & Atkins, M. S. & Stanton, B. (2006). Effects of home and school computer use on school readiness and cognitive development among Head Start children: A randomized controlled pilot trial. *Merrill-Palmer Quarterly* 52(2), 239-263.

Iacus, S. M., King, G., & Porro, G. (2012). Causal inference without balance checking: Coarsened Exact Matching. *Political Analysis*, *20*(1), 1-24.

Macaruso, P. & Rodman, A. (2011). Efficacy of computer-assisted instruction for the

development of early literacy skills in young children. *Reading Psychology*, *32*(2), 172-196.

Miller, P., Votruba-Drzal, E., & Setodji, C. M. (2013). Family income and early achievement across the urban–rural continuum. *Developmental Psychology*, *49*(8), 1452.

National Early Literacy Panel. (2008). *Developing early literacy: Report of the National Early Literacy Panel*. Washington, DC: National Institute for Literacy. Retrieved from http://www.nifl.gov/earlychildhood/NELP/NELPreport.html

Nores, M., & Barnett, W. S. (2014). *Access to high quality early care and education: Readiness and opportunity gaps in America*. CEELO Policy Report. New Brunswick, NJ: National Institute for Early Education Research. Retrieved from http://ceelo.org/wp-content/uploads/2014/05/ceelo_policy_report_access_quality_ece.pdf

Paris, S. G., Pearson, P. D., Cervetti, G., Carpenter, R., Paris, A. H., DeGroot, J., … Bashore-Berg, T. (2004). Assessing the effectiveness of summer reading programs. In G. D. Borman & M. Boulay (Eds.), *Summer Learning: Research, Policies, and Programs* (pp. 121-161). New York, NY: Routledge.

Phillips, R., Harper, S., & Gamble, S. (2007). Summer programming in rural communities: Unique challenges. *New Directions for Youth Development*, *2007*(114), 65-73.

Tollerson, K., & Guckert, S. (2016). *Eliminating the summer slide: Enabling students to make academic growth over the summer with iPads*. Chicago, IL: eSpark Learning.

Appendix A

Table A.1

Cohort 1 – Letter Naming Fluency (LNF) Matched Sample Demographics

	UPSTART Summer Program (N = 297)		Comparison Students (N =297)	
	N	%	N	%
Gender				
Male	149	50.17	149	50.17
Female	148	49.83	148	49.83
Race				
White	284	95.62	284	95.62
Hispanic	6	2.02	6	2.02
Other	7	2.36	7	2.36
Language				
English	289	97.31	291	97.98
Spanish	3	1.01	3	1.01
Other	5	1.68	3	1.01
English Language Learner				
Yes	289	97.31	289	97.31
No	8	2.69	8	2.69
Special education student				
Yes	12	4.76	12	4.76
No	240	95.24	240	95.24
Low income student				
Yes	120	47.62	120	47.62
No	132	52.38	132	52.38
Title 1 student				
Yes	230	91.27	230	91.27
No	22	8.73	22	8.73

Table A.2

Cohort 1 – Phoneme Segmentation Fluency (PSF) Matched Sample Demographics

	UPSTART Summer Program (N = 318)		Comparison Students (N = 318)	
	N	%	N	%
Gender				
Male	165	51.89	165	51.89
Female	153	48.11	153	48.11
Race				
White	295	92.77	295	92.77
Hispanic	13	4.09	13	4.09
Other	10	3.14	10	3.14
Language				
English	304	95.60	303	95.28
Spanish	5	1.57	5	1.57
Other	9	2.83	10	3.14
English Language Learner				
Yes	13	4.09	13	4.09
No	305	95.91	305	95.91
Special education student				
Yes	10	96.38	10	3.62
No	266	3.62	266	96.38
Low income student				
Yes	136	50.72	136	50.72
No	140	49.28	140	49.28
Title 1 student				
Yes	247	89.49	247	89.49
No	29	10.51	29	10.51

Table A.3

Cohort 1 – Nonsense Word Fluency - Correct Letter Sounds (NWF-CLS) Sample Demographics

	UPSTART Summer Program (N = 300)		Comparison Students (N = 300)	
	N	%	N	%
Gender				
Male	153	51.00	153	51.00
Female	147	49.00	147	49.00
Race				
White	282	94.00	282	94.00
Hispanic	10	3.33	10	3.33
Other	8	2.66	8	2.66
Language				
English	289	96.33	290	96.67
Spanish	4	1.33	4	1.33
Other	7	2.33	6	2.00
English Language Learner				
Yes	9	3.00	9	3.00
No	291	97.00	291	97.00
Special education student				
Yes	11	4.30	11	4.30
No	245	95.70	245	95.70
Low income student				
Yes	116	45.31	116	45.31
No	140	54.69	140	54.69
Title 1 student				
Yes	225	87.89	225	87.89
No	31	26.17	31	26.17

Table A.4

Cohort 1 – Nonsense Word Fluency - Whole Words Read (NWF-WWR) Sample Demographics

	UPSTART Summer Program (N = 363)		Comparison Students (N = 363)	
	N	%	N	%
Gender				
Male	188	51.79	188	51.79
Female	175	48.21	175	48.21
Race				
White	334	92.01	334	92.01
Hispanic	15	4.13	15	4.13
Other	14	3.86	14	3.86
Language				
English	348	95.87	347	95.59
Spanish	6	1.65	6	1.65
Other	9	2.48	10	2.75
English Language Learner				
Yes	13	3.58	13	3.58
No	350	96.42	350	96.42
Special education student				
Yes	26	8.25	26	8.25
No	289	91.75	289	91.75
Low income student				
Yes	152	48.25	152	48.25
No	163	51.75	163	51.75
Title 1 student				
Yes	235	88.26	235	88.26
No	37	11.75	37	11.75

Table A.5

Cohort 1 – Composite Sample Demographics

	UPSTART Summer Program (N = 333)		Comparison Students (N = 333)	
	N	%	N	%
Gender				
Male	169	50.75	169	50.75
Female	164	49.25	164	49.25
Race				
White	318	95.50	318	95.50
Hispanic	9	2.70	9	2.70
Other	6	1.80	6	1.80
Language				
English	325	97.60	326	97.90
Spanish	4	1.20	3	0.90
Other	4	1.20	4	1.20
English Language Learner				
Yes	326	97.90	326	97.90
No	7	2.10	7	2.10
Special education student				
Yes	11	3.82	11	3.82
No	277	96.18	277	96.18
Low income student				
Yes	127	44.10	127	44.10
No	161	55.90	161	55.90
Title 1 student				
Yes	255	88.55	255	88.55
No	33	11.46	33	11.46

Table A.6

Cohort 2 – Nonsense Word Fluency- Whole Words Read (NWF-WWR) Matched Sample Demographics

	UPSTART Summer Program (N = 190)		Comparison Students (N =190	
	N	%	N	%
Gender				
Male	93	48.95	93	48.95
Female	97	51.05	97	51.05
Race				
White	183	96.32	183	96.32
Hispanic	5	2.63	5	2.63
Other	2	1.05	2	1.05
English Language Learner				
Yes	4	2.11	4	2.11
No	186	97.89	186	97.89
Special education student				
Yes	14	7.37	14	7.37
No	176	92.63	176	92.63
Low income student				
Yes	89	46.84	89	46.84
No	101	53.16	101	53.16
Title 1 student				
Yes	165	86.84	165	86.84
No	25	13.16	25	13.16

Table A.7

Cohort 2 – Nonsense Word Fluency- Correct Letter Sounds (NWF-CLS) Matched Sample Demographics

	UPSTART Summer Program (N = 222)		Comparison Students (N = 222)	
	N	%	N	%
Gender				
Male	109	49.10	109	49.10
Female	113	50.90	113	50.90
Race				
White	210	94.59	210	94.59
Hispanic	7	3.15	7	3.15
Other	5	2.25	5	2.25
English Language Learner				
Yes	7	3.15	7	3.15
No	215	96.85	215	96.85
Special education student				
Yes	16	7.21	16	7.21
No	206	92.79	206	92.79
Low income student				
Yes	105	47.30	105	47.30
No	117	52.70	117	52.70
Title 1 student				
Yes	193	86.94	193	86.94
No	29	13.06	29	13.06

Table A.8

Cohort 2 – Oral Reading Fluency (ORF) Sample Demographics

	UPSTART Summer Program (N = 251)		Comparison Students (N = 251)	
	N	%	N	%
Gender				
Male	120	47.81	120	47.81
Female	131	52.19	131	52.19
Race				
White	240	95.62	240	95.62
Hispanic	7	2.79	7	2.79
Other	4	1.60	4	1.60
English Language Learner				
Yes	6	2.39	6	2.39
No	245	97.61	245	97.61
Special education student				
Yes	19	7.57	19	7.57
No	232	92.43	232	92.43
Low income student				
Yes	107	42.63	107	42.63
No	144	57.37	144	57.37
Title 1 student				
Yes	186	87.65	186	87.65
No	31	12.35	31	12.35

Table A.9

Cohort 2 – Composite Sample Demographics

	UPSTART Summer Program (N = 235)		Comparison Students (N = 235)	
	N	%	N	%
Gender				
Male	108	45.96	108	45.96
Female	127	54.04	127	54.04
Race				
White	226	96.17	226	96.17
Hispanic	7	2.98	7	2.98
Other	2	0.85	2	0.85
English Language Learner				
Yes	6	2.55	6	2.55
No	229	97.45	229	97.45
Special education student				
Yes	221	94.04	221	94.04
No	14	5.96	14	5.96
Low income student				
Yes	97	41.28	97	41.28
No	138	58.72	138	58.72
Title 1 student				
Yes	206	87.66	206	87.66
No	29	12.34	29	12.34

Table A.10

Cohort 3 – Oral Reading Fluency Matched Sample Demographics

	UPSTART Summer Program (N = 222)		Comparison Students (N = 222)	
	N	%	N	%
Gender				
Male	108	48.65	108	48.65
Female	114	51.35	114	51.35
Race				
White	216	97.30	216	97.30
Hispanic	5	2.25	5	2.25
Other	1	0.45	1	0.45
English Language Learner				
Yes	2	0.90	2	0.90
No	220	99.10	220	99.10
Special education student				
Yes	11	4.95	11	4.95
No	211	95.05	211	95.05
Low income student				
Yes	78	35.14	78	35.14
No	144	64.86	144	64.86
Title 1 student				
Yes	205	92.35	205	92.35
No	17	7.66	17	7.66

Appendix B

Table B.1

Pre-program baseline and post-program outcome correlations

	Pre-Test Measure	Post-Test Measure	Correlation
Cohort 1			
	comp	comp	r=0.802, p<0.001
	lnf	lnf	r=0.749, p<0.001
	psf	psf	r=0.558, p<0.001
	nwf_cls	nwf_cls	r=0.799, p<0.001
	nfw_wwr	nfw_wwr	r=0.755, p<0.001
Cohort 2			
	comp	comp	r=0.926, p<0.001
	dorf	dorf	r=0.937, p<0.001
	nwf_cls	nwf_cls	r=0.827, p<0.001
	nfw_wwr	nfw_wwr	r=0.812, p<0.001
Cohort 3			
	dorf	**dorf**	**r=0.936, p<0.001**

Note: Measure names are the following: LNF (Letter Naming Fluency); PSF (Phoneme Segmentation Fluency); NWF-CLS (Nonsense Word Fluency - Correct Letter Sounds); NWF-WWR (Nonsense Word Fluency - Whole Words Read); ORF (Oral Reading Fluency); COMP (Acadience Reading Composite).

Appendix C

Table C.1

Regression Analysis of Predictors of Cohort 1 Phenome Segmentation Fluency Learning Loss (N=551)

Variable	*B*	*SE B*	*β*
(Constant)	25.29	4.71	
Treatment	1.26	.86	.05
Baseline	-0.63	.06	-.51**
Low-income status	.10	1.10	.004
White racial group	4.11	2.34	.08
R^2	.43		
F	8.02**		

* p<0.05; ** p<0.01

Table C.2

Regression Analysis of Predictors of Cohort 2 Nonsense Word Fluency-Correct Letter Sounds Learning Loss (N=444)

Variable	*B*	*SE B*	*β*
(Constant)	-9.47	8.73	
Treatment	1.80	1.94	.04
Baseline	-.15	.03	-.24**
Low-income status	-1.88	2.58	-.04
White racial group	6.33	5.24	.07
R^2	0.20		
F	1.99**		

* p<0.05; ** p<0.01

Table C.3

Regression Analysis of Predictors of Cohort 2 Nonsense Word Fluency- Whole Words Read Learning Loss (N=380)

Variable	*B*	*SE B*	*β*
(Constant)	6.55	4.78	
Treatment	.88	.84	.05
Baseline	-.17	.04	-.27**
Low-income status	-1.17	1.11	-.07
White racial group	-5.15	3.25	-.11
R^2	.24		
F	2.24**		

* p<0.05; ** p<0.01

Table C.4

Regression Analysis of Predictors of Cohort 2 Oral Reading Fluency Learning Loss (N=502)

Variable	*B*	*SE B*	*β*
(Constant)	20.41	6.23	
Treatment	1.16	1.20	.04
Baseline	-.07	.03	-.14**
Low-income status	.08	1.58	.00
White racial group	-9.20	3.68	-.14*
R^2	.16		
F	1.58**		

* p<0.05; ** p<0.01

Table C.5

Regression Analysis of Predictors of Cohort 2 Composite Learning Loss (N=470)

Variable	*B*	*SE B*	*β*
(Constant)	-14.50	2.98	
Treatment	4.45	2.30	.07
Baseline	-.04	.03	-.09
Low-income status	3.84	3.72	.06
White racial group	3.10	9.55	.018
R^2	.15		
F	1.57*		

* p<0.05; ** p<0.01

Table C.6

Regression Analysis of Predictors of Cohort 3 Oral Reading Fluency Learning Loss (N=444)

Variable	*B*	*SE B*	*β*
(Constant)	-4.69	6.34	
Treatment	.62	1.40	.02
Baseline	-.10	.03	-.21**
Low-income status	1.34	1.77	.04
White racial group	-4.17	5.19	-.04
R^2	.29		
F	3.42**		

* p<0.05; ** p<0.01