Utah UPSTART Program Evaluation
Program Impacts on Early Literacy

Fourth Year Results

Cohort 4 Technical Report

February 2014

Submitted by:

Evaluation and Training Institute
11500 W. Olympic Blvd.
Suite 626
Los Angeles, CA 90064

All correspondence should be sent to:
Jon Hobbs, Ph.D.
jhobbs@eticonsulting.org

Contents

EXECUTIVE SUMMARY **1**

Evaluation Design 1
Data Collection 2
Data Analysis 2
Results 3

INTRODUCTION **6**

BACKGROUND: THE C3 EVALUATION **6**

C4 EVALUATION DESIGN **7**

Research Questions 8
Outcome Measures 10
Data Collection 11

PRESCHOOL DATA ANALYSIS **12**

Attrition Analysis 12
Analysis Strategy: Do UPSTART children have better literacy skills at Kindergarten than control group children? 13
Analysis Strategy: Do UPSTART children show stronger literacy growth rates from preschool to Kindergarten than control group children? 14
Analysis of Implementation Time 14

RESULTS **15**

UPSTART Implementation 15
UPSTART Enrollment 15
UPSTART Equipment Provided 16
UPSTART Graduates 17
UPSTART Usage 17
UPSTART Usage and Literacy Outcomes 21
UPSTART Literacy Outcomes 29
Do UPSTART students have better literacy skills at Kindergarten than control students? 29
Do UPSTART students show stronger literacy growth rates from preschool to kindergarten than control students? 37

SUMMARY AND CONCLUSIONS **40**

Data Collection 40
Data Analysis 40
UPSTART Implementation Results 41
UPSTART Impact Results 42

APPENDICES **44**

Executive Summary

Established as a pilot demonstration project by the Utah state legislature, UPSTART uses educational technology in a home-based approach to develop the school readiness skills of preschool children. The Waterford Institute provided documentation for a fourth-year UPSTART enrollment of 1,250 children. A majority (63%) of the 1,250 preschool children that enrolled in the fourth year of UPSTART were from low income families, according to data provided by the Waterford Institute. Slightly more girls (51%) were enrolled than boys (49%). In terms of ethnicity, the vast majority of the Cohort 4 (C4) enrollment w Caucasian (76%) with 20% being Hispanic. The ethnicity of the remaining 4% of the C4 enrollment was of Asian, African, Native American, Pacific Islander, or unknown origin.

Evaluation Design

The evaluation of UPSTART's fourth year of implementation used a pretest-posttest control group design to assess the program's impact on developing the children's early literacy skills in preschool. Other objectives included documenting the extent to which participants used the computerized curriculum; establishing the relationship between curriculum usage and literacy outcomes; and documenting the degree to which the C4 participants met the program's curriculum usage criteria, otherwise referred to as the "graduation" rate.

An ordinary least squares (OLS) regression approach was used to estimate posttest differences in the development of literacy skills between a sample of UPSTART participants (the treatment group) and a group of similar nonparticipants (the control group) in the year prior to enrollment in kindergarten. The children were measured on two tests of early literacy skills: the Brigance Inventory of Early Development and the Bader Reading and Language Assessment. Covariates used in the analyses to adjust for initial between group differences included pretest scores on the respective tests and selected demographic characteristics that independently influenced posttest outcomes. Additionally, differences between the treatment and control groups in their growth rates on the two tests were examined.

The effect of UPSTART usage on literacy skill development was examined for UPSTART participants using an analysis of covariance in which usage levels were split into quartiles based on the usage distribution of the preschool analysis sample. The statistical models controlled for the child's initial level of literacy development, as measured by the pretest score on each of the two respective tests (i.e., the Brigance and the Bader). The effect of UPSTART usage on literacy skill development was assessed by comparing the adjusted mean posttest performance on the Brigance and the Bader at each usage quartile with the fourth quartile of usage. The purpose of the analysis was to determine whether literacy development increased with increasing usage of the UPSTART curriculum.

Descriptive statistics were also computed to describe the population of students that enrolled in the fourth year of UPSTART (i.e., Cohort 4). The descriptors included student demographics, the equipment that C4 participants received, hours of UPSTART curriculum usage, and the graduation status of C4 students. Graduates were participants who met the program's curriculum usage criteria.

Data Collection

Two hundred and fifty four year-old children were recruited for the C4 study. Two groups of students were recruited: treatment (who had enrolled in UPSTART during Year 4 of the program), and control (who had not enrolled in the UPSTART program). Families were given a financial incentive to participate in the evaluation and participation was optional for each group. Treatment student were randomly selected from a list of UPSTART participants. Control families were recruited using a variety of methods, including: distributing flyers at day care facilities, including public pre-kindergarten locations, referral sampling and advertisements.

After two waves of testing were completed, 220 cases (children) had matched data from pre-program and post-program measures. Of the 220 cases used in the analyses, there were 117 treatment group children and 103 control group children. The children's parents were given an intake questionnaire (see Appendix A) at the time their children were pretested on the Brigance and Bader prior to their participation in UPSTART. The children were posttested on the Brigance and Bader in July and August of 2013. The intake interview certified that 203 of the recruited children had no prior exposure to the UPSTART online curriculum. Complete pretest and posttest Brigance data were obtained and analyzed for all 203 of these children (101 treatment group children and 102 control children). On the Bader, complete pretest and posttest data were obtained for 172 children (79 treatment children and 93 control children). Measurement attrition on the Bader (16%) was caused primarily by children not being able to score on the pretest. There was no measurement attrition with the Brigance sample.

Data Analysis

To determine whether UPSTART children had developed better literacy skills at kindergarten entry compared to control group children, group equivalence on the pretests was first examined using independent sample t-tests. Relationships between the demographics and posttest scores were then examined using correlation analyses. Next, posttest differences between the treatment and control groups were examined for both the Brigance and Bader. Finally, posttest differences were re-examined by adjusting for initial differences between the treatment and control groups with the use of multiple regression analysis. The regressions used a hierarchical block design in which the pretest and a set of demographic covariates were entered first, followed by the treatment-control group indicator.

To determine whether UPSTART students exhibited stronger literacy growth rates relative to control students from preschool to kindergarten, paired sample t-tests were run to obtain gain scores (i.e., posttest minus pretest) for the matched Brigance and Bader treatment groups on the total test and each of the subtests. Statistically significant growth rates were determined by examining confidence intervals for the treatment and control groups for each test measure at the 99% confidence interval.

The effect of UPSTART usage on reading proficiency was examined for UPSTART participants using an analysis of covariance (ANCOVA) in which usage levels were split into quartiles based on the usage distribution of the preschool analysis sample. The final ANCOVA models statistically controlled for initial literacy skills as measured by the pretest on each respective measure. The ANCOVA analyses estimated the effects of usage at quartiles one through three compared with usage at the fourth quartile, controlling for initial levels of literacy development.

Results

As in previous years, most of the C4 participants (69%) received a computer drive with the UPSTART curriculum loaded on it. Approximately 11% of the fourth year participants received a computer loan and a free Internet subscription to help them access the UPSTART curriculum. Approximately another 8% of the C4 participants were loaned a personal computer to use at home while participating in UPSTART. The remaining 12% of the fourth year participants were provided with various combinations of educational technology to enable them to access the UPSTART curriculum, including wireless and cellular devices.

Findings about UPSTART usage are summarized below.

- The C4 preschool analysis samples had a mean of approximately 76-78 hours of UPSTART curriculum use over the fourth year of the project. This compares with an average of approximately 75 hours of instruction for program "graduates" and approximately 71 hours of instruction for all students enrolled in UPSTART in the fourth year of the program.

- UPSTART curriculum use was significantly correlated with literacy skills at the beginning of kindergarten as measured by the Brigance ($r = .49$) and the Bader ($r = .44$). Controlling for initial literacy skills, the correlation of UPSTART use with kindergarten outcomes was somewhat lower: $r=.35$ for the Brigance sample and $r=.38$ for the Bader sample.

- Controlling for initial levels of literacy development, UPSTART use accounted for 13%-15% of the variance in literacy skills measured by the Bader and Brigance posttests respectively at kindergarten entry.

- There is an apparent linear trend in literacy development with increasing use of the UPSTART curriculum.

- The UPSTART graduation rate continued to hold at 94% in Year 4. This is the same level of program attainment as realized in Year 3.

- UPSTART graduate status in the fourth year of the program was significantly correlated with hours of instruction (r=.55) as well as with literacy outcomes measured by the Brigance (r=.31) and the Bader (r=.35) at the beginning of kindergarten.

- The UPSTART graduation rate has leveled off at 94% in Years 3 and 4. This indicates that UPSTART is continuing to make very good progress in achieving the curriculum usage goals set for program implementation.

Findings about UPSTART's impact on literacy skills measured at the beginning of kindergarten are summarized below.

- UPSTART participation in the fourth year of the program generally show small to medium size effects on improving the phonics skills of young preschool children. C4 effects as measured by the Brigance at the beginning of kindergarten were observed on the Visual Discrimination subtest, the Letter Sounds subtest, the Basic Pre-Primer Vocabulary subtest and on the overall Brigance. The largest impact observed in UPSTART's fourth year of operation was in the development of vocabulary. These findings replicates results found in the third year evaluation.

- Overall, treatment group children scored an average of 29 points higher on the Brigance posttest (regression adjusted) compared to control group children. This finding is similar to the overall impact of UPSTART as found in the third year of the evaluation.

- Overall, UPSTART achieved medium size effects on improving the phonological awareness skills of participants in Year 4 of the program as measured by the Total Bader assessment. In terms of specific domains of phonological awareness, treatment group effects were observed on two of the three Bader subtests: Phoneme Blending and Phoneme Segmentation.

- On average, treatment group children scored an average of almost six points higher on the Bader posttest (regression adjusted) compared to control group children. This finding is similar to the overall impact of UPSTART as found in the third year of the evaluation.

- UPSTART participants showed significantly stronger growth rates compared to control group children on the *Total Brigance* and four of the ten subtests, including: *Vocabulary, Survival Sight Words, Letter Sounds and Reciting the Alphabet*. These results are similar to the findings about Brigance growth rates in the third year evaluation.

- UPSTART participants showed significantly stronger growth rates compared to control children on the *Total Bader* and two of the three Bader subtests, specifically *Phoneme Blending* and *Phoneme Segmenting*.

- Program effects for improvements in knowledge of letter sounds and vocabulary development are two of the early reading skills that the UPSTART curriculum focuses on. Similarly, program effects for phonemic blending and segmenting reflect UPSTART's focus on developing children's phonological awareness.

Based on the fourth year results, the evidence suggests that UPSTART's use of education technology in a home based approach has considerable merit for facilitating the development of school readiness in young preschool children.

Introduction

UPSTART is a pilot project established by the Utah state legislature that uses a home-based education technology approach to develop the school readiness skills of preschool children. In its fourth year of operation during the 2012-13 school year, the project's implementation contractor – the Waterford Institute – enrolled 1,250 preschool children and provided them a game formatted program of early literacy instruction delivered by personal computers and the Internet, designed to prepare them academically for kindergarten. The 1,250 children enrolled in the fourth year cohort, hereafter referred to as C4, participated in the UPSTART program from August 2012 through June 2013.

The evaluation of UPSTART's fourth year of implementation used a pretest-posttest control group design to assess the program's impact on developing the children's early literacy skills in preschool. Other objectives included (a) documenting the extent to which participants used the computerized curriculum; (b) establishing the relationship between curriculum use and literacy outcomes; and (c) documenting the program's completion or "graduation" rate as measured by the proportion of the enrollment that met the criteria established for use of the program's curriculum.

Slightly more C4 girls (51%) were enrolled than boys (49%). In terms of ethnicity, the vast majority (76%) of the C4 enrollment was Caucasian, with 20% of the children being of Hispanic origin. The ethnicity of the remaining 4% of the C4 enrollment was composed of children from African, Asian, Native American, Pacific Islander and unknown backgrounds. The primary language spoken by the vast majority of the C4 children was English (85%). Approximately 14% of the C4 children spoke Spanish and 1% spoke other languages. Six percent of the C4 children had a diagnosed disability, mostly speech impairments.

A majority (63%) of the 1,250 preschool children that enrolled in the fourth year of UPSTART were from low-income families, according to data provided by the Waterford Institute. Most commonly, the C4 parents had some college (31%) or had achieved a bachelor's degree (38%). The vast majority of the C4 parents were married (approximately 90%).

Background: The C3 Evaluation

Previously, a pretest-posttest control group design was used in evaluating the third year (C3) of UPSTART. In the C3 evaluation, children were pretested in the fall of their preschool year (2011) and posttested in the spring of the preschool year (2012) using two assessments: the Brigance Inventory of Early Development (the Brigance) and the Bader Reading and Language Assessment (the Bader). The Brigance was selected as a measure of early language and academic

skills development and the Bader was selected as a measure of phonological awareness. In the C3 evaluation, complete pretest and posttest data were obtained for a sample of 112-129 UPSTART children (the treatment group), depending on the test, and 120-130 non-UPSTART children (the control group), depending on the test. *The same evaluation design was used in the C4 evaluation during the 2012-13 preschool year,* but with smaller sample sizes.

In the third year of UPSTART, 94% of the C3 enrollment met the program's curriculum usage standards and were considered to have "graduated" from UPSTART. Hours of instruction logged in the UPSTART curriculum in the third year was found to be significantly and positively correlated with literacy achievement on the Brigance (r=.40) and the Bader (r=.20). C3 students logged an average of 67 hours of instruction in the program during 2011-12. Further analysis controlling for initial levels of literacy development showed that increasing levels of curriculum use were related to overall early literacy achievement as measured by the Brigance. However, this relationship could not be established for the Bader in the C3 evaluation, possibly due to it being a more limited assessment scale with a weaker measured relationship to early literacy development.

The C3 evaluation statistical models controlled for initial levels of literacy development and further showed that participation in UPSTART had a medium size effect on improving the treatment group children's development of early literacy skills compared to control children. Overall, treatment group children scored approximately 28 points higher on the Brigance posttest compared to control children. The largest effect measured by the Brigance was on vocabulary acquisition.

C3 statistical modeling showed a strong program impact for improving the phonological awareness skills measured by the Bader, which favored the treatment group children by over 6 points compared to the control group children. The largest impact measured by the Bader was on improving the phoneme segmentation ability of UPSTART students.

C4 Evaluation Design

The Cohort 4 evaluation continued to use the quasi-experimental research design variant of the nonequivalent comparison group design described previously. Specifically, the design implemented in the C4 evaluation (as in C2 and C3) used a treatment group and an untreated comparison group, with both pretest and posttest data collected on the same children over a 12 month interval during the year prior to enrollment in kindergarten. The design is diagramed below. NR indicates that the evaluation was a quasi-experiment since the children were not randomly assigned to groups.

The C4 study recruited 220 four year-old children; 117 treatment group children who had enrolled in UPSTART for Year 4 of the program and 103 nonparticipating control group children. The control children were recruited from local preschools in Utah and the treatment children were recruited from families enrolling in UPSTART.

In the diagram below, T stands for 4 year-old children who received the UPSTART preschool program during its third year of operation, and C stands for 4 year-old comparison group children who did not participate in UPSTART. The "X" indicates that the UPSTART children received the Waterford Early Learning Program prior to kindergarten and that the children from the control group did not. O1 indicates pretest measurements taken in the fall of 2012 for C4 treatment and control group children. O2 indicates posttest measurements taken in June and July of 2013 for the C4 treatment and control group children.

NR	T	01	X	O2
NR	C	O1		02

The use of both a pretest and a comparison group facilitates our ability to examine potential threats to validity that could jeopardize a clear interpretation of the results.[1] Because the study is not a randomized control trial, the groups are nonequivalent by definition, and consequently selection bias can be assumed to operate to some degree in some manner. The pretest 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). The pretest also allows us to examine the nature and degree of attrition in the study and whether it differentially affects one group more than the other.

Research Questions

We hypothesized that if UPSTART has no effect on improving early literacy skills, then the preschool children who participated in UPSTART – the treatment group – would be expected to perform at the same level as the comparison group on posttest measures of early literacy development at the beginning of kindergarten. If UPSTART does have an effect on improving early literacy, then the treatment group should perform significantly better than the comparison group on the posttest at the beginning of kindergarten. For purposes of triangulation, we also wanted to take a slightly different look at the data by examining growth rates from pretest to posttest. If UPSTART shows stronger literacy growth rates, then the treatment group would be expected to show greater gain scores (posttest score minus pretest score) relative to the comparison group on the various subtests and total test scores.

[1] See Shadish, Cook, and Campbell (2002). *Experimental and Quasi-Experimental Designs for Generalized Causal Inference*. Boston: Houghton Mifflin Company.

With respect to concerns for school readiness, our research questions for the C4 evaluation study were as follows:

RQ1: D*o UPSTART students have better early literacy skills at kindergarten compared to control group students?*

If the answer is yes, then we would expect to see:

T > C @ O2 (controlling for differences at O1)

If the answer is no, then we would expect to see:

T = C @ O2

RQ2: D*o UPSTART students show stronger literacy growth rates from preschool to kindergarten compared to control group students?*

If the answer is yes, then we would expect to see:

T > C @ O2- O1 (growth)

If the answer is no, then we would expect to see:

T = C @ O2-O1 (growth)

In the preschool analysis, the outcomes of interest were measures of early literacy skills relevant to emerging readers such as phonological awareness, letter recognition, letter sound knowledge and vocabulary development.

The Utah State Office of Education (USOE) and the Utah State Legislature were also interested in outcomes related to the implementation of UPSTART. Research questions along this line included:

RQ3: *What was the extent of UPSTART curriculum use in terms of the amount of exposure per participant, as measured in minutes or hours of instruction per week?*

RQ4: *What percent of the participants completed the full implementation program (i.e., "graduated" as defined by the Waterford Institute)?*

RQ5: *How does the level of UPSTART curriculum use relate to reading readiness outcomes?*

Data for research questions 3 and 4 were obtained from records maintained by the Waterford Institute and are answered in this report by descriptive statistics. The answer to Research Question 5 was derived from statistical analyses of the relationship between exposure to the

computer-assisted program of instruction (measured by program records documenting minutes of computer use for each enrolled student) and the measured literacy outcomes of interest.

Outcome Measures

The reading skills taught by the Waterford Early Learning Program at Level 1 of the curriculum[2] include:

- Phonological Awareness: phonemic segmenting and blending.
- Phonics: letter name knowledge, sound knowledge, and word reading.
- Comprehension and Vocabulary: vocabulary knowledge.
- Language Concepts: oral reading fluency.

The Brigance. The Brigance Inventory of Educational Development was selected as an early literacy measure of phonics and vocabulary knowledge and as a measure of pre-kindergarten academic and cognitive skills. Ten of the Brigance scales were administered from the *language development and academic/cognitive domains*, as described below.

The Brigance language development scales included the:

- *Expressive Objects subtest*: the child is asked to name pictures shown by an assessor. (Total possible subtest score = 27)
- *Receptive Objects subtest*: the child is asked to point to pictures named by an assessor. (Total possible subtest score = 27)
- *Expressive Grammar subtest*: the child is assessed on the ability to use plural *s*, *ing*, prepositions, and interpret and talk about an illustration. (Total possible subtest score = 12)

The Brigance academic and cognitive literacy scales included the:

- *Visual Discrimination* subtest: the assessment focuses on the child's ability to identify similarities and differences between forms, uppercase letters, lowercase letters, and words. (Total possible subtest score = 20)
- *Recites Alphabet* subtest: the child is asked to recite the alphabet. (Total possible subtest score = 26)
- *Lowercase Letter Knowledge* subtest: the child is asked to name and recognize (point to) lower case letters presented by an assessor. (Total possible subtest score = 52)
- *Sounds of Lowercase Letters* subtest: the child is assessed on the ability to produce sounds of lowercase letters. (Total possible subtest score = 26)
- *Auditory Discrimination* subtest: the assessment focuses on the child's ability to identify if two words sound the same or different. (Total possible subtest score = 10)

[2] Level One is the beginning point of the curriculum where the preschool child begins as a nonreader and is introduced to skills designed to teach the child to read.

- *Survival Sight Words* subtest: the assessment focuses on the child's ability to read survival sight words that appear on signs in public places. (Total possible subtest score = 16)
- *Basic Pre-Primer Vocabulary* subtest: the assessment focuses on the child's ability to read basic vocabulary words found in pre-primer reading programs. (Total possible subtest score = 24)
- *Total Brigance:* sum of the language and cognitive subtest scores. (Total possible score = 240)

As shown above, the version of the Brigance used in the C4 evaluation is a fairly comprehensive early literacy assessment comprised of 10 subtests in which the total test ranges from a score of zero to a score of 240. The Brigance is weighted toward the academic/cognitive literacy domain which accounts for approximately 73% of the total test score.

The Bader. The Bader was selected as a measure of *phonological awareness*. The Bader is comprised of three subtests, as follows:

- *Rhyme Recognition*: the child is asked to say yes if a pair of words presented orally by the assessor end the same way or to say no if the word pair do not end the same. (Total possible subtest score = 10)
- *Phonemic Blending*: the child is presented with a sequence of phonemes and is asked to say the word they constitute. (Total possible subtest score = 8)
- *Phoneme Segmentation*: the child is presented with a word and is asked to say the word sounds that make up the word in correct sequence. (Total possible subtest score = 8)
- Total Bader: sum of the Bader subscale scores. (Total possible Bader score = 26)

As revealed above, the Bader employs a relatively narrow test scale and measures the child's phonological awareness, considered an important predictor of later reading ability. Phonological awareness involves the child's ability to detect the sound structure of spoken words at three levels: rhyming, syllables, and phonemes.

Data Collection

Two hundred and fifty four year-old children were recruited for the C4 study. Two groups of students were recruited: treatment (who had enrolled in UPSTART during Year 4 of the program), and control (who had not enrolled in the UPSTART program). Families were given a financial incentive to participate in the evaluation and participation was optional for each group. Treatment students were randomly selected from a list of UPSTART participants, and phone recruitment was used to setup testing appointments. Control families were recruited using a variety of methods, including: distributing flyers at low-income day care facilities, including public pre-kindergarten locations, referral sampling and advertisements. Control families contacted the evaluators and were given information about the evaluation. In addition, control

families were screened to ensure that control group characteristics were similar to those of the treatment group, and that no special learning needs existed.

The children's parents were given an intake questionnaire (see Appendix A) in the fall of 2012 at the time their children were pretested on the Brigance and Bader. The children were posttested on the Brigance and Bader in the summer of 2013.

Preschool Data Analysis

A preschool student data file was developed based on data collected from the intake questionnaire and from the pretest and posttest administrations of the Brigance and Bader. The final analysis file was based on the subset of children with matched pretest and posttest data, and who had not previously used the UPSTART computerized learning program as documented through the intake interview.

Attrition Analysis

Of the 220 preschool children recruited for the C4 evaluation study, it was determined through the intake interview that 17 of the children had previously used the UPSTART online learning program. This validity check resulted in these 17 cases being removed from the final analysis sample. All of the 203 cases that passed this validity check also provided complete Brigance testing data. However, 21 of these cases could not provide matched pretest-posttest data on the Bader. This resulted in a final Bader test sample of 172 children. The C4 attrition results are summarized in Table 1.

Table 1
C4 Attrition Summary

Attrition Indicator	*Treatment Group N*	*Control Group N*	*Total Sample N*
Recruited	117	103	220
Passed Validity Check	101	102	203
Matched Brigance Pre/Post	101	102	203
Matched Bader Pre/Post	79	93	172

The data in Table 1 show that the C4 evaluation started off with a recruitment base of 220 children. The validity check in the intake interview provided data suggesting that 17 of the 220 children probably had prior exposure to the UPSTART curriculum, which invalidated their participation in the evaluation. All 203 of the valid cases provided Brigance test data whereas only 172 of the 203 valid cases could provide matched pretest-posttest data from the Bader. The Bader measurement attrition rate was 15.3% and was primarily caused by the children not being able to score on the pretest. The measurement attrition reported for the Bader in Year 4 (15%) is substantially better than that reported in Year 3 (24%). Additionally, the lack of any

measurement attrition for the Brigance in Year 4 is a substantial improvement over last year's measurement attrition rate of 15%.

The final analysis in the C4 evaluation used the data collected from those children who passed the validity screening and were able to provide matched pretest and posttest data on the Brigance and the Bader. The principal effect of the observed study attrition may have been primarily to reduce statistical power for the Bader preschool analysis.

Analysis Strategy: Do UPSTART children have better literacy skills at Kindergarten than control group children?

The general strategy for determining whether there was an impact of the UPSTART preschool experience on young children's literacy skills was to compare a sample of program participants with a similar group of nonparticipants on Brigance and Bader posttest scores collected at the beginning of kindergarten. This strategy assumes that the two groups are initially similar on factors that influence the literacy skills measured at kindergarten. These factors could include initial differences between the groups on measured literacy skills (e.g., pretest scores) as well as demographic factors that differentiate the treatment and control groups (e.g., the mother's marital status) if they are significantly related to posttest performance.

If the treatment and control groups are essentially similar at the beginning of UPSTART on factors affecting posttest literacy outcomes of interest, then any observed differences on the posttest can be attributed to participation in UPSTART. Alternatively, if there is significant initial nonequivalence between the groups, then statistical adjustments to the posttest outcomes using regression analysis will be necessary in leveling the playing field so that more accurate and fair comparisons can be made.

The equivalence of the treatment and control groups in the final analysis samples were examined on the basis of the Brigance and Bader pretest scores and on the basis of those demographic characteristics that were significantly related to the posttests. Group equivalence on the pretests was examined using independent sample t-tests. Relationships between the demographics and the posttest scores were examined using correlation analyses.[3] Initial between group differences were, in fact, found on pretest measures of early literacy and for demographic characteristics affecting the total posttest scores on the Brigance and Bader. This necessitated a final set of analyses using multiple regression analysis.

Posttest differences between the treatment and control groups were first examined for both the Brigance and Bader using independent sample t-tests. Ultimately, posttest differences were re-

[3] It was necessary to transform a number of the demographic measures from nominal measures to scale measures by creating "dummy variables" on the basis of the dominant characteristics of the sample. For example, parent's marital status was transformed into whether the parent was married or not, or percent married.

examined by adjusting for initial differences between the treatment and control groups with the use of multiple regression analysis. The regressions used a hierarchical block design in which the pretest AND a set of demographic covariates were entered first, followed by the treatment-control group comparison. Effect size estimates are graphically presented for all posttest differences between the treatment and control groups on the Brigance and Bader.

Analysis Strategy: Do UPSTART children show stronger literacy growth rates from preschool to Kindergarten than control group children?

To determine whether UPSTART students show stronger literacy growth rates from preschool to kindergarten compared to control students, paired sample t-tests were run to compare pretest and posttest gain scores for the matched Brigance and Bader *treatment groups* on the total test and each of the subtests. The same analysis was performed with the Brigance and Bader matched *control groups*. Statistically significant growth rates were determined by examining confidence intervals for the treatment and control group gain scores for each test measure at the 99% confidence interval.[4] Bar charts are displayed for each set of gain score comparisons.

Analysis of Implementation Time

An analysis of covariance (ANCOVA) was used to determine the relationship between the amount of instruction received by UPSTART participants and literacy outcomes. An ordinal version (ordered categories) of UPSTART use (transformed to hours of instruction) was used to see what the impact of instructional time in the program was on literacy outcomes as measured by total scores on the Brigance and Bader posttests. This was accomplished by creating a new variable called Usage Group in which hours of instruction was factored into four levels corresponding to quartiles. The ANCOVA was run separately for the Brigance and the Bader with Usage Group as the independent variable and the respective pretests as a covariate.

[4] To guard against Type I error (falsely rejecting the null hypothesis) in conducting tests of statistical inference (e.g., t-tests and multiple regression analysis) the criterion for statistical significance was based on the error rate for the collection of comparison required by the Brigance and the Bader. For example, comparisons among the ten Brigance subtest means and the total test at the .05 level can result in at least half of the statistical tests being significant by chance: 11 (.05) = .55. Increasing the rigor of the significance criterion by moving to the .01 level for the Brigance still leaves the possibility that the effective significance level for the collection of comparisons is .11, not .01: 11 (.01) = .11). This consideration resulted in a decision rule to set the confidence level at 99% and $p<.01$ for the collection of comparisons across the Bader and Brigance. For further detail, see Kirk R.E. (1968). *Experimental Design: Procedures for the Behavioral Sciences*. Belmont, CA: Wadsworth Publishing Company.

Results

Findings are reported first in answering research questions 3, 4, and 5 about the implementation of the UPSTART program in its fourth year of operation. We also report data describing the education technology equipment provided to enrolled children by UPSTART. We then report findings on the impact of UPSTART to answer research question 1 and 2 about the extent to which UPSTART may have facilitated the literacy development of children who participated in the home-based education technology preschool program compared to children who did not participate in the program.

UPSTART Implementation

Findings reviewed under UPSTART implementation include enrollment in the fourth year, equipment provided to enrolled families by UPSTART, usage of the UPSTART curriculum in terms of instructional time logged, the proportion of UPSTART students considered to have "graduated" from the program, and the relationship between levels of UPSTART curriculum usage and literacy outcomes.

UPSTART Enrollment

The Waterford Institute provided documentation for a fourth-year UPSTART enrollment of 1,250 children. Some basic demographic characteristics of the C4 population are presented below in Table 2 compared to the C4 analysis sample as estimated from the Brigance subgroup (n=101).

Table 2
Basic Demographic Characteristics: C4 population vs. C4 Analysis Sample

Demographic Categories		***All UPSTART (N=1,250)***	***Analysis Sample (N=101)***
Child's Gender	Boys	49.2%	49.5%
	Girls	50.8%	50.5%
Child's Ethnicity	Caucasian	75.5%	89.0%
	Hispanic	19.5%	4.0%
	Asian/Pacific Islander	1.7%	1.0%
	Black	0.7%	0.0%
	Native American	0.9%	0.0%
	Multiracial	--	5.0%
	Other	1.6%	1.0%
Child's Primary Language	English	84.6%	100.0%
	Spanish	14.3%	0.0%
	Other	1.1%	0.0%
	Some High School	7.2%	0.0%

Demographic Categories		*All UPSTART (N=1,250)*	*Analysis Sample (N=101)*
Parent Educational Attainment	High School Graduate	13.8%	5.0%
	Some College	31.1%	43.6%
	College Graduate	37.9%	47.5%
	Advanced Degree	6.7%	4.0%
	Unknown	3.3%	0.0%
Parent Marital Status	Married	89.5%	93.1%
	Other	10.5%	6.9%

As shown in Table 2, the analysis sample was somewhat of a more advantaged subgroup compared to the C4 population. For example, the C4 population is under-represented in the analysis sample in terms of Hispanics and Spanish speakers. However, this is somewhat problematic from a language standpoint in terms of test-taking logistics. The analysis sample also tends to be somewhat more highly educated than the C4 population. Both examples illustrate the need for the analysis sample to be more representative of the enrolled UPSTART population.

UPSTART Equipment Provided

The kind of education technology supports provided to children enrolled in UPSTART is shown in Table 3 for all 1,250 students enrolled in the fourth year and for the C4 preschool analysis sample (using the Brigance test). The majority of the fourth year UPSTART students (69%) received a computer drive with the UPSTART curriculum loaded on it. This allowed families to access the UPSTART curriculum from their home computers. Similarly, the students in the C4 preschool analysis sample most often (77.2%) also received a computer drive with the curriculum loaded on it.

UPSTART provided personal computers and free Internet subscriptions to 11.3% of the C4 students; this was also the case with about 7.9% of the C4 preschool analysis sample. Another 7.6% of the C4 program participants were given access to a home computer for free while they participated in the program. In the C4 analysis sample, 3% were given access to a home computer for free while they participated in the program. The remaining 12% of the C4 enrollment received various combinations of computer technology to enable them to access the UPSTART curriculum.

Table 3
Equipment provided to C4 Participants

Equipment Provided	*All UPSTART (N=1,250)*	*Analysis Sample (N=101)*
Drive	69.0%	77.2%

Equipment Provided	*All UPSTART (N=1,250)*	*Analysis Sample (N=101)*
Computer & Internet	11.3%	7.9%
Computer	7.6%	3.0%
Computer & Wireless	5.4%	5.0%
Computer & Cellular	3.6%	5.0%
Internet & Drive	2.3%	2.0%
Cellular & Drive	0.6%	0.0%
Computer with Wireless & Internet	0.2%	0.0%

UPSTART Graduates

Of the 1,250 children documented as enrolled in UPSTART in the fourth year of the program, the Waterford Institute classified 1,169 as children who had met the program's usage criteria and were thereby considered to be graduates of the program. The usage criteria included (a) logging more than 1,000 minutes (16.67 hours of instruction) with the UPSTART curriculum and (b) averaging at least one hour of instruction per week while they were participating in the program. By these criteria, Cohort 4 achieved a *graduation rate of 94%* (i.e., 1,169/1,250 = .935, rounded to 94%).

UPSTART graduate status was significantly correlated with hours of instruction (r = .55) as well as with Total Brigance posttest scores in kindergarten (r = .31) and with Total Bader posttest scores in kindergarten (r = .35).

UPSTART Usage

The hours of instruction observed for all children documented to be enrolled in the fourth year of UPSTART are summarized in Table 4 compared to "graduates" and the children in the C4 preschool analysis samples. The average level of usage for all students enrolled in the fourth year of UPSTART was approximately 71 hours of instruction. The C4 academic year covered approximately 45 weeks of instruction, beginning August 1, 2012 and ending June 28, 2013.

Five of the enrolled families were provided instructional equipment (i.e., computers, an Internet subscription, and a computer drive), but the enrolled children in these families did not log any instructional time in the UPSTART curriculum during Year 4 of the program. For enrolled families whose children did use the curriculum, the average duration in the program was approximately 40 weeks.

The children in the C4 analysis samples used the UPSTART curriculum for approximately 76-78 hours of instruction on the average (see Table 4). The levels of usage for each of the test samples

were slightly higher than the average UPSTART graduate and 6-7 hours more than the average participant in Year 4 of the program.

Table 4
C4 Hours of UPSTART Instruction

Group	*N*	*Mean*	*SD*	*Range*
All UPSTART	1,250	71.19	25.78	00.00 – 190.76
UPSTART Graduates	1,169	74.94	21.98	17.24 - 190.76
Brigance Analysis Sample	100	76.48	26.34	0.54 – 151.26
Bader Analysis Sample	78	78.40	26.68	0.54 – 151.26

The histograms in Figures 1-4 show the distribution of hours of instruction for the total C4 population (Figure 1), the C4 graduates (Figure 2), and the C4 analysis samples (Figure 3 for the Brigance sample and Figure 4 for the Bader sample). All four histograms show hours of instruction to be essentially normally distributed.

C4 Population Usage. In the C4 population (see Figure 1), UPSTART curriculum usage was normally distributed with an average usage level of approximately 71 hours. As previously noted, five of the enrolled children logged zero hours of instruction during their time in UPSTART (i.e., up to six weeks). At the other end of the spectrum, two children logged over 190 hours of instruction.

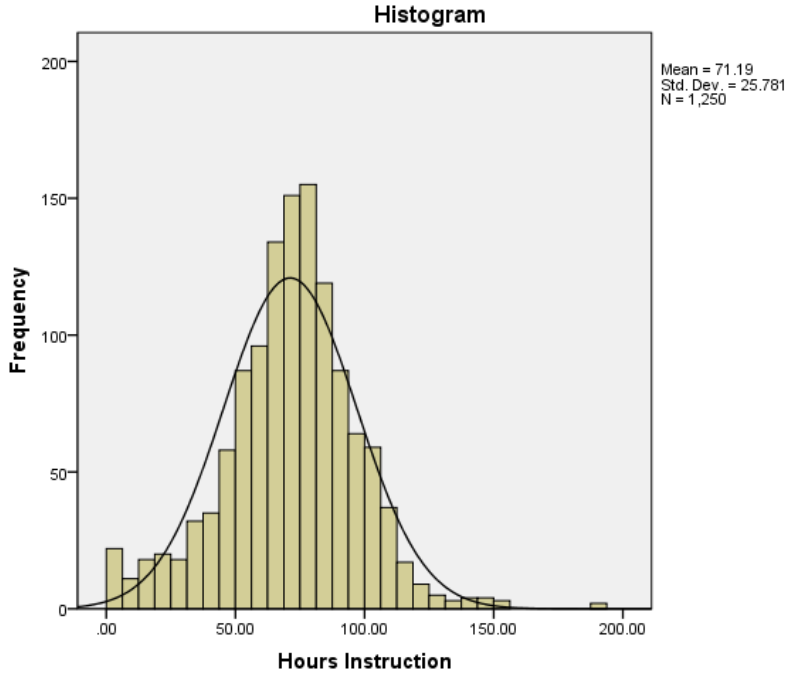


Figure 1. Hours of Instruction for All Students Enrolled in UPSTART in Year 4

The bottom quartile of the C4 population completed 57.04 hours of instruction or less. The midpoint of the C4 population distribution (the median) was 72.55 hours of instruction. The top quartile of the C4 population completed in excess of 86.06 hours of instruction.

C4 Graduate Usage. UPSTART curriculum usage for the subset of graduates was normally distributed with an average usage level of 74.94 hours of instruction. As noted previously, two of the C4 participants – both graduates -- logged over 190 hours of instruction. The bottom quartile of the C4 graduates ranged from 17.24 hours to 60.72 hours of usage. The midpoint of the C4 graduate distribution (the median) was 74.36 hours of instruction. The top quartile of the C4 graduates completed in excess of 87.54 hours of instruction.

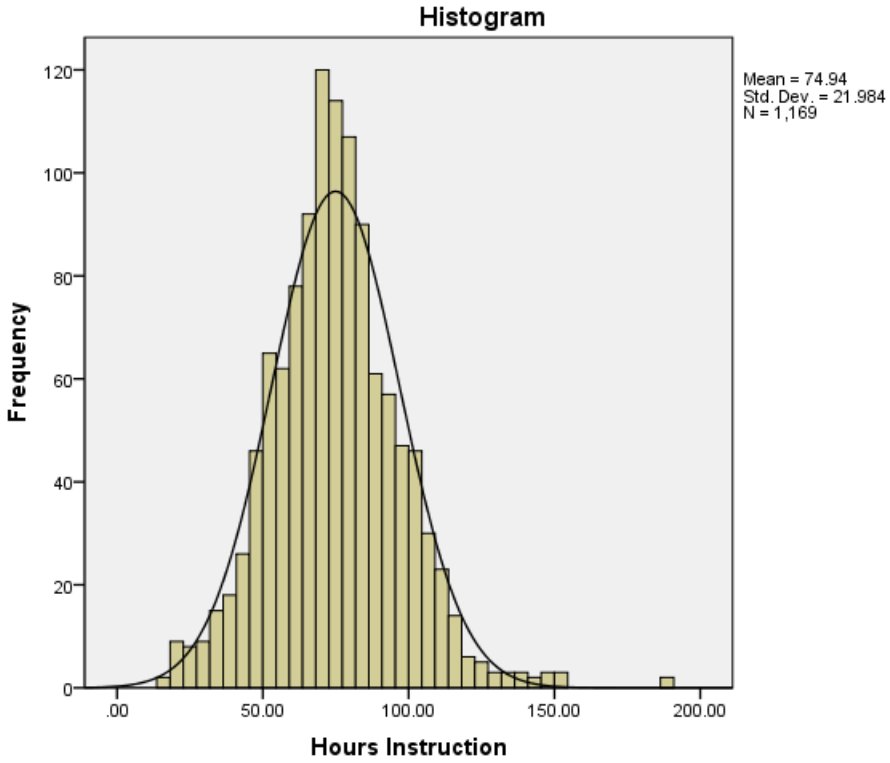


Figure 2. Hours of Instruction for UPSTART Graduates in Year 4

C4 Analysis Sample Usage. UPSTART curriculum usage for the C4 analysis samples was essentially normally distributed. The average level of usage for the Brigance sample was approximately 77 hours of instruction; for the Bader sample, the average level of usage was approximately 78 hours. Curriculum usage for children in the analysis samples ranged from less than one hour to over 150 hours of instruction.

The Brigance Analysis Sample

UPSTART usage for the C4 Brigance analysis sample (see Figure 3) is essentially normally distributed with a mean of 76.48 hours of instruction and a standard deviation of 26.34 hours.

The Brigance analysis sample's median is 76.75 hours of instruction. For the usage analysis with the Brigance sample, hours of instruction are distributed as follows by quartile of usage:

- 1st Quartile: 0.54 hours to 64.67 hours
- 2nd Quartile: 64.68 hours to 76.74 hours
- 3rd Quartile: 76.75 hours to 91.14 hours
- 4th Quartile: 91.15 hours to 151.25 hours

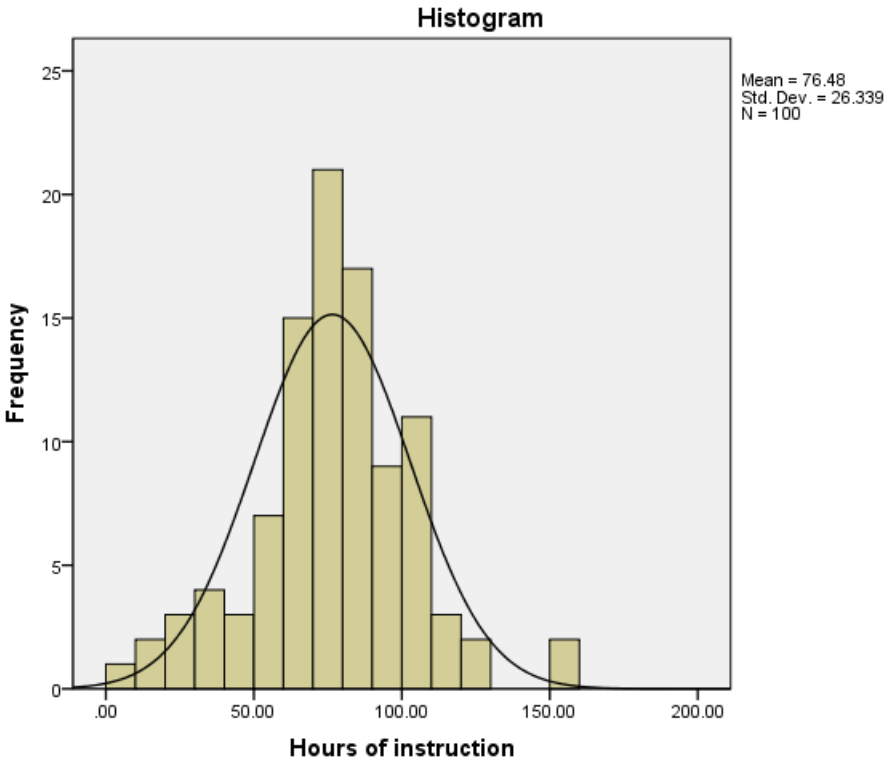


Figure 3. Hours of Instruction for C4 Brigance Analysis Sample

The Bader Analysis Sample

UPSTART usage for the C4 Bader analysis sample (see Figure 4) is also normally distributed with a mean of 78.4 hours of instruction and a standard deviation of 26.68 hours. The Bader analysis sample's median is 79.43 hours of instruction. For the usage analysis with the Bader sample, hours of instruction are distributed as follows by quartile of usage:

- 1st Quartile: 0.54 hours to 65.11 hours
- 2nd Quartile: 65.12 hours to 79.43 hours
- 3rd Quartile: 79.44 hours to 92.7 hours
- 4th Quartile: 92.71 hours to 151.25 hours

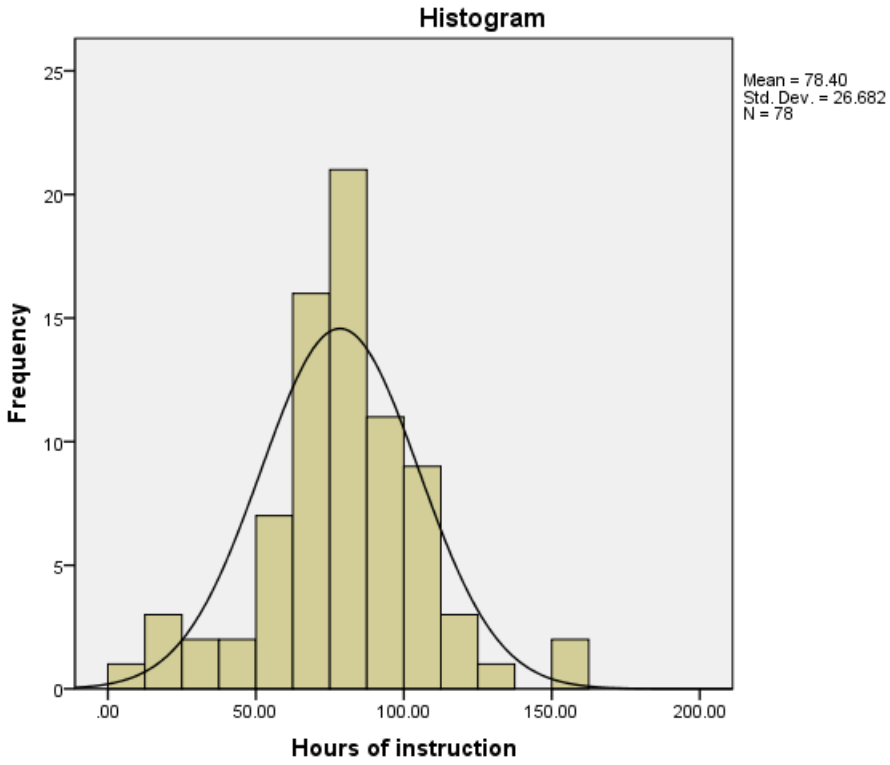


Figure 4. Hours of Instruction for C4 Bader Analysis Sample

UPSTART Usage and Literacy Outcomes

As in previous years, the fourth year evaluation of UPSTART found curriculum usage to be significantly and positively related to literacy outcomes as measured by total posttest scores on both the Brigance the Bader. The correlation between UPSTART usage and literacy outcomes measured by the Brigance Total Posttest was positive and statistically significant (r=.49, p<.01, n=100). Similarly, the correlation between UPSTART usage and literacy outcomes measured by the Bader Total Posttest was also positive and statistically significant (r= .44, p<.01, n=78). The relationship between usage and literacy outcomes was examined further as discussed below.

UPSTART Usage as a Predictor of Brigance Outcomes. A simple regression of usage on Brigance posttest scores showed that usage (measured in hours) accounted for 23% of the variance in literacy outcomes measured by the Brigance (adjusted R^2 = .23). See Table 5 below for the regression model summary and Table 6 for the Analysis of Variance (ANOVA) summary. The relationship between usage and Brigance posttest scores was moderately strong as indicated by the standardized regression coefficient in Table 7 below (Beta = .49) and statistically significant (p<.01).

Table 5: Usage Model Summary for Brigance

R	R Square	Adjusted R Square	Std. Error of the Estimate
.491	.241	.233	37.515

The independent variable is Hours of instruction.

Table 6: ANOVA of Usage for Brigance Outcomes

	Sum of Squares	df	Mean Square	F	Sig.
Regression	43812.009	1	43812.009	31.130	.000
Residual	137924.631	98	1407.394		
Total	181736.640	99			

The independent variable is Hours of instruction.

Table 7: Usage Coefficients for Brigance Outcomes

	Unstandardized Coefficients		Standardized Coefficients	t	Sig.
	B	Std. Error	Beta		
Hours of instruction	.799	.143	.491	5.579	.000
(Constant)	117.479	11.573		10.152	.000

The curve-fitting plot presented in Figure 5 below shows a substantial linear relationship between UPSTART usage (measured in hours of instruction) and Brigance posttest scores.

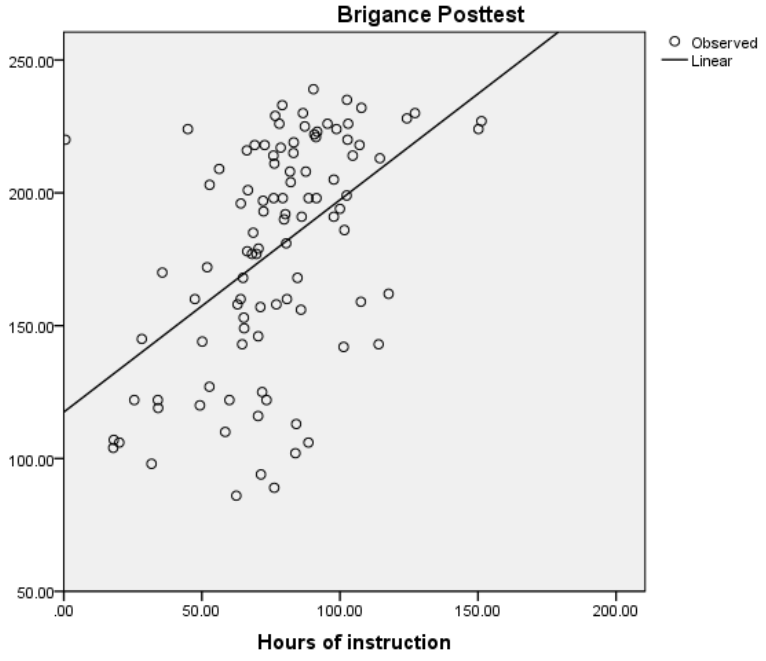


Figure 5. Plot of Hours of Instruction and Brigance Posttest Scores

UPSTART Usage as a Predictor of Bader Outcomes. The same procedures were performed with the C4 UPSTART usage data and the Bader posttest scores. Similar results were found: 18% of the variance in Bader posttest outcomes could be accounted for by UPSTART usage (adjusted $R^2 = .18$) and the relationship between usage and Bader posttest scores was moderately strong (Beta = .44) and statistically significant ($p<.01$). See Table 8 below for the regression model summary, Table 9 for the Analysis of Variance (ANOVA) summary, and Table 10 for the regression coefficients showing the impact of usage on Bader posttest scores.

Table 8
Usage Model Summary for Bader

R	R Square	Adjusted R Square	Std. Error of the Estimate
.441	.195	.184	6.232

The independent variable is Hours of instruction.

Table 9
ANOVA of Usage for Bader Outcomes

	Sum of Squares	df	Mean Square	F	Sig.
Regression	712.915	1	712.915	18.354	.000
Residual	2952.072	76	38.843		
Total	3664.987	77			

The independent variable is Hours of instruction.

Table 10
Usage Coefficients for Bader Outcomes

	Unstandardized Coefficients		Standardized Coefficients	t	Sig.
	B	Std. Error	Beta		
Hours of instruction	.114	.027	.441	4.284	.000
(Constant)	8.072	2.203		3.664	.000

Dependent variable: Bader Posttest

The curve fitting plot presented in Figure 6 shows a moderately strong linear relationship between UPSTART usage and the development of phonological awareness skills as measured by the Bader assessment (standardized regression coefficient for usage = 0.44, p<.01).

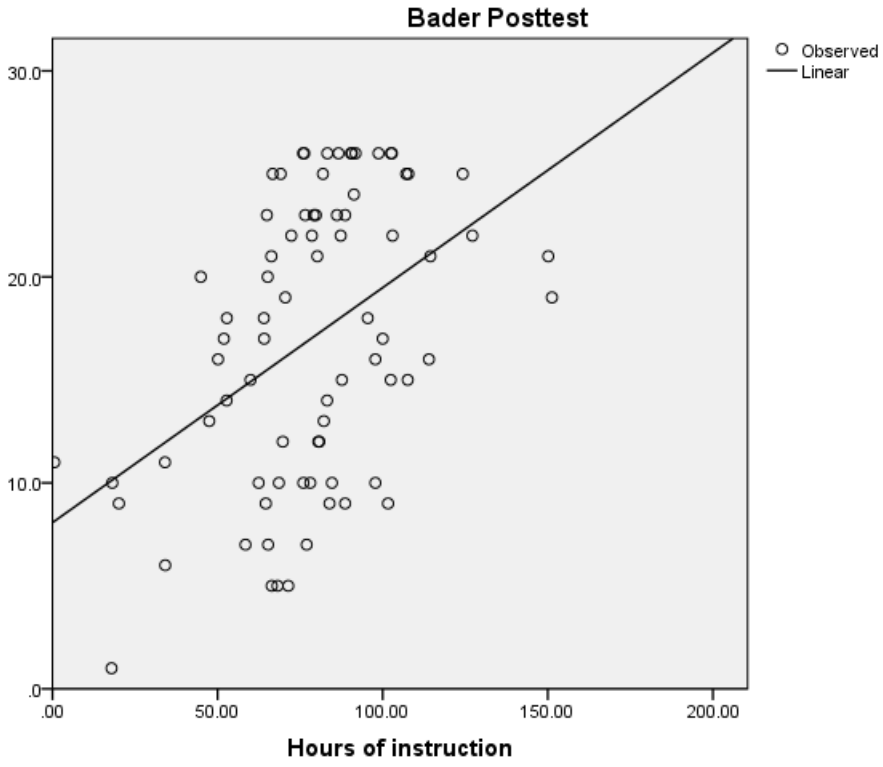


Figure 6. Plot of Hours of Instruction and Bader Posttest Scores.

Analysis of Covariance

The level of initial literacy development (as measured by the assessment pretests) significantly influences both UPSTART usage (r=.37 in the Brigance sample and r=.25 in the Bader sample) and posttest performance at kindergarten entry (r=.69 in the Brigance sample and r=.44 in the Bader sample). Because usage effects are confounded by pretest scores[5] we wanted to re-examine the relationship between UPSTART usage and literacy outcomes at kindergarten entry by controlling for initial levels of literacy development. This was accomplished with an analysis of covariance (ANCOVA) in which we tested to see if literacy outcomes at kindergarten entry increased with UPSTART usage quartile (measured in hours of instruction as described

[5] When pretest scores are controlled for, the correlation between usage and Brigance posttest scores drops from r=.49 (simple correlation) to r=.35 (partial correlation). Similarly, when pretest scores are controlled for, the correlation between usage and Bader posttest scores drops from r=.44 (simple correlation) to r=.38 (partial correlation).

previously for the analysis samples), controlling for initial levels of literacy development (as measured by the respective pretests for the Brigance and the Bader).

Brigance ANCOVA. The Brigance pretest was used as a control covariate in the C4 ANCOVA. The ANCOVA results for the Brigance test sample shown in Table 11 reveal that the usage factor is statistically significant ($p<.01$). The amount of variance in Brigance Posttest scores accounted for by usage, controlling for prior literacy achievement, is approximately 15% (see the Partial Eta Squared statistic column).

Table 11

UPSTART Usage as a predictor of Brigance Total Posttest Scores

(Dependent Variable: Brigance Posttest; Adjusted R Squared Model = .53, Computed using alpha = .05)

Source	*Type III Sum of Squares*	*df*	*Mean Square*	*F*	*Sig.*	*Partial Eta Squared*	*Noncent. Parameter*	*Observed Power*[b]
Corrected Model	99732.719	4	24933.180	28.885	.000	.549	115.538	1.000
Intercept	120015.328	1	120015.328	139.035	.000	.594	139.035	1.000
Brigance_Pre	50324.479	1	50324.479	58.300	.000	.380	58.300	1.000
Brigance_Usage	14540.663	3	4846.888	5.615	.001	.151	16.845	.937
Error	82003.921	95	863.199					
Total	3370104.000	100						
Corrected Total	181736.640	99						

In Table 12 below, each usage group is identified by its quartile value, 1 through 4. The covariance model compares the effects of each level of usage with the fourth quartile level of usage for the C4 analysis sample, controlling for the influence of initial literacy skills, and displays the difference in Brigance total posttest scores in the column labeled B – expressed as a regression coefficient. The parameter estimates in Table 12 suggest a linear trend in literacy achievement (as measured by total Brigance posttest scores) with increasing usage.

Table 12

Parameter Estimates: Brigance Total Posttest Score - Beginning Kindergarten

Parameter	*B*	*Std. Error*	*t*	*Sig.*	*Partial Eta Squared*	*Observed Power*
Intercept	126.060	11.960	10.540	.000	.539	1.000
Brigance_Pre	.621	.081	7.635	.000	.380	1.000
Usage_Group=1.00	-35.135	8.912	-3.943	.000	.141	.974
Usage_Group=2.00	-18.126	8.541	-2.122	.036	.045	.556
Usage_Group=3.00	-8.709	8.354	-1.043	.300	.011	.178
Usage_Group=4.00	0	.	.	.	.	

The covariance-adjusted Brigance posttest means can be seen more clearly by usage quartile in Table 13. These data indicate that literacy achievement differs significantly between usage quartiles 1 and 3 and between quartiles 1 and 4. While there is an apparent linear trend in the data, with literacy achievement increasing with usage levels, the difference between Brigance posttest means at usage quartiles 3 and 4 are not statistically significant nor is the difference in posttest means between usage quartiles 2 and 3 statistically significant.

Table 13

Parameter Estimates: Brigance Total Posttest Scores by Usage Quartile

Dependent Variable: Brigance Posttest

Brigance Usage Group	Mean	Std. Error	95% Confidence Interval	
			Lower Bound	Upper Bound
Quartile 1	158.917[a]	6.119	146.769	171.065
Quartile 2	175.926[a]	5.894	164.225	187.628
Quartile 3	185.344[a]	5.913	173.606	197.082
Quartile 4	194.053[a]	6.068	182.007	206.098

a. Covariates appearing in the model are evaluated at the following values:
Brigance Pretest = 109.5200.

The trend in literacy achievement levels across usage quartiles is shown in Figure 7 below.

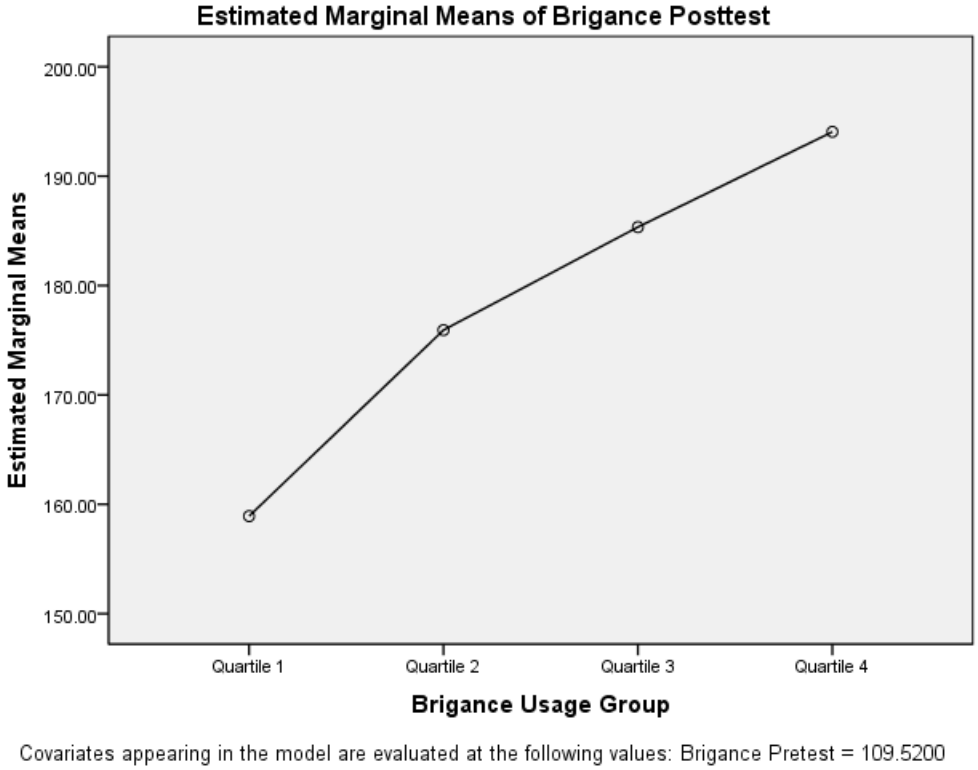


Figure 7. Brigance Literacy Development by UPSTART Usage Quartiles

Bader ANCOVA. The Bader pretest was used as a control covariate in the C4 ANCOVA for the Bader test sample. The ANCOVA results for the Bader test sample shown in Table 14 reveal that the usage factor is statistically significant (p=.01). The amount of variance in Bader Posttest scores accounted for by usage, controlling for prior literacy achievement, is approximately 13% (see the Partial Eta Squared statistic column).

Table 14
Tests of Between Subjects Effects: Bader Total Posttest - Beginning Kindergarten

Dependent Variable: Bader Postttest

Source	*Type III Sum of Squares*	*df*	*Mean Square*	*F*	*Sig.*	*Partial Eta Squared*	*Noncent. Parameter*	*Observed Power[b]*
Corrected Model	1193.792	4	298.448	8.816	.000	.326	35.265	.999
Intercept	4099.387	1	4099.387	121.097	.000	.624	121.097	1.000
Bader_Pre	620.999	1	620.999	18.345	.000	.201	18.345	.988
Bader_Usage	366.755	3	122.252	3.611	.017	.129	10.834	.773
Error	2471.195	73	33.852					
Total	2624.000	78						
Corrected Total	3664.987	77						

Adjusted R Squared Model = .29 Computed using alpha = .05

In Table 15 below, each usage group is identified by its quartile value, 1 through 4. The covariance model compares the effects of each level of usage with the fourth quartile level of usage for the C4 analysis sample, controlling for the influence of initial literacy skills, and displays the difference in Bader total posttest scores in the column labeled B – expressed as a regression coefficient. The parameter estimates in Table 15 suggest a linear trend in literacy achievement (as measured by total Brigance posttest scores) with increasing usage. However, differences beyond Quartile 1 are not statistically significant.

Table 15
Parameter Estimates: Bader Total Posttest Score - Beginning Kindergarten

Parameter	*B*	*Std. Error*	*t*	*Sig.*	*Partial Eta Squared*	*Noncent. Parameter*	*Observed Power[a]*
Intercept	15.119	1.708	8.851	.000	.518	8.851	1.000
Bader_Pre	.607	.142	4.283	.000	.201	4.283	.988
Usage_Group =1.00	-5.672	1.906	-2.976	.004	.108	2.976	.836
Usage_Group=2.00	-2.760	1.873	-1.474	.145	.145	1.474	.307
Usage_Group=3.00	-0.600	1.864	-0.322	.748	.748	.322	.062
Usage_Group=4.00	0	.	.	.	.	.	.

The covariance-adjusted Brigance posttest means can be seen more clearly by usage quartile in Table 16. These data indicate that literacy achievement differs significantly between usage quartile 1 and quartiles 3 and 4. While there is an apparent linear trend in the data, with literacy achievement increasing with usage levels, the differences in literacy achievement between adjacent quartiles is not statistically significant. Rather, the data indicate that literacy achievement differs significantly between usage quartiles 1 and 3 and between quartiles 1 and 4. The difference between Bader posttest means at usage quartiles 3 and 4 are not statistically significant nor is the difference in posttest means between usage quartiles 2 and 3 statistically significant.

Table 16

Parameter Estimates: Bader Total Posttest Scores by Usage Quartile

Dependent Variable :Bader Posttest

			95% Confidence Interval	
Bader Usage Group	Mean	Std. Error	Lower Bound	Upper Bound
Quartile 1	13.584	1.344	10.904	16.264
Quartile 2	16.496	1.304	13.898	19.094
Quartile 3	18.656	1.308	16.048	21.263
Quartile 4	19.256	1.339	16.588	21.924

The apparent linear trend in the development of phonological awareness skills (as measured by the Bader Posttest at kindergarten entry) with UPSTART usage is shown in Figure 8.

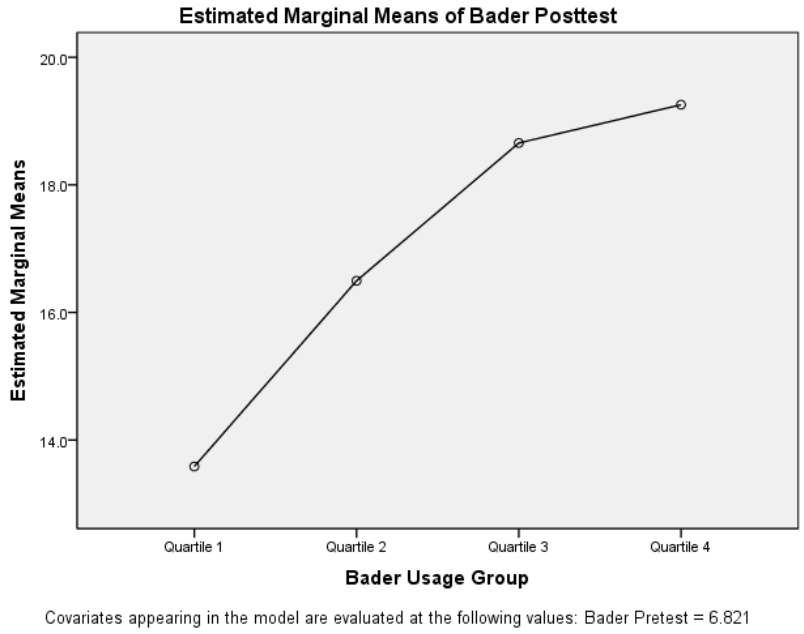


Figure 8. Bader Literacy Development by UPSTART Usage Quartiles

UPSTART Literacy Outcomes

In this section, the main research questions of interest to the C4 study are addressed:

Research Question 1: *Do UPSTART students have better literacy skills at kindergarten than control students?*

Research Question 2: *Do UPSTART students show stronger literacy growth rates from preschool to kindergarten than control students?*

For each of these two questions, results for the Brigance are reviewed first, followed by the results for the Bader.

Do UPSTART students have better literacy skills at Kindergarten than control students?

As discussed previously, the analytic strategy for answering Research Question 1 proceeded through the following phases:

- Pretest Analysis
- Covariate Analysis
    - Identifying significant pre-existing differences between the treatment and control groups
    - Identifying significant posttest predictors that differentiate the treatment and control groups
- Posttest Analysis
- Multiple Regression Analysis

Brigance Pretest Results

The control group children scored significantly higher on the overall Brigance at pretest compared to the UPSTART treatment group children. The average Brigance pretest difference was over 16 points. Significant Brigance subtest differences favoring the control children were observed on the Recites Alphabet test, the Letter Knowledge test and the Letter Sounds test (see Table 17 for details). These differences were probably related to the fact that substantially more of the control children were attending preschool at the time of the pretest (80% vs. 31%). This stands to reason since the control children were primarily recruited through preschools. Thus, there were identifiable pretest differences between the treatment and control group children on the Brigance.

Table 17
Brigance Pretest Analysis of Treatment-Control Group Differences

Brigance Pretest	*Group*	*N*	*Mean*	*SD*	*t*	*Significance*
Expressive Objects	Treatment	101	24.495	3.877	-1.869	NS
	Control	102	25.294	1.891		
Receptive Objects	Treatment	101	26.822	0.517	0.639	NS
	Control	102	26.735	1.258		
Expressive Grammar	Treatment	101	8.901	1.584	-0.653	NS
	Control	102	9.029	1.189		
Visual Discrimination	Treatment	101	13.515	4.969	0.049	NS
	Control	102	13.480	5.090		
Recites Alphabet	Treatment	101	8.624	8.715	-2.947	**
	Control	102	12.490	9.944		
Lowercase Letter Knowledge	Treatment	101	15.505	19.020	-2.493	**
	Control	102	22.402	20.382		
Sounds of Lowercase Letters	Treatment	101	4.505	7.497	-2.515	**
	Control	102	7.559	9.674		
Auditory Discrimination	Treatment	101	4.515	3.425	-2.364	NS
	Control	102	5.676	3.574		
Survival Sight Words	Treatment	101	1.228	1.535	-1.194	NS
	Control	102	1.549	2.236		
Basic Pre-primer Vocabulary	Treatment	101	1.079	3.984	-0.394	NS
	Control	102	1.314	4.469		
Total Brigance	Treatment	101	109.188	39.216	-2.832	**
	Control	102	123.529	42.934		

**p≤.01

There were several demographic differences between the treatment and control children in the Brigance analysis sample that were related to posttest outcomes. These included the child's ethnicity (if Hispanic), the child's primary language (if English), and whether or not the parent was married at the time of the in-take interview. These three variables plus total Brigance pretest scores were used as covariates in a regression analysis to adjust posttest outcomes for pre-existing between group differences. The Brigance pretest was retained as a statistical control variable in the final regression analysis. See Appendix B and Appendix C for further details.

Brigance Posttest Results

Posttest results showed that the UPSTART treatment group performed significantly better than the control children on three of the Brigance subtests*: Visual Discrimination, Letter Sounds, and Vocabulary.* Posttest results also showed that the UPSTART treatment group performed significantly better than the control children on the *Total Brigance.* On the *Total Brigance*, the treatment group children outscored the control group children by an average of 16.44 points. The Brigance posttest results are shown below in Table 18. Graphs depicting pretest and posttest mean scores for Brigance subtests are included in Appendix D.

Table 18
Brigance Posttest Analysis of Treatment-Control Group Differences

Brigance Posttest	*Group*	*N*	*Mean*	*SD*	*t*	*Significance*
Expressive Objects	Treatment	101	25.980	.979	-0.138	NS
	Control	102	26.000	1.062		
Receptive Objects	Treatment	101	26.990	.099	1.000	NS
	Control	102	26.971	.169		
Expressive Grammar	Treatment	101	10.188	1.230	1.522	NS
	Control	102	9.922	1.264		
Visual Discrimination	Treatment	101	18.139	2.069	3.294	**
	Control	102	16.912	3.134		
Recites Alphabet	Treatment	101	18.891	8.956	1.647	NS
	Control	102	16.706	9.931		
Lowercase Letter Knowledge	Treatment	101	39.337	16.638	1.277	NS
	Control	102	36.176	18.566		
Sounds of Lowercase Letters	Treatment	101	17.198	8.918	2.643	**
	Control	102	13.588	10.486		
Auditory Discrimination	Treatment	101	7.594	3.076	1.327	NS
	Control	102	7.000	3.297		
Survival Sight Words	Treatment	101	3.703	3.968	1.330	NS
	Control	102	3.020	3.318		
Basic Pre-primer Vocabulary	Treatment	101	10.000	9.583	3.815	**
	Control	102	5.284	7.942		
Total Brigance	Treatment	101	178.019	42.974	2.669	**
	Control	102	161.578	44.778		

**p≤.01

Using the data from Table 18, effect sizes[6] were calculated to show the magnitude of UPSTART's impact at posttest as measured by each of the 10 Brigance subtests and the Total Brigance. The effect size (ES) estimates are presented in Table 19 below, and show the magnitude of the average performance difference in standard deviation units between the C4 treatment group and the control group on each of the Brigance assessments administered in the C4 evaluation.

The ES estimates for Visual Discrimination, Letter Sounds, and the Total Brigance range from .34 to .39 and indicate that the treatment group's scores are on the order of a third of a standard deviation larger than control group scores. These differences would be considered small effects by Cohen.[7] The ES for vocabulary is .59 and would be considered a medium size effect by Cohen. On average, treatment group children scored 16 points higher on the total Brigance at posttest compared to control group children.

[6] An effect size was calculated for each test as the treatment group mean minus the control group mean divided by the control group standard deviation.

[7] See Chapter 2 in Cohen, J. (1988). *Statistical Power Analysis for the Behavioral Sciences*. Hillsdale, NJ: Erlbaum

Table 19: Brigance Effect Size Estimates

Test	*Effect Size*	*Significance*	*Magnitude of Effect*
Expressive Objects	-0.018	NS	
Receptive Objects	0.111	NS	
Expressive Grammar	0.210	NS	
Visual Discrimination	0.391	**	Small
Recites Alphabet	0.220	NS	
Letter Knowledge	0.170	NS	
Letter Sounds	0.344	**	Small
Auditory Discrimination	0.180	NS	
Survival Sight Words	0.205	NS	
Basic Pre-Primer Vocabulary	0.593	**	Medium
Total Brigance	0.367	**	Small

**p≤.01

Figure 9 shows the Brigance effect size estimates by total test and subtest in bar chart format.

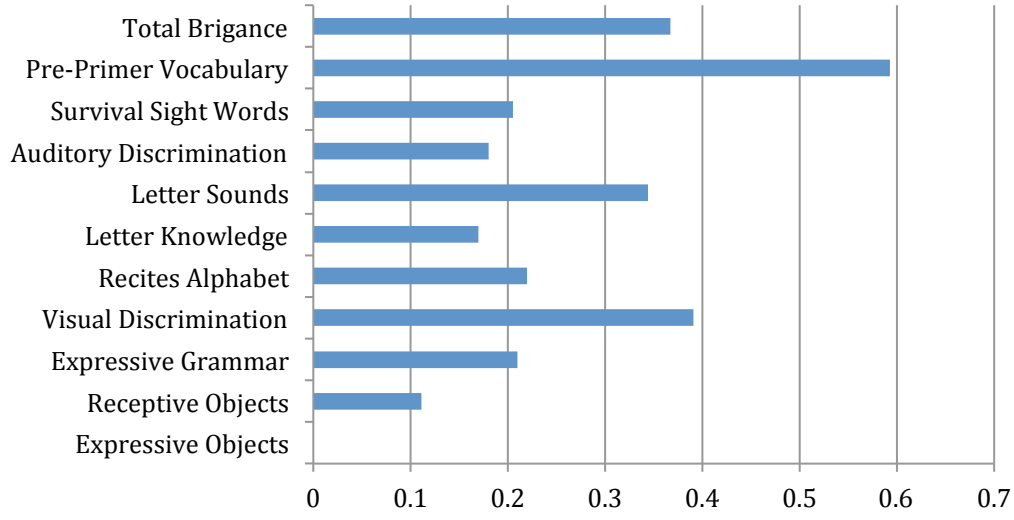


Figure 9: UPSTART Impact as Measured by the Brigance in Effect Size Units

Brigance Posttest Regression Results

Adjusting for initial differences in literacy skills between the treatment and control groups through the use of multiple regression analysis, it was found that the treatment group children outscored the control group children on the *overall Brigance* posttest by 29.43 points. The regression-adjusted posttest difference on the Total Brigance is substantially larger than the raw t-test results reviewed earlier (i.e., an average difference of approximately 29 points versus 16 points).

The final Brigance regression model[8] is shown in Tables 20 and 21.

Table 20
OLS ANOVA Summary Table for Total Brigance Posttest – Beginning of Kindergarten

Model	*Sum of Squares*	*df*	*Mean Square*	*F*	*Sig.*
Regression	228550.548	2	114275.274	132.593	.000
Residual	172370.624	200	861.853		
Total	400921.172	202			

Predictors: Total Brigance Pretest, Study Group (Treatment vs. Control)

Table 21
OLS Regression Coefficients for Total Brigance Posttest – Beginning of Kindergarten

Model	*Unstandardized Coefficients*		*Standardized Coefficients*			*Partial*
	B	Std. Error	Beta	t	*Sig.*	*Correlation*
(Constant)	61.792	6.957		8.882	.000	
Brigance Pretest	.795	.050	.746	15.788	.000	.745
Study Group	29.431	4.202	.331	7.004	.000	.444

Adjusted R Square = .56 Partial R Square = .20

The observed effect size for the Brigance treatment group using regression adjusted estimates is found by computing the partial R square statistic for Study Group, which in this case has a value of .20 (i.e., .444 squared = .197). A partial R square value of .20 is in the medium effects size range (see Cohen, 1988; Chapter 9). Note that the regression-based estimate of UPSTART's overall impact on literacy growth is larger (a medium effect) than the observed raw impact (a small effect) because the regression procedure statistically adjusted for the initial inequalities in literacy skill levels between the treatment group and the control group.

[8] The preliminary regression model showed that the effect of three additional demographic covariates (child's ethnicity if Hispanic, child's primary language if English, and parent's marital status) were statistically non-significant when entered into the regression equation with the Brigance pretest.

Bader Pretest Results

Similar to the Brigance results, the control group children also scored significantly higher on the Bader pretests compared to the treatment group children entering UPSTART. As with the Brigance, this probably reflects greater levels of participation in daycare and preschool among the control children (78% vs. 30% for the Bader sample). Table 22 shows that the control group children scored significantly higher on the three Bader subtests at the pretest as well as on the overall Bader at pretest.

Table 22
Bader Pretest Analysis of Treatment-Control Group Differences

Bader Pretest	*Group*	*N*	*Mean*	*SD*	*t*	*Significance*
Rhyme Recognition	Treatment	79	5.24	2.98	-2.19	*
	Control	93	6.23	2.91		
Phoneme Blending	Treatment	79	1.10	2.28	-2.91	**
	Control	93	2.31	3.16		
Phoneme Segmenting	Treatment	79	0.41	1.26	-1.98	*
	Control	93	0.94	2.20		
Total Bader	Treatment	79	6.75	4.78	-3.26	**
	Control	93	9.47	6.19		

* $p \leq .05$; ** $p \leq .01$

There were a number of demographic differences between the treatment and control children in the Bader analysis sample (see Appendix E). The demographic differences shown in Table 23 are those that were significantly related to Bader posttest outcomes (see Appendix C). In general, these initial differences favored the UPSTART treatment group.

Table 23
Bader Sample Demographic Differences Related to Posttest Outcomes

Demographic	*Group*	*N*	*Mean*	*SD*	*t*	*Significance*
Parent's educational attainment	Treatment	79	3.59	.63	2.21	*
	Control	92	3.35	.80		
Parent's primary language is English	Treatment	79	.99	.11	2.30	*
	Control	93	.91	.28		
Child's primary language is English	Treatment	79	1.00	.00	2.20	*
	Control	93	.94	.25		
Child is Hispanic	Treatment	79	.03	.16	-2.84	**
	Control	93	.14	.35		
Parent is Married	Treatment	79	.95	.22	2.63	**
	Control	93	.83	.38		

* $p \leq .05$; ** $p \leq .01$

The above five demographics along with the Bader pretest were used as covariates in a multiple regression analysis (see below) to adjust posttest scores for initial between group differences.

Bader Posttest Results

Bader posttest results showed a statistically significant treatment group effect for Phoneme Blending and Phoneme Segmentation as well as for the Total Bader; see Table 24 below. The mean observed (unadjusted) difference between the treatment and control group on the Total Bader posttest – which favored the treatment group – was 3.89 points. Graphs depicting pretest and posttest mean scores for Bader subtests are included in Appendix F.

Table 24
Bader Posttest Analysis of Treatment-Control Group Differences

Bader Posttest	***Group***	***N***	***Mean***	***SD***	***t***	***Significance***
Rhyme Recognition	Treatment	79	7.86	2.53	1.27	NS
	Control	93	7.31	3.16		
Phoneme Blending	Treatment	79	5.09	2.96	3.35	**
	Control	93	3.48	3.32		
Phoneme Segmenting	Treatment	79	3.98	3.42	3.45	**
	Control	93	2.24	3.13		
Total Bader	Treatment	79	16.92	6.90	3.52	**
	Control	93	13.03	7.70		

** p<.01

Using Cohen's standardized difference score conventions, effect size estimates for the Bader posttest results are presented below in Table 25, and show the magnitude of the average performance difference in standard deviation units between the C4 treatment group and the control group on each of the Bader assessments administered in the C4 evaluation.

Table 25
Bader Effect Size Estimates

Test	***Effect Size***	***Significance***	***Magnitude of Effect***
Rhyme Recognition	0.174	NS	
Phonemic Blending	0.484	**	Small
Phoneme Segmentation	0.555	**	Medium
Total Bader	0.519	**	Medium

** p<.01

Figure 10 shows the Bader effect size estimates by total test and subtest in bar chart format. The effect size for Phonemic Blending technically falls within the small effects range (ES < .50) but as can be seen in the Figure 10 bar chart, it approaches the medium effect size range and essentially has the appearance of a medium size effect.

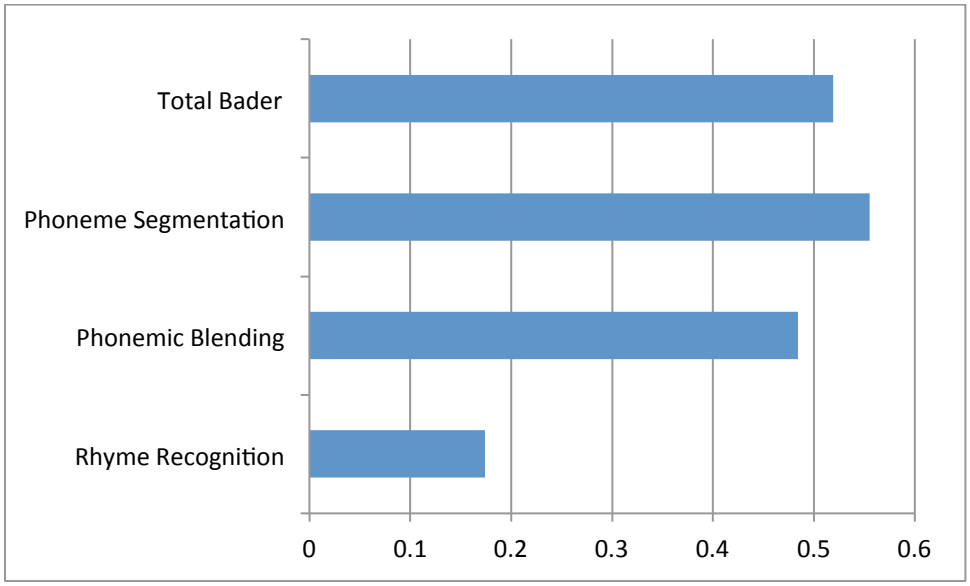


Figure 10: UPSTART's Impact as Measured by the Bader in Effect Size Units

Bader Posttest Regression Results

Adjusting for the initial between group differences using multiple regression analysis, it was found that the treatment group outscored the control group on the Total Bader by 5.92 points on the average. The final Bader regression model[9] is shown below in Tables 26 and 27.

Table 26
OLS ANOVA Summary Table for Total Bader Posttest – Beginning of Kindergarten

Model	*Sum of Squares*	*df*	*Mean Square*	*F*	*Sig.*
Regression	3580.472	2	1790.236	50.824	.000
Residual	5952.941	169	35.225		
Total	9533.413	171			

Predictors: Bader Pretest and Study Group

[9] Preliminary regression models showed that the effect of the demographic covariates (parent's education, child's ethnicity if Hispanic, parent and child's primary language if English, and parent's marital status if married) were statistically non-significant when entered into the regression equation with the total Bader pretest.

Table 27
OLS Regression Coefficients for Total Bader Posttest – Beginning of Kindergarten

	Unstandardized Coefficients		*Standardized Coefficients*			*Partial*
	B	Std. Error	Beta	t	*Sig.*	*Correlation*
(Constant)	5.990	.987		6.069		
Total Pretest	.743	.081	.571	9.126	.000	.575
Study Group	*5.919*	.935	.396	6.331	.000	.438

Adjusted R Square = .37 Partial R Square = .19

The partial R square for Study Group in the Bader regression analysis is .19, which suggests a medium effect size (see Cohen, 1988; Chapter 9) for the C4 UPSTART program in helping to develop children's phonological awareness as measured by the overall Bader. It should also be noted that the statistically adjusted overall treatment effect of 5.92 is larger than the raw difference score reported previously in the discussion of the t-test findings (i.e., 3.89 points).

Do UPSTART students show stronger literacy growth rates from preschool to kindergarten than control students?

Paired samples t-tests were performed to examine growth rates as measured by the Brigance and the Bader total test batteries and subtests for the treatment and control group children. Growth rates for the treatment and control children were compared based on the observed difference scores between the posttest and the pretest. Significant differences in growth rates were estimated on the basis of whether or not the confidence intervals of the treatment and control groups overlapped at the 99% Confidence Interval of the Mean Growth Rate.

Brigance Growth Score Results

There was statistically significant growth from pretest to posttest for the matched Brigance treatment group sample (N=101) on the Total Brigance and on all ten subtests. Similar results were observed for the matched Brigance control group (N=102) except for the Receptive Objects subtest which did not show significant growth from pretest to posttest.

Growth rates were significantly different at the 99% CI between the treatment and control group for the *overall Brigance* and four subtests: *Vocabulary, Sight Words, Letter Sounds, and Recites Alphabet.* All of these differences in growth rates favored the UPSTART treatment group. These results are shown in Table 28 below.

Table 28
Treatment-Control Group Differences in Growth Rates on the Brigance

Brigance Test	***Control Group***		***Treatment Group***		***Significance p≤.01***
	Mean Growth	99% CI Growth Rate	Mean Growth	99% CI Growth Rate	
Expressive Objects	.706	.301 – 1.111	1.485	.522 – 2.448	NS
Receptive Objects	.235	-.092 – .563	.168	.030 - .307	NS
Expressive Grammar	.892	.490 - 1.294	1.287	.848 – 1.726	NS
Visual Discrimination	3.431	2.206 – 4.657	4.624	3.447 – 5.800	NS
Recites Alphabet	4.216	1.605 -6.827	10.267	7.848 – 12.687	**
Letter Knowledge	13.775	9.558 -17.992	23.332	13.997 – 28.666	NS
Letter Sounds	6.029	3.964 – 8.105	12.693	10.495 – 14.892	**
Auditory Discrimination	1.324	.112 – 2.535	3.079	2.021 – 4.138	NS
Survival Sight Words	1.471	.850 -2.091	2.475	1.667 – 3.284	**
Basic Vocabulary	3.971	2.351 -5.591	8.921	6.608 – 11.234	**
Total Brigance	36.049	28.737 - 43.361	68.832	60.294 - 77.369	**

The differences in growth rates between the treatment and control groups are also shown in bar chart format below in Figure 11.

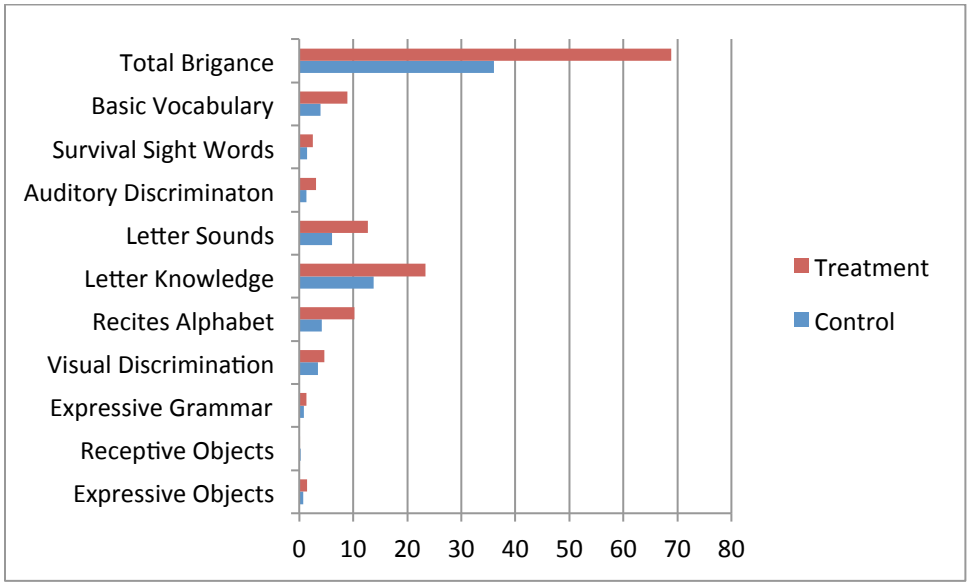


Figure 11. Growth Rate Comparisons on the Brigance

Bader Growth Score Results

There was statistically significant growth from pretest to posttest for the matched Bader treatment group sample (N=79) and for the matched Bader control group sample (N=93) on the Total Bader and all of the Bader subtests. Additionally, the UPSTART treatment group showed significantly stronger growth rates (statistically significant at the 99% CI) relative to the control

group on the Total *Bader* and on two of the three Bader subtests. Specifically, the UPSTART treatment group showed stronger growth rates from pretest to posttest on the Phoneme Blending and Phoneme Segmenting subtests as well as on the overall Bader test. Growth rates for rhyme recognition were not statistically significant between the treatment and control groups at the 99% CI. These results are shown in Table 29.

Table 29
Treatment-Control Group Differences in Growth Rates on the Bader

Bader Test	***Control Group***		***Treatment Group***		***Significance p≤.01***
	Mean Growth	99% CI Growth Rate	Mean Growth	99% CI Growth Rate	
Rhyme Recognition	1.086	.063 – 2.110	2.620	1.632 – 3.608	NS
Phoneme Blending	1.172	.449 – 1.900	3.987	3.138 - 4.836	**
Phoneme Segmenting	1.301	.567 - 2.046	3.570	2.585 – 4.555	**
Total Bader	3.559	1.927 - 5.191	10.177	8.332 - 12.022	**

Figure 12 uses bar charts to compare the growth rates of the treatment and control group as measured by the Total Bader and each of its subtests from pretest to posttest for the matched samples.

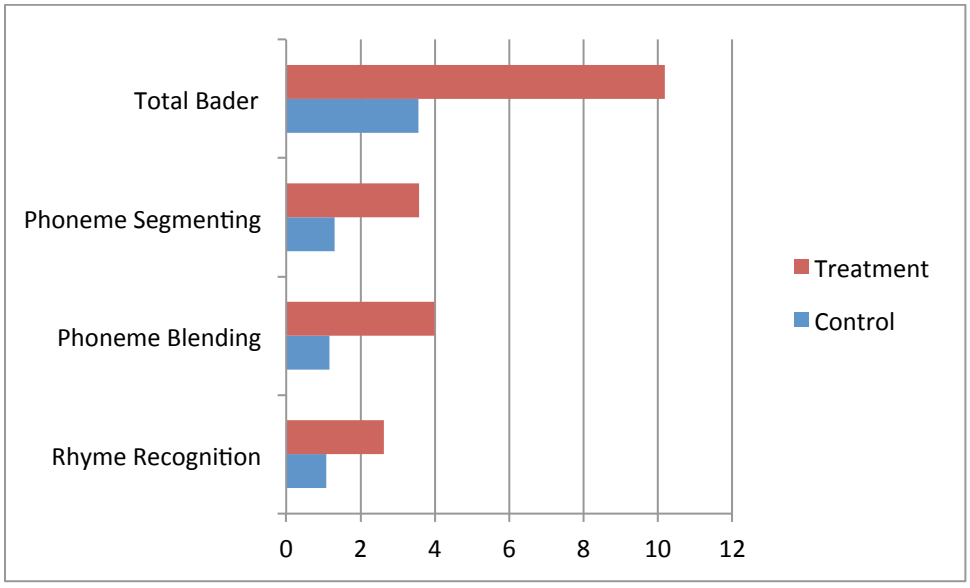


Figure12. Growth Rate Comparisons on the Bader

Summary and Conclusions

This final section of the Cohort 4 (C4) evaluation report summarizes:

- The data that were collected and analyzed;
- The analysis methods employed;
- C4 findings and trends in UPSTART implementation and usage; and
- C4 findings and trends in UPSTART's impact on the development of early literacy skills.

Data Collection

220 four year-old children were recruited for the C4 study; 117 treatment group children who had enrolled in UPSTART for Year 4 of the program and 103 control group children who had not enrolled in the UPSTART program. The children's parents were given an intake questionnaire (see Appendix A) at the time their children were pretested on the Brigance and Bader. The children were posttested on the Brigance and Bader in the summer of 2013.

The intake interview certified that 203 of the recruited children had no prior exposure to the UPSTART online curriculum. Complete pretest and posttest Brigance data were obtained and analyzed for all 203 of these children (101 treatment group children and 102 control children). On the Bader, complete pretest and posttest data were obtained for 172 children (79 treatment children and 93 control children). Measurement attrition on the Bader (16%) was caused primarily by children not being able to score on the pretest. There was no measurement attrition with the Brigance sample.

Data Analysis

To determine whether UPSTART children have better literacy skills at kindergarten compared to control group children, group equivalence on the pretests was first examined using independent sample t-tests. Relationships between the demographics and the posttest scores were then examined using correlation analyses.[10] Next, posttest differences between the treatment and control groups were examined for both the Brigance and Bader. Finally, posttest differences were re-examined by adjusting for initial differences between the treatment and control groups with the use of multiple regression analysis. The regressions used a hierarchical block design in which the pretest and a set of demographic covariates were entered first, followed by the treatment-control group.

To determine whether UPSTART students show stronger literacy growth rates from preschool to kindergarten relative to control students, paired sample t-tests were run to compare pretest and posttest scores for the matched Brigance and Bader treatment groups on the total test and each of the subtests. Statistically significant growth rates were determined by examining confidence

[10] It was necessary to transform a number of the demographic measures from nominal measures to scale measures by creating "dummy variables" on the basis of the dominant characteristics of the sample. For example, parent's marital status was transformed into whether the parent was married or not, or percent married.

intervals for the treatment and control groups for each test measure at the 99% confidence interval

The effect of UPSTART usage on reading proficiency was examined for UPSTART participants using an analysis of covariance (ANCOVA) in which usage levels were split into quartiles based on the usage distribution of the preschool analysis sample. The final ANCOVA models statistically controlled for initial literacy skills as measured by the pretest on each respective measure. The ANCOVA analyses estimated the effects of usage at quartiles one through three compared with usage at the fourth quartile, controlling for the children's initial level of literacy development.

UPSTART Implementation Results

The Waterford Institute provided documentation for a fourth-year UPSTART enrollment of 1, 250 children. A majority (63%) of the 1,250 preschool children that enrolled in the fourth year of UPSTART were from low-income families, according to data provided by the Waterford Institute. Slightly more girls (51%) were enrolled than boys (49%). In terms of ethnicity, the vast majority of the C4 enrollment was Caucasian (76%), 20% were Hispanic, and the remaining 4% were of Asian descent, African American, Native American, Pacific Islander, or of unknown origin.

As in previous years, most of the C4 participants (69%) received a computer drive with the UPSTART curriculum loaded on it. Approximately 11% of the fourth year participants received a computer loan and a free Internet subscription to help them access the UPSTART curriculum. Another 7.6% of the C4 participants were loaned a personal computer to use at home while participating in UPSTART. The remaining 12% of the fourth year participants were provided with various combinations of educational technology to enable them to access the UPSTART curriculum, including wireless and cellular devices.

Findings about UPSTART usage are summarized below.

- The C4 preschool analysis samples had a mean of approximately 76-78 hours of UPSTART curriculum usage over the fourth year of the project. This compares with an average of approximately 75 hours of instruction for program "graduates" and approximately 71 hours of instruction for all students enrolled in UPSTART in the fourth year of the program.

- UPSTART curriculum usage was significantly correlated with literacy skills at the beginning of kindergarten as measured by the Brigance ($r = .49$) and the Bader ($r = .44$). Controlling for initial literacy skills, the correlation of UPSTART usage with kindergarten outcomes was somewhat lower: $r=.35$ for the Brigance sample and $r=.38$ for the Bader sample.

- Controlling for initial levels of literacy development, UPSTART usage accounted for 13%-15% of the variance in literacy skills measured by the Bader and Brigance posttests respectively at kindergarten entry.

- There is an apparent linear trend in literacy development with increasing usage of the UPSTART curriculum.

- The UPSTART graduation rate continued to hold at 94% in Year 4. This is the same level of program attainment as realized in Year 3.

- UPSTART graduate status in the fourth year of the program was significantly correlated with hours of instruction (r=.55) as well as with literacy outcomes measured by the Brigance (r=.31) and the Bader (r=.35) at the beginning of kindergarten.

UPSTART Impact Results

- UPSTART participation in the fourth year of the program generally show small to medium size effects on improving the phonics skills of young preschool children. C4 effects as measured by the Brigance at the beginning of kindergarten were observed on the *Visual Discrimination* subtest, the *Letter Sounds* subtest, the Basic Pre-Primer *Vocabulary* subtest and on the Total Brigance assessment.

- The largest impact observed in UPSTART's fourth year of operation was in the development of *vocabulary*. This finding replicates results found in the third year evaluation.

- Overall, treatment group children scored an average of 29 points higher on the Brigance posttest (regression adjusted) compared to control group children. This finding is similar to the overall impact of UPSTART as found in the third year of the evaluation.

- Overall, UPSTART achieved medium size effects on improving the phonological awareness skills of participants in Year 4 of the program as measured by the Total Bader assessment.

- On average, treatment group children scored an average of almost six points higher on the Bader posttest (regression adjusted) compared to control group children. In terms of specific domains of phonological awareness, treatment group effects were observed on two of the three Bader subtests: *Phoneme Blending and Phoneme Segmentation*.

- UPSTART participants showed significantly stronger growth rates compared to control group children on the *Total Brigance* and four of the ten subtests, including: *Vocabulary, Survival Sight*

Words, Letter Sounds and Reciting the Alphabet. These results are similar to the findings about Brigance growth rates in the third year evaluation.

- UPSTART participants showed significantly stronger growth rates compared to control children on the *Total Bader* and two of the three Bader subtests, specifically *Phoneme Blending* and *Phoneme Segmenting*.

- Program effects for improvements in knowledge of letter sounds and vocabulary development are two of the early reading skills that the UPSTART curriculum focuses on. Similarly, program effects for phonemic blending and segmenting reflect UPSTART's focus on developing children's phonological awareness.

APPENDICES

Appendix A: Parent Intake Form

UPSTART Evaluation
Parent Intake Form

Please check the appropriate response with an "**X**". Choose only ONE response for each question.

1. Have any of your children participated in the UPSTART program in the past?
   □1 Yes □2 No

   1a. If yes, did your 4-year-old also use the program?
   □1 Yes □2 No

Child Information

2. What is your child's birthday? _____

3. What is your child's gender?
   □1 Male □2 Female

4. What year will your child be entering Kindergarten?
   □1 2011 □2 2012

5. What is your child's ethnicity?
   □1 Hispanic □2 Native American/Alaskan Native □3 Asian/Pacific Islander
   □4 Caucasian □5 African American □6 Other:_____

6. What is your child's primary language?
   □1 English □2 Spanish □3 Portuguese □4 Chinese
   □5 German □6 Japanese □7 Other:_____

7. Is your child currently attending a daycare/preschool?
   □1 Yes □2 No

   7a. **If yes**, approximately how many hours a week does your child attend a daycare/preschool?
   □1 less than 10 hours □2 10-19 hours □4 20-24 hours
   □5 25-29 hours □6 30-34 hours □7 35 or more hours

8. Does your child have access to a computer in your house?
   □1 Yes □2 No

9. Does your child use a computer in her/his day care or preschool?
   □1 Yes □2 No □3 Not Applicable (not in day care or preschool)

10. How comfortable is your child using a computer?
    □1 Very comfortable □2 Somewhat comfortable □3 Somewhat uncomfortable
    □4 Not comfortable □5 Very uncomfortable

Caregiver Information

11. What is your relation to the participating child?
    □1 Mother □2 Father □3 Grandmother □4 Grandfather
    □5 Step Father □6 Step Mother □7 Other:_____

12. What is your ethnicity?
    □1 Hispanic □2 Native American/Alaskan Native □3 Asian/Pacific Islander
    □4 Caucasian □5 African American □6 Other:_____

13. What is your primary language?
    □1 English □2 Spanish □3 Portuguese □4 Chinese
    □5 German □6 Japanese □7 Other:_____

14. What is the highest level of education you have completed?
    □1 Did not complete high school □2 High school diploma/GED □3 High school
    □4 Some college □5 Bachelor's degree □6 Masters degree □7 Doctorate

15. What is your paid employment status:
    □1 Full time □2 Part time □3 Not working

16. What is your spouse's paid employment status:
    □1 Full time □2 Part time □3 Not working
    □4 Not Applicable (single parent)

17. What is your marital status?
    □1 Married □2 Separated □3 Divorced □4 Unmarried

18. How many people live in your home (including you and all your children)?
    □1 One □2 Two □3 Three □4 Four □5 Five □6 Six or more

19. What is your total household annual income?
    □1 under $10,000 □2 $10,000-$24,999 □3 $25,000-$49,999
    □4 $50,000-$74,999 □5 $75,000-$99,000 □6 $100,000 or more

Thank you for participating in the Utah UPSTART Evaluation!

Appendix B: Group Differences for Brigance Tests

Brigance Sample: Treatment – Control Differences on Demographics

Covariate	*Group*	*N*	*Mean*	*SD*	*t*	*Significance*
Child is female	Treatment	101	.50	.50	-.91	NS
	Control	102	.57	.50		
Child is Caucasian	Treatment	101	.89	.31	1.90	NS
	Control	102	..79	.41		
Child is Hispanic	Treatment	101	.04	.20	-2.67	**
	Control	102	.15	.36		
Child's primary language is English	Treatment	101	1.00	.00	2.93	**
	Control	102	.92	.27		
Child currently attends preschool or daycare	Treatment	100	.31	.47	-7.97	**
	Control	100	.80	.40		
Child has access to a computer at home	Treatment	101	.97	.17	2.02	*
	Control	101	.90	.30		
Child comfort level with computers	Treatment	99	1.88	.82	-.54	NS
	Control	94	1.95	.92		
Parent is Caucasian	Treatment	101	.93	.26	-.01	NS
	Control	102	.85	.36		
Parent's primary language is English	Treatment	101	.97	.17	1.77	NS
	Control	102	.91	.29		
Parent Educational Attainment (recoded)[11]	Treatment	101	3.51	.66	1.78	NS
	Control	101	3.32	.84		
Parent is married	Treatment	101	.93	.26	2.55	**
	Control	101	.81	.39		
Parent is working	Treatment	101	.39	.49	-.85	NS
	Control	101	.46	.50		
Household size	Treatment	100	5.08	.92	1.91	*
	Control	101	4.81	1.07		
Household annual income category	Treatment	100	3.75	1.12	-.19	NS
	Control	101	3.78	1.26		

**p≤.01
*p≤.05

[11] 1 = High School Dropout; 2 = High School Graduate, 3 = Some College; 4 = College Graduate; 5 = Graduate Degree

Appendix C: Posttest Correlations

Pearson Correlations with Total Posttest Scores

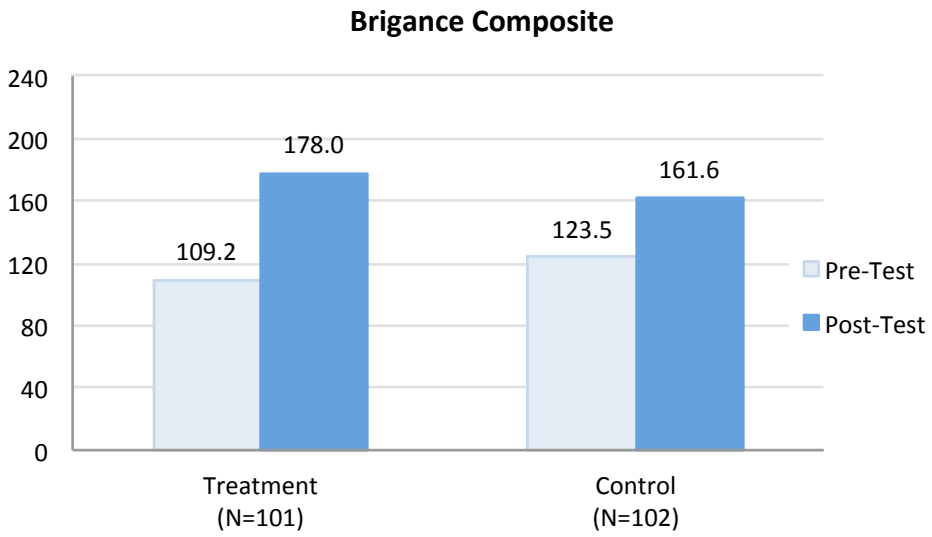
Variable	***Brigance***	***Bader***
Study Group[12]	.19**	.26**
Pretest	.68**	.48**
Child is Female	.11	.12
Child is Caucasian	.17**	.10
Child is Hispanic	-.21**	-.19**
Child attends daycare/preschool	-.13	-.11
Child has computer access at home	.05	. .07
Child's computer comfort	-.11	.01
Parent is employed	-.06	-.06
Household size	.02	-.02
Household income	.16*	.17*
Child's primary language is English	.16*	.18*
Parent is Caucasian	.15*	.12
Parent's primary language is English	.11	.17*
Parent Educational Attainment	.22**	.17*
Parent is married	.24**	.20**

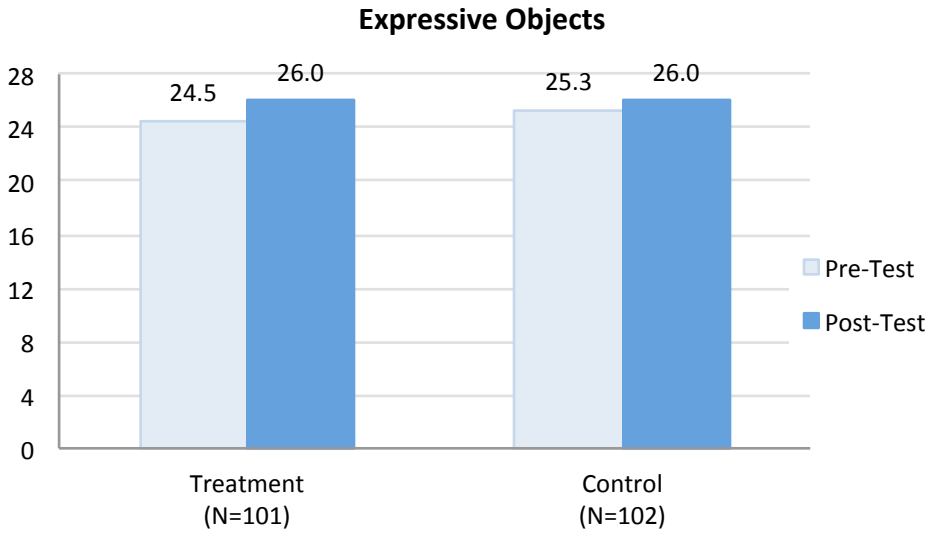
**p≤.01 *p≤.01

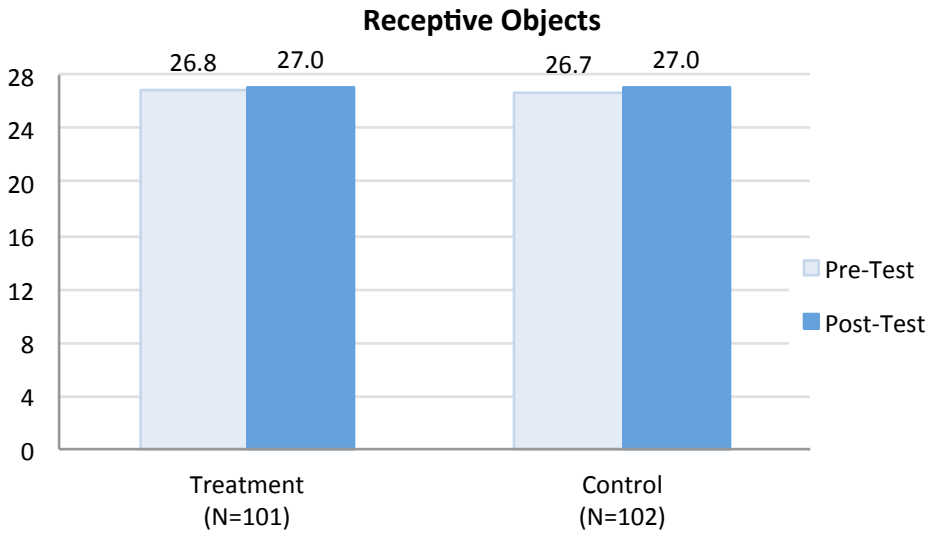
[12] Coded 1 if Treatment Group and 0 if Control Group

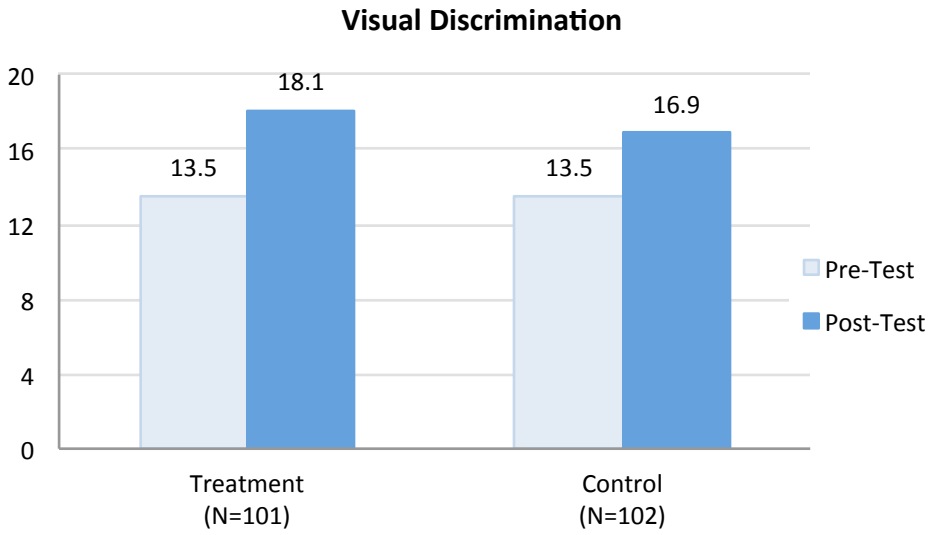
Appendix D: Brigance Subtest Mean Scores

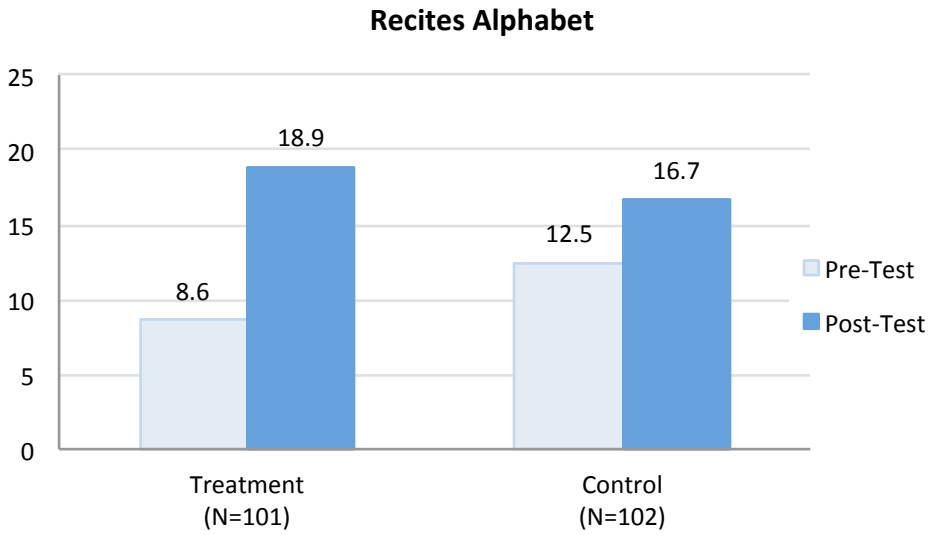
Brigance Subtest Mean Graphs: Pre- and Posttest Scores

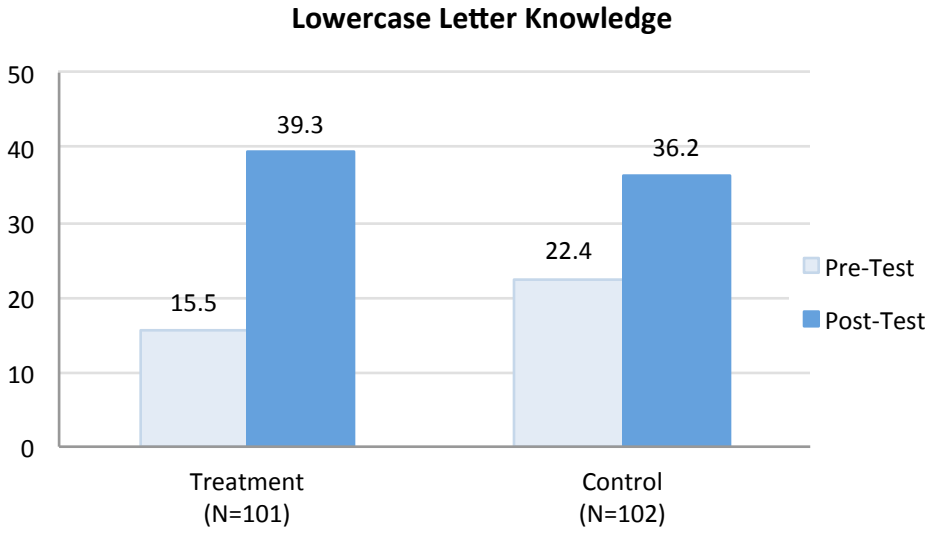


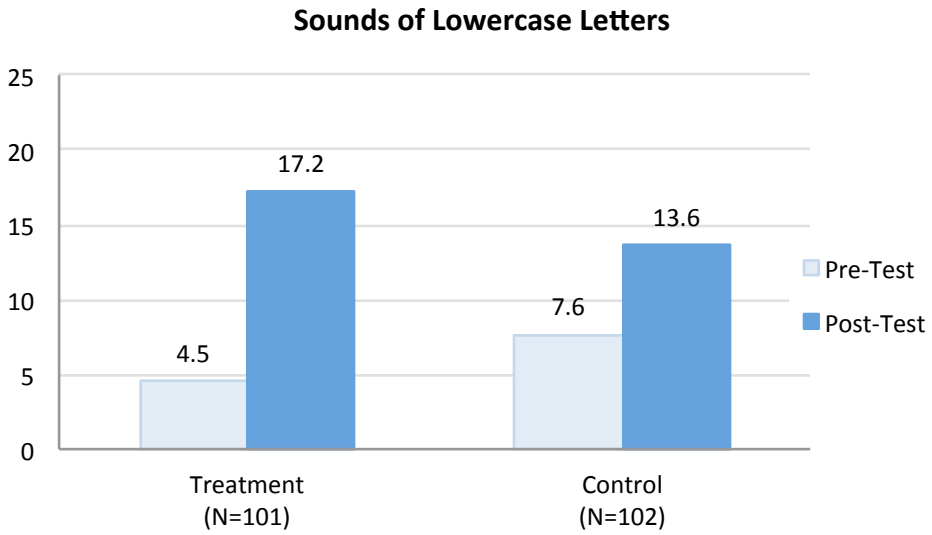


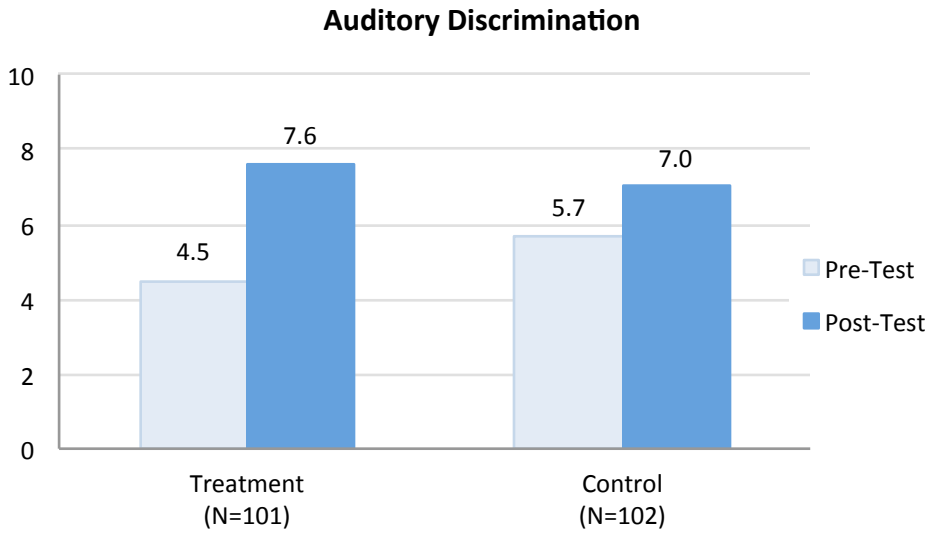

Receptive Objects
26.8
27.0
26.7
27.0
28
24
20
16
12
8
4
0
Pre-Test
Post-Test
Treatment
(N=101)
Control
(N=102)

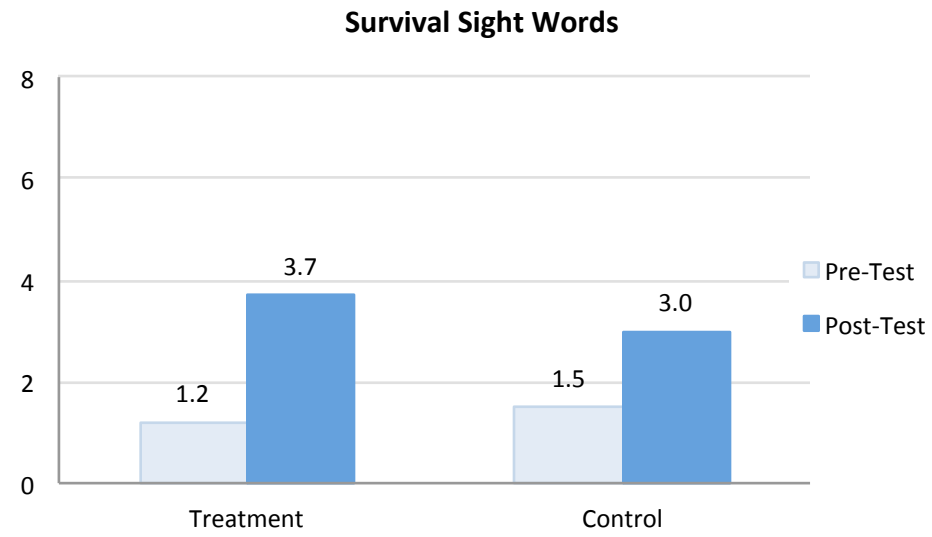

Visual Discrimination
18.1
16.9
13.5
13.5
20
16
12
8
4
0
Pre-Test
Post-Test
Treatment
(N=101)
Control
(N=102)

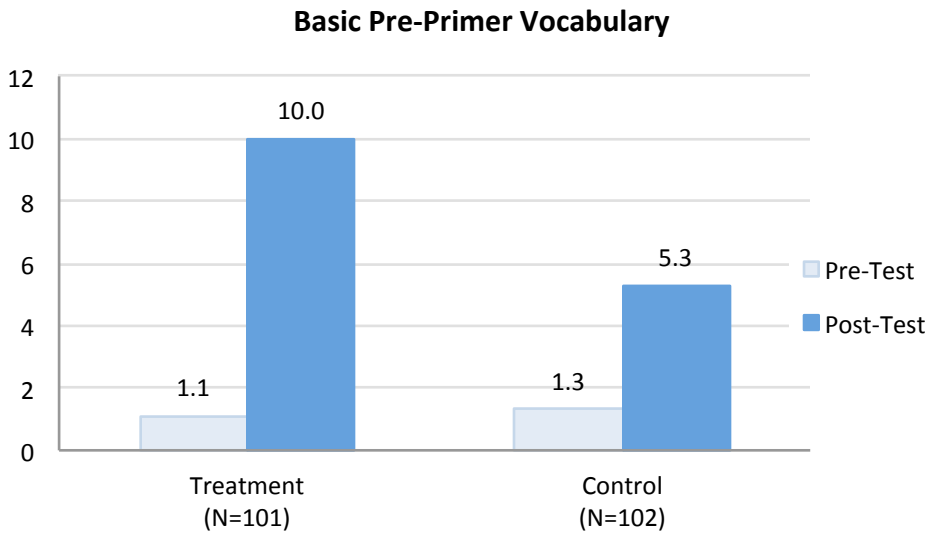

Recites Alphabet
25
20
15
10
5
0
18.9
16.7
12.5
8.6
Pre-Test
Post-Test
Treatment
(N=101)
Control
(N=102)


Lowercase Letter Knowledge
50
40
30
20
10
0
39.3
36.2
22.4
15.5
Pre-Test
Post-Test
Treatment
(N=101)
Control
(N=102)


Sounds of Lowercase Letters
25
20
15
10
5
0
17.2
13.6
7.6
4.5
Pre-Test
Post-Test
Treatment
(N=101)
Control
(N=102)


Auditory Discrimination
10
8
6
4
2
0
7.6
7.0
5.7
4.5
Pre-Test
Post-Test
Treatment
(N=101)
Control
(N=102)


Survival Sight Words
8
6
4
2
0
3.7
1.2
3.0
1.5
Pre-Test
Post-Test
Treatment
Control


Basic Pre-Primer Vocabulary
12
10
8
6
4
2
0
10.0
1.1
5.3
1.3
Pre-Test
Post-Test
Treatment
(N=101)
Control
(N=102)

Appendix E: Group Differences for Bader Tests

Bader Sample: Treatment – Control Differences on Demographics

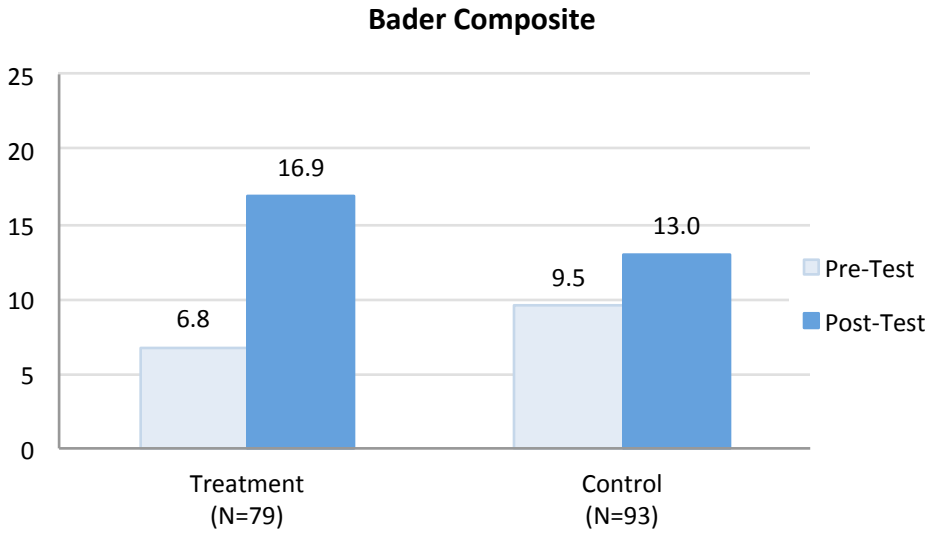
Covariate	*Group*	*N*	*Mean*	*SD*	*t*	*Significance*
Child is Female	Treatment	79	.53	.50	-.78	NS
	Control	93	.59	.49		
Child is Caucasian	Treatment	79	.90	.30	1.73	NS
	Control	93	.81	.40		
Child is Hispanic	Treatment	79	.03	.16	-2.84	**
	Control	93	.14	.35		
Child's primary language is English	Treatment	79	1.00	.00	2.52	**
	Control	93	.94	.25		
Currently attending preschool or daycare	Treatment	79	.30	.46	-7.01	**
	Control	91	.78	.42		
Child has access to a computer at home	Treatment	79	.96	.19	1.58	NS
	Control	92	.90	.30		
Child comfort level with computers	Treatment	77	1.91	.86	-0.07	NS
	Control	86	1.92	.90		
Parent is Caucasian	Treatment	79	.95	.22	2.03	*
	Control	93	.86	.35		
Parent's primary language is English	Treatment	79	.99	.11	2.30	*
	Control	93	.91	.28		
Parent Educational Attainment (recoded)[13]	Treatment	79	3.59	.63	2.21	*
	Control	92	3.35	.80		
Parent is married	Treatment	79	.95	.22	2.63	**
	Control	92	.83	.38		
Parent is employed	Treatment	79	.38	.49	-1.12	NS
	Control	92	.47	.50		
Household size	Treatment	78	5.05	.91	1.74	NS
	Control	92	4.78	1.08		
Household annual income category	Treatment	78	3.86	1.14	0.24	NS
	Control	92	3.82	1.23		

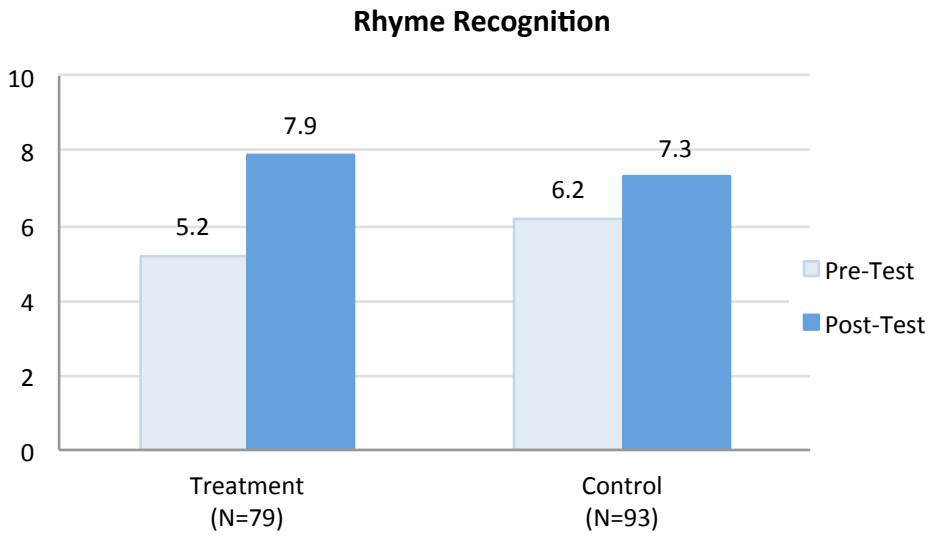
**p≤.01
*p≤.05

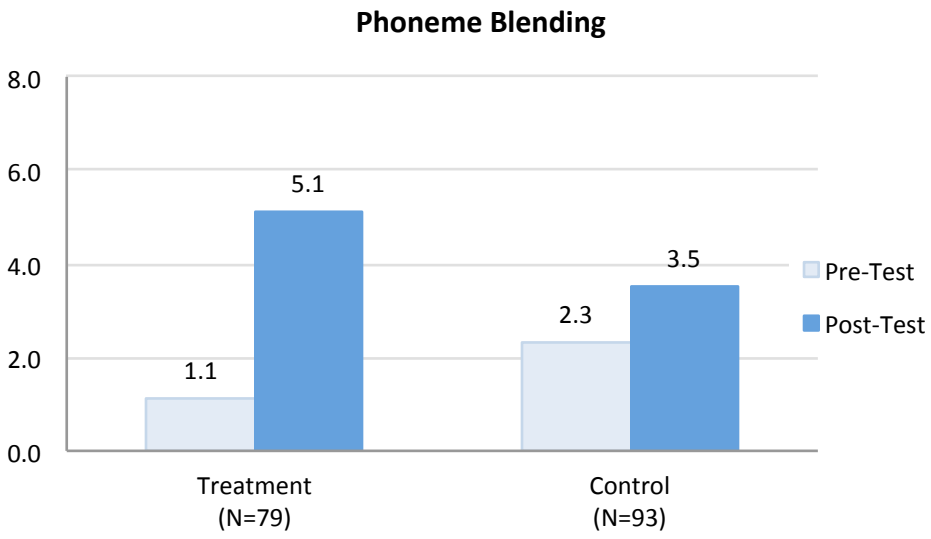
[13] 1 = High School Dropout; 2 = High School Graduate, 3 = Some College; 4 = College Graduate; 5 = Graduate Degree

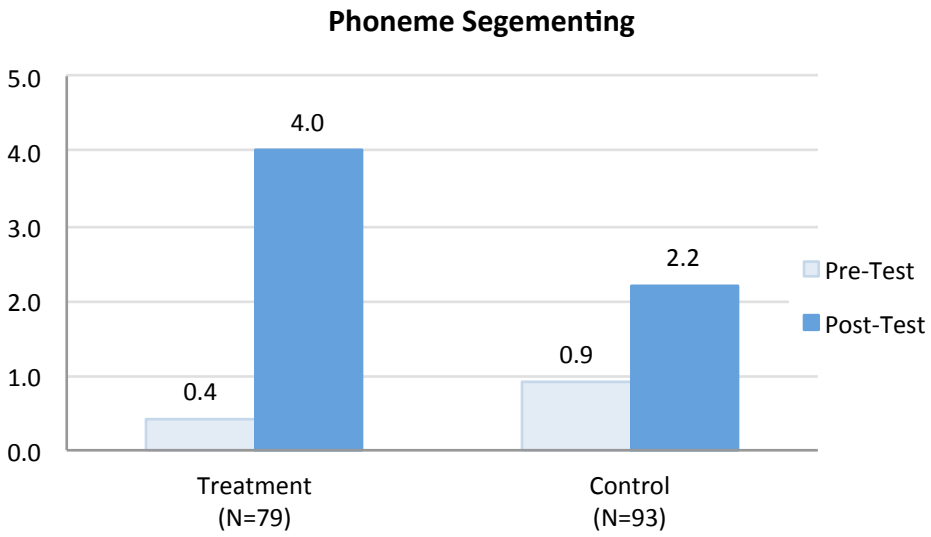
Appendix F: Bader Subtest Mean Scores

Bader Subtest Mean Graphs: Pre- and Posttest Scores






Phoneme Blending
8.0
6.0
4.0
2.0
0.0
5.1
3.5
2.3
1.1
Pre-Test
Post-Test
Treatment
(N=79)
Control
(N=93)


Phoneme Segementing
5.0
4.0
3.0
2.0
1.0
0.0
4.0
2.2
0.9
0.4
Pre-Test
Post-Test
Treatment
(N=79)
Control
(N=93)