Utah UPSTART Program Evaluation Program Impacts on Early Literacy

Third Year Results
Cohort 3 Technical Report

March 2013

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

Established as a pilot demonstration project by the Utah state legislature, UPSTART uses home-based educational technology to develop the school readiness skills of preschool children. The Waterford Institute provided documentation for a third-year UPSTART enrollment of 1,168 children. A majority (64%) of the 1,168 preschool children that enrolled in the third year of UPSTART were from low-income families, according to data provided by the Waterford Institute. Slightly more boys (53%) were enrolled than girls (47%). In terms of ethnicity, the vast majority of the Cohort 3 (C3) enrollment was Caucasian (79%), 16% were Hispanic, 3% were of Asian descent, 1% were Black, and 1% were Native American. Ethnicity was unknown for 1% of the third year enrollment.

Evaluation Design

The evaluation of UPSTART's third 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 C3 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, 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 (the Brigance and the Bader) and selected demographic characteristics that independently of usage were found to influence posttest outcomes. 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.

Descriptive statistics were also computed to describe the population of students that enrolled in the third year of UPSTART (i.e., Cohort 3). The descriptors included student demographics, the equipment that Cohort 3 (abbreviated as C3) students received, hours of UPSTART curriculum usage, and the graduation status of C3 students. Graduates were participants that met the program's curriculum usage criteria.

Data Collection

305 four year-old children were recruited for the C3 study; 151 treatment group children who had enrolled in UPSTART for Year 3 of the program and 154 control group children who had not enrolled in the UPSTART program. The children's parents were administered 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 approximately one year after their pretest. Complete pretest and posttest data were obtained and analyzed for 259 children on the Brigance (129 treatment group children and 130 control children) and for 232 children on the Bader (112 treatment children and 120 control children).

Data Analysis

To determine whether UPSTART children have better literacy skills at kindergarten compared to control group children, group equivalence on the pretests was examined using independent sample t-tests. Relationships between the demographics and the 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 was entered first, followed by a set of demographic covariates, followed by the treatment-control group.

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 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 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 and selected demographic characteristics (e.g., access to a computer at home). The ANCOVA analyses estimated the effects of usage at quartiles one through three compared with

2

usage at the fourth quartile, controlling for covariates entered in the final model. Importantly, it was determined that the ANCOVA analyses were statistically underpowered with the size of the final analysis samples obtained from the Bader (N=112) and the Brigance (N=129).

Results

As in previous years, most of the C3 participants (71%) received a computer drive with the UPSTART curriculum loaded on it. Approximately 12% of the third year participants received a computer loan and free internet access to help them access the UPSTART curriculum. Another 7% of the C3 participants were loaned a personal computer to use at home while participating in UPSTART. The remaining 10% of the third 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.

- Compared to previous years, the C3 participants logged a greater number of instructional hours with the UPSTART curriculum. The C3 preschool analysis samples had a mean of approximately 75 hours of UPSTART curriculum usage over the third year of the project. This compares with an average of approximately 70 hours of instruction for program "graduates" and approximately 67 hours of instruction for all students enrolled in UPSTART in the third year of the program.

- Length of participation in the UPSTART curriculum was significantly and positively correlated with literacy skills at the beginning of kindergarten as measured by the Brigance (r=.40) and the Bader (r=.20).

- Early language and cognitive literacy skills measured by the Brigance generally improved with increasing levels of UPSTART curriculum usage. This was not the case with the phonological awareness skills measured by the Bader, which did not improve significantly with increasing levels of UPSTART curriculum usage. Both of these findings replicate the C2 usage results.

- The UPSTART graduation rate has continued to rise each year: from 59% in Year 1, to 76% in Year 2, to 94% in Year 3. This indicates that UPSTART is making very good progress in achieving the curriculum usage goals set for program implementation.

- UPSTART graduate status in the third year of the program was significantly correlated with hours of instruction but not with literacy outcomes measured by the Brigance or the Bader at the beginning of kindergarten. Both of these findings replicate the C2 usage results.

Findings about UPSTART's impact on literacy skills measured at the beginning of Kindergarten are summarized below. *Compared to C2, the C3 impacts were stronger and observed across the early literacy domains measured by the administered assessments.*

- UPSTART participation generally continues to show moderately strong impacts on improving the phonics skills of young preschool children. With one exception, these effects were observed across all domains assessed by the Brigance with the C3 group at the beginning of Kindergarten. The largest impact observed in UPSTART's third year of operation was in the development of vocabulary. Overall, treatment group children scored an average of 28 points higher on the Brigance posttest compared to control group children.

- Overall, UPSTART achieved large impacts on improving the phonological awareness skills of participants in Year 3 of the program. Treatment group effects were observed across all three of the phonological awareness domains measured by the Bader: the ability to recognize rhymes; the ability to blend phonemes; and the ability to segment phonemes.

- UPSTART participants showed significantly stronger growth rates compared to control group children on the Total Brigance and five of the ten subtests, including: Vocabulary, Sight Words, Auditory Discrimination, Letter Sounds and Reciting the Alphabet.

- UPSTART participants showed significantly stronger growth rates compared to control children on the Total Bader and all of the Bader subtests, including Rhyme Recognition, Phoneme Blending and Phoneme Segmenting.

Based on the third 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.

Recommendations
Programmatically, the one recommendation that could help the participants improve their phonics and beginning reader skills is to offer incentives for reaching usage milestones that appear correlated with literacy achievement. These might include rewarding the attainment of three successive levels of curriculum usage, awarded at the following milestones:
- 65 hours
- 75 hours
- 85 hours

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 third year of operation during the 2011-12 school year, the project's implementation contractor – the Waterford Institute – enrolled 1,168 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 1,168 children enrolled in the third year cohort, hereafter referred to as C3, participated in the UPSTART program from October 2011 through June 2012.

The evaluation of UPSTART's third 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 usage 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 usage of the program's curriculum.

Slightly more C3 boys (53%) were enrolled than girls (47%). In terms of ethnicity, the vast majority (79%) of the C3 enrollment was Caucasian, 16% of the children were Hispanic, 3% were of Asian descent, 1% were Black, and 1% were Native American. Ethnicity for approximately 1% of the third year enrollment was unknown. The primary language spoken by the vast majority of the C3 children was English (88%); approximately 11% of the C3 children spoke Spanish. Seven percent of the C3 children had a diagnosed disability, mostly speech impairments.

A majority (64%) of the 1,168 preschool children that enrolled in the third year of UPSTART were from low-income families, according to data provided by the Waterford Institute. Most commonly, the C3 parents had some college (40%) but did not achieve a college degree. The vast majority of the C3 parents were married.

Background: The Year 2 Evaluation

Previously, a pretest-posttest control group design was used in evaluating the second year (C2) of UPSTART. In the C2 evaluation, children were pretested in the fall of their preschool year (2010) and posttested in the spring of the preschool year (2011) 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 C2 evaluation, complete pretest and posttest data were obtained for a sample of 76-77 UPSTART children (the treatment group), depending on the test, and 82 non-UPSTART children (the control group). *The same evaluation design was used in the C3 evaluation during the 2011-12 preschool year but with larger sample sizes.*

In the second year of UPSTART, slightly more than three-quarters (76%) of the C2 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 second year was found to be significantly and positively correlated with literacy achievement on the Brigance (r=.44) and the Bader (r=.22). C2 students logged an average of 49 hours of instruction in the program during 2010-11. Further analysis controlling for initial levels of literacy development showed that increasing levels of curriculum usage were related to overall early literacy achievement as measured by the Brigance. However, this relationship could not be established for the Bader in the C2 evaluation, possibly due to it being a more limited assessment scale with a weaker measured relationship to early literacy development.

The C2 evaluation statistical models controlled for initial levels of literacy development and further showed that participation in UPSTART had a moderately strong impact on improving the treatment group children's knowledge of alphabetic letters and letter sounds compared to control children. Overall, treatment group children scored approximately 25 points higher on the Brigance posttest compared to control children. C2 statistical modeling also showed a relatively small program impact for improving the early literacy skills measured by the Bader, which overall favored the treatment group children by 4 points compared to the control group children. This impact could be accounted for by significantly better performance in the ability of treatment students to blend phonemes on the Bader posttest.

C3 Evaluation Design

The Cohort 3 evaluation continued to use the quasi-experimental research design variant of the nonequivalent comparison group design described above. Specifically, the design implemented in the C3 evaluation (as in C2) 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 control children were recruited from local preschools and the treatment children were recruited from families enrolling in UPSTART. The C3 study recruited 305 four year-old children; 151 treatment group children who had enrolled in UPSTART for Year 3 of the program and 154

nonparticipating control group children. This was a 60% increase in the size of the C3 study sample compared to the C2 evaluation.[1]

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 2011 for C3 treatment and control group children; for the C2 controls, pretests were administered in the fall of 2010. O2 indicates posttest measurements taken in the fall of 2011 for the C2 controls and in the fall of 2012 for the C3 treatment and control group children. Regardless of the cohort, the interval between pretests and posttests was approximately 12 months.

```
NR    T      01     X      O2
---------------------------------------------
NR    C      O1            02
```
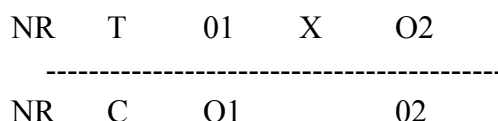
The use of both a pretest and a comparison group facilitates our ability to examine potential threats to validity which could jeopardize a clear interpretation of the results.[2] Because the study is not a randomized control trial, the groups are nonequivalent by definition, and consequently selection 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 it 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

---

[1] The larger evaluation sample was accomplished in part by recruiting 60% more treatment group children from the C3 enrollment; by recruiting 62 nonparticipating children for the control group through local preschool referrals; and by including the data already collected on the 92 control children from the C2 study.
[2] See Shadish, Cook, and Campbell (2002). *Experimental and Quasi-Experimental Designs for Generalized Causal Inference*. Boston: Houghton Mifflin Company.

expected to show greater gain scores (posttest score minus pretest score) relative to the comparison group on the various subtests and total test scores.

With respect to concerns for school readiness, our research questions for the C3 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 usage 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 usage 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 usage 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[3] 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)

---

[3] 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.

- *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)
- *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 C3 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

As previously stated, 305 four year-old children were recruited for the C3 study; 151 treatment group children who had enrolled in UPSTART for Year 3 of the program; 60 control group children who had not enrolled in the UPSTART program; and 91 controls who had previously been tested in the C2 evaluation. The children's parents were administered an intake questionnaire (see Appendix A) in the fall of 2010 and 2011at the time their children were pretested on the Brigance and Bader. The order of testing was Brigance first, followed by the

Bader. The children were posttested on the Brigance and Bader in the fall of 2011 (C2 controls) and the fall of 2012 (all C3 treatment and controls).

Complete pretest and posttest data for the C3 evaluation were obtained for 259 children on the Brigance (129 treatment children and 130 control children) and for 232 children on the Bader (112 treatment children and 120 control children). Study attrition due to incomplete test data is discussed below.

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 complete pretest and posttest data.

Attrition Analysis

An initial sample of 305 preschool children was recruited for the C3 evaluation study. Three children in the initial control group sample had prior exposure to the UPSTART curriculum, probably from siblings that previously participated in the evaluation as members of the treatment group. These three cases were deleted from the analysis file. That left a revised base sample of 151 treatment group children and 151 control group children for a total sample of 302 children who were 4 years old at the time of their pretest with the Brigance and the Bader.

The final analysis samples consisted of children who had valid matched pretest and posttest scores on the Brigance and on the Bader. Invalidating events included (a) failing to show up for posttesting and (b) not being able to score on the test. The latter was reflected by a score of zero on the test, indicating a "test floor effect" and that the test was too difficult for the child. The observed attrition on each test was examined using an SPSS case summary analysis, the results of which are presented in Table 1 for the Brigance and in Table 2 for the Bader.

Table 1
C3 Brigance Testing

Attrition Indicator	Treatment Group N	Control Group N	Total Sample N
Pretest Base	151	151	302
Valid Pretests	151	151	302
Posttest Base	129	130	259
Valid Posttests	129	130	259
Matched Pretests/ Posttests	129	130	259

Table 1 shows that the C3 evaluation started off with a pretest base sample of 151 treatment children and 151 control group children for a total of 302 children pretested. All of the children

11

were able to score on the Brigance Pretest. Approximately 15% of the pretested children did not return for posttesting, about equally divided among the treatment and control children. All of the children who posttested on the Brigance were able to score on the test.

Brigance attrition was computed to be 15%. The C3 evaluation started with a recruitment base of 305 children and ended up with a final Brigance analysis sample of 259 children. Thus:

- 259/305 = .849
- 1-.849 = .15 = 15%

Table 2
C3 Bader Testing

Attrition Indicator	Treatment Group N	Control Group N	Total Sample N
Pretest Base	151	151	302
Valid Pretests	130	142	272
Posttest Base	129	130	259
Valid Posttests	128	127	255
Matched Pretests/Posttests	112	120	232

Table 2 shows a more complex set of events surrounding attrition on the Bader, which begins with a pretest base of 151 children per group. We then encounter the situation where a number of the children cannot score on the Bader pretest. Bader pretest floor effects are more prevalent among the C3 treatment group children whereas failing to show up for the Bader posttest is more of a problem with the control group families. The Bader analysis sample ends up valid matched pretest and posttest scores on 232 children: 112 treatment group children and 120 control group children.

Bader attrition was computed to be 24%. The C3 evaluation started with a recruitment base of 305 children and ended up with a final Bader analysis sample of 232 children. Thus:
- 232/305 = .76
- 1-.76 = .24 = 24%

The final analysis in the C3 evaluation used the data collected from those children with a valid matched pretest and posttest 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 UPSTART 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 which are significantly related to the posttests (e.g., the mother's marital status).

If the treatment and control groups are essentially similar at the beginning 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.[4] Initial between group differences were found on 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-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 was entered first, followed by a set of demographic covariates, followed by the treatment-control group. Effect size estimates are graphically presented for all Brigance and Bader posttest differences between the treatment and control groups.

---

[4] 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.

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.[5] 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 usage (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.

Results

Findings are reported first in answering research questions 3, 4, and 5 about the implementation of the UPSTART program in its third 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 questions 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.

---

[5] 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.

UPSTART Implementation

Findings reviewed under UPSTART implementation include enrollment in the third 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 third-year UPSTART enrollment of 1,168 children. Some basic demographic characteristics of the C3 population are presented below in Table 3 compared to the C3 analysis sample as estimated from the Brigance subgroup (n=129).

Table 3
Basic Demographic Characteristics: C3 population vs. C3 Analysis Sample

Demographic Categories		All UPSTART (N=1,168)	Analysis Sample (N=129)
Child's Gender	Boys	52.9%	56.6%
	Girls	47.1%	43.4%
Child's Ethnicity	Caucasian	78.9%	89.8%
	Hispanic	15.6%	3.1%
	Asian	2.7%	0.8%
	Black	0.9%	0.8%
	Native American	0.8%	--
	Other	1.1%	5.5%
Child's Primary Language	English	87.9%	99.2%
	Spanish	11.4%	--
	Other	0.7%	0.8%
Parent Educational Attainment	High School	17.7%	5.4%
	Some College	39.6%	24.8%
	College Graduate	34.5%	51.2%
	Advanced Degree	8.1%	18.6%
Parent Marital Status	Married	90.2%	95.3%
	Other	9.8%	4.7%

As shown in Table 3, the analysis sample was somewhat of a more advantaged subgroup compared to the C3 population. For example, the C3 population is under-represented in the analysis sample in terms of Hispanics and Spanish speakers. The analysis sample is also more highly educated than the C3 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 4 for all 1,168 students enrolled in the third year and for the C3 preschool sample (using the pretest base). The majority of the third year UPSTART students (approximately 71%) 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 C3 preschool sample most often (81%) also received a computer drive with the curriculum loaded on it.

UPSTART provided personal computers to almost 12% of the C3 students and gave them free access to the Internet while they used the equipment. This was also the case with about 5% of the C3 preschool sample. Another 7% of the C3 program participants were given access to a home computer for free while they participated in the program. Among the C3 preschool sample, 5% were given access to a home computer for free while they participated in the program. The remaining 5% of the C3 enrollment received various combinations of computer technology to enable them to access the UPSTART curriculum.

Table 4
Percent of C3 Students Provided Equipment by UPSTART

Equipment Provided	All UPSTART	Preschool Sample
Drive	71.3	81.5
Computer & Internet	11.6	5.3
Computer	7.4	5.3
Computer & Wireless	3.9	4.6
Computer & Cellular	2.5	0.7
Internet & Drive	1.7	2.0
Cellular & Drive	1.1	--
Computer with Wireless & Internet	0.3	0.7
Other	0.1	--
Sample Size	N = 1,168	N = 151

UPSTART Graduates

Of the 1,168 children documented as enrolled in UPSTART in the third year of the program, the Waterford Institute classified 1,099 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 3 achieved a *graduation rate of 94%* (i.e., 1,099/1,168 = .94, or 94%).

It should be noted that 98% of the preschool analysis sample was composed of program graduates. For both the Brigance and Bader analysis samples, UPSTART graduate status in year 3 of the program was significantly correlated with hours of instruction (r=.38 and r=.43 respectively, p <.01) but not with total posttest scores on the Brigance or the Bader.

UPSTART Usage

The hours of instruction observed for all children documented to be enrolled in the third year of UPSTART are summarized in Table 5 compared to "graduates" and the children in the C3 preschool analysis samples. The average level of usage for all students enrolled in the third year of UPSTART was approximately 67 hours of instruction. The C3 academic year covered 39 weeks of instruction, beginning the week of October 3, 2011 and ending the week of June 25, 2012. The children in the C3 analysis samples used the UPSTART curriculum for approximately 75 hours of instruction on the average. C3 graduates used the UPSTART curriculum for approximately 70 hours of instruction on the average.

Table 5
C3 Hours of UPSTART Instruction

Group	N	Mean	SD	Range
All UPSTART	1,168	66.52	23.79	<1 – 175.43
UPSTART Graduates	1,099	69.64	20.69	17.05 - 175.43
Brigance Analysis Sample	129	74.50	17.47	9.32 – 123.18
Bader Analysis Sample	112	75.68	17.04	9.32 – 123.18

The histograms in Figures 1-4 show the distribution of hours of instruction for the total C3 population (Figure 1), the C3 graduates (Figure 2), and the C3 analysis samples (Figure 3 for the Brigance and Figure 4 for the Bader). All four histograms show hours of instruction to be essentially normally distributed. However, each group has slight variations in skewness, as described below.

C3 Population Usage. In the C3 population (see Figure 1), UPSTART curriculum usage is normally distributed with an average usage level of approximately 67 hours. However, there is a very slight negative skew in the instructional hours distribution (skew statistic = -.051) with a small "pile-up" at the left end of the scale representing a subgroup four children with zero hours of instruction.

Figure 1. Hours of Instruction for All Students Enrolled in UPSTART in Year 3
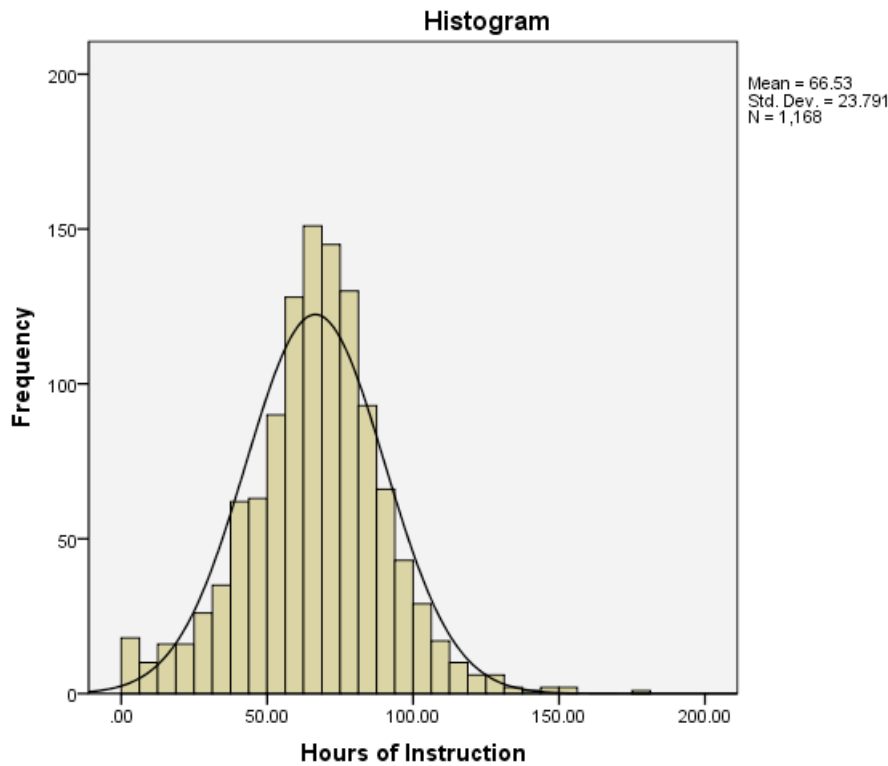
The bottom quartile of the C3 population completed less than 53.5 hours of instruction. The midpoint of the C3 population distribution (the median) was 67.25 hours of instruction. The top quartile of the C3 population completed in excess of 80.67 hours of instruction.

C3 Graduate Usage. UPSTART curriculum usage for the subset of graduates was normally distributed with an average usage level of approximately 70 hours. The graduate usage distribution has a slight positive skew (skew statistic = .480) with two graduates logging over 150 hours and one student logging over 175 hours. The bottom quartile of the C3 graduates ranges from 17 hours to 56.9 hours. The midpoint of the C3 graduate distribution (the median) was 69 hours of instruction. The top quartile of the C3 graduates completed in excess of 81.6 hours of instruction.

*Figure 2. Hours of Instruction for UPSTART Graduates in Year 3*
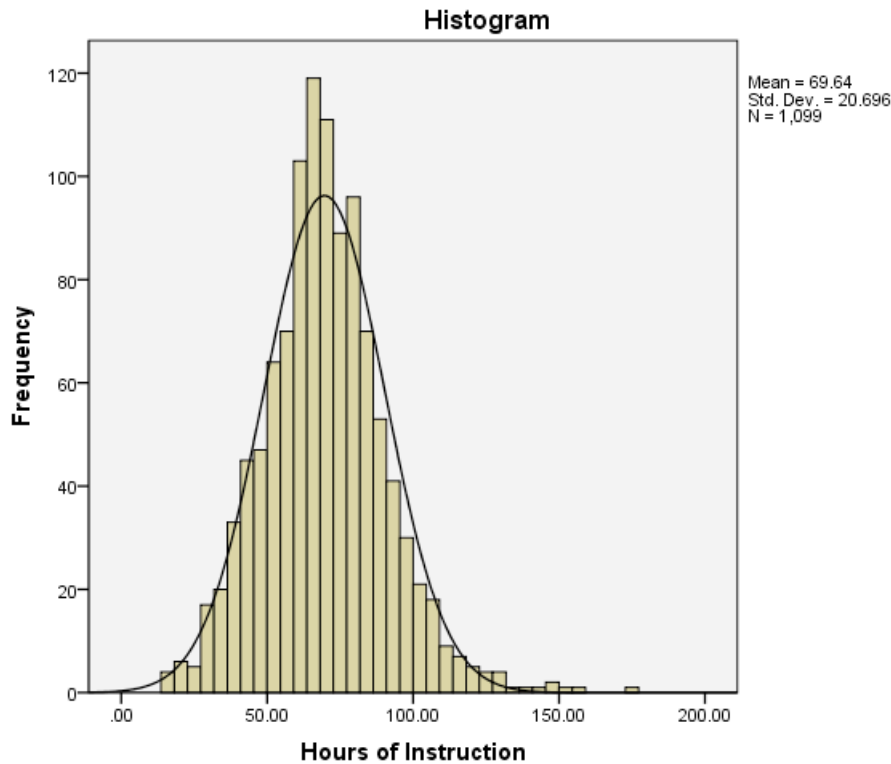
C3 Analysis Sample Usage. UPSTART curriculum usage for the C3 analysis samples were normally distributed with slight negative skews. The average usage level for the Brigance sample was approximately 75 hours of instruction; for the Bader sample, the average usage was approximately 76 hours. The analysis samples show a very slight negative skew with one child logging only 9.32 hours of instruction.

UPSTART usage for the C3 Brigance analysis sample (see Figure 3) is left-skewed (skew statistic = -0.080) with a mean of approximately 75 hours of instruction and a standard deviation of 17 hours. The Brigance analysis sample's median is approximately 73 hours of instruction. For the usage analysis with the Brigance sample, hours of instruction are distributed as follows by quartile of usage:

- 1st Quartile: 9.3 hours to 62.3 hours
- 2nd Quartile: 62.4 hours to 73.4 hours
- 3rd Quartile: 73.47 hours to 83.65 hours
- 4th Quartile: 83.87 hours to 123.2 hours

Figure 3. Hours of Instruction for C3 Brigance Analysis Sample
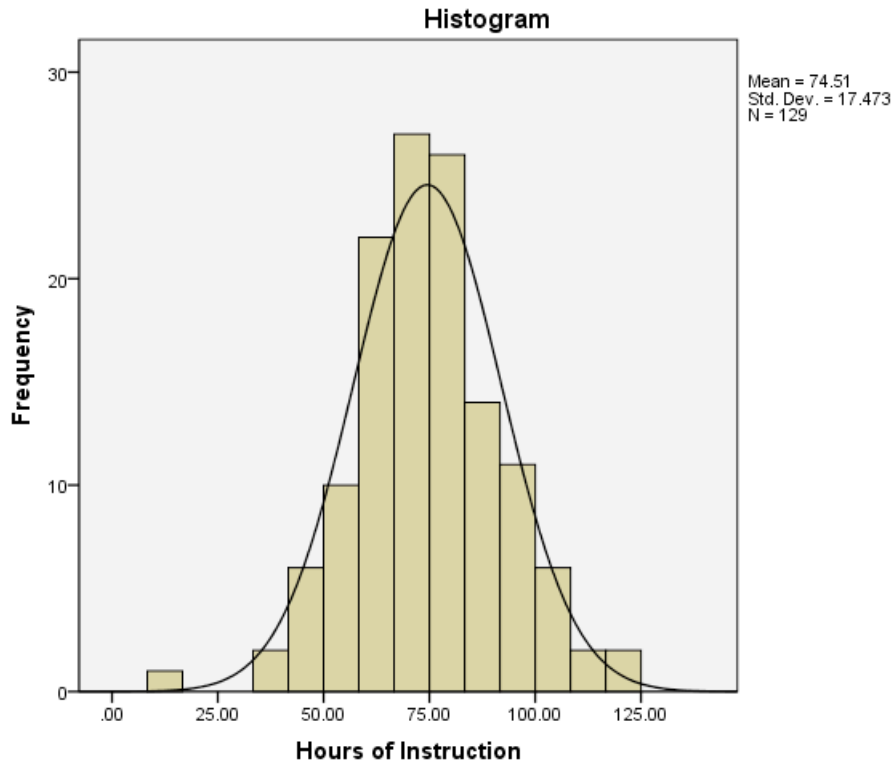
UPSTART usage for the C3 Bader analysis sample (see Figure 4) is left-skewed (skew statistic = -0.200) with a mean of approximately 76 hours of instruction and a standard deviation of 17 hours. The Bader analysis sample's median is approximately 75 hours of instruction. For the usage analysis with the Bader sample, hours of instruction are distributed as follows by quartile of usage:

- 1st Quartile: 9.3 hours to 64.5 hours
- 2nd Quartile: 64.6 hours to 74.6 hours
- 3rd Quartile: 74.7 hours to 84.7 hours
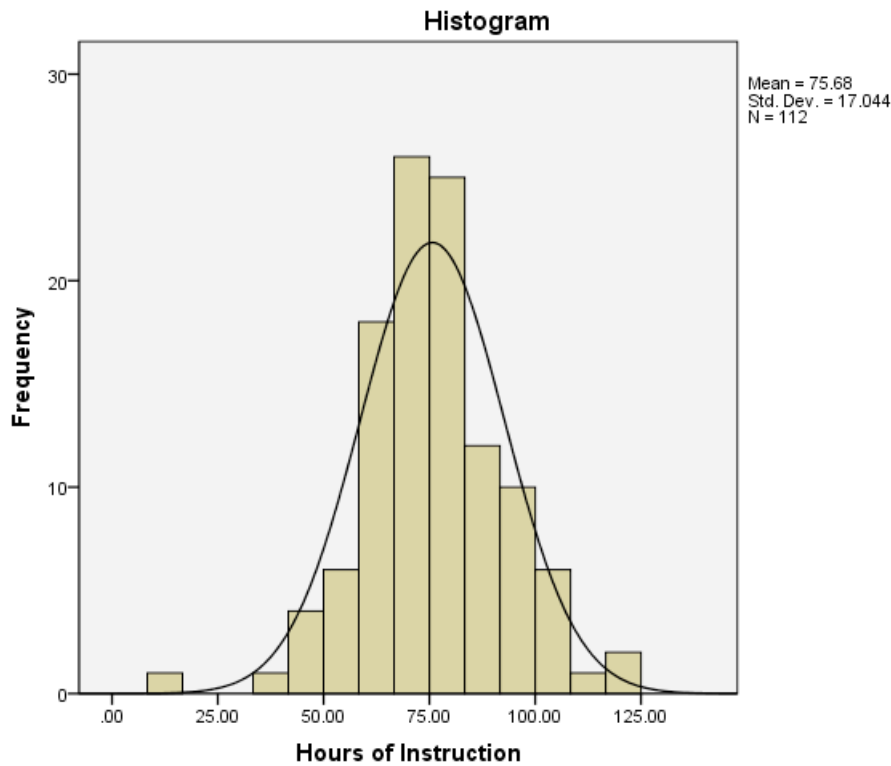- 4th Quartile: 84.8 hours to 123.2 hours

Figure 4. Hours of Instruction for C3 Bader Analysis Sample

How UPSTART Usage Relates to Literacy Outcomes

The preschool analysis established that UPSTART curriculum usage is positively and significantly correlated with literacy outcomes as measured by total posttest scores on the Brigance and the Bader. The relationship between UPSTART usage and literacy outcomes measured by the Brigance Total Posttest was moderately strong ($r=.40$, $p<.01$) whereas the relationship between UPSTART usage and literacy outcomes measured by the Bader Total Posttest were not quite so robust ($r=.20$, $p<.05$). Correspondingly, it was found that increases in total literacy scores on the Brigance were significantly related to increasing levels of UPSTART curriculum usage, but not so for the Bader. These two findings are discussed below.

UPSTART Usage Effects Measured by the Brigance. Table 6 shows that UPSTART usage is significantly and positively related[6] to posttest literacy outcomes measured by the Brigance Total Posttest, statistically controlling for initial levels of literacy as measured by total scores on the Brigance pretest and the availability of a computer for use in the home prior to enrollment in

---

[6] The Partial Eta Square statistic for Usage Group in Table 6 indicates that increasing exposure to the UPSTART curriculum accounts for about 7% of the literacy skills measured by the Brigance at the beginning of kindergarten. This suggests a relatively small effect of UPSTART usage by itself. See further discussion of usage effects.

UPSTART.[7] Sample size for the Brigance usage analysis was 129 and the observed statistical power was somewhat less than optimal[8] for assessing the usage factor usage employing an analysis of covariance (ANCOVA) approach with two covariates.

Table 6
Tests of Between Subjects Effects: Brigance Total Posttest - Beginning K-

Source	Type III Sum of Squares	df	Mean Square	F	Sig.	Partial Eta Squared
Corrected Model	102123.061	5	20424.612	35.226	.000	.589
Intercept	36329.219	1	36329.219	62.656	.000	.337
Brigance_Pre	53580.241	1	53580.241	92.408	.000	.429
Computer_Home	2868.738	1	2868.738	4.948	.028	.039
Usage Group	5550.811	3	1850.270	3.191	.026	.072
Error	71317.962	123	579.821			
Total	5040132.000	129				
Corrected Total	173441.023	128				

Adjusted R Squared = .57; Computed using alpha = .05

In Table 7 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 C3 analysis sample, controlling for the influence of initial literacy skills and the effect of prior computer access in the home, and displays the difference in Brigance total posttest scores in the column labeled B – expressed as a regression coefficient. In terms of usage effects, the parameter estimates in Table 7 essentially indicate that literacy achievement as measured by total scores on the Brigance are significantly different between usage quartiles 1 and 4 -- a difference of 19 points. Table 7 also shows that Brigance literacy achievement is not significantly different between usage quartiles 2, 3, and 4.

---

[7] Exploratory analyses that controlled for hours of instruction indicated that Brigance posttest performance was most directly influenced by initial literacy ability, prior access to a computer in the home, the child's ethnicity (if Hispanic), and household size. The latter two covariates proved to be statistically non-significant as control measures when entered into the regression model in conjunction with pretest scores and the prior availability of a computer in the home.

[8] The observed power for the analysis was .76; the desired power for a given analysis is .80 or higher.

Table 7
Parameter Estimates: Brigance Total Posttest Score - Beginning K-

Parameter	B	Std. Error	t	Sig.	Partial Eta Squared	Observed Power
Intercept	105.775	13.710	7.715	.000	.326	1.00
Brigance_Pre	.547	.057	9.613	.000	.429	1.00
Computer_Home	25.060	11.266	2.224	.028	.039	.598
[Usage_Group=1.00]	-19.174	6.653	-2.882	.005	.063	.816
[Usage_Group=2.00]	-6.729	6.100	-1.103	.272	.010	.195
[Usage_Group=3.00]	-3.150	6.134	-0.514	.609	.002	.080
[Usage_Group=4.00]	0	.	.	.	.	.

The covariance-adjusted Brigance posttest means can be seen more clearly by usage quartile in Table 8. Participants in Quartile 1(less than 62 hours of instruction) have a mean posttest score of 182.29 on the Brigance which is significantly lower than Quartile 4 participants (84 hours or more of instruction) who have a significantly higher mean posttest score --by 19 points -- of 201.47 on the Brigance.

Table 8
Parameter Estimates: Brigance Total Posttest Scores by Usage Quartile

Estimates

Dependent Variable: Brigance Total Posttest

Hrs Instruction Quartile	Brigance Posttest Mean	Std. Error	95% Confidence Interval	
			Lower Bound	Upper Bound
Quartile 1	182.293[a]	4.554	173.278	191.308
Quartile 2	194.738[a]	4.214	186.396	203.080
Quartile 3	198.317[a]	4.329	189.748	206.866
Quartile 4	201.467[a]	4.402	192.754	210.181

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

Total Pretest = 130.82 and child has access to a computer at home = .96.

The covariance adjusted mean posttest literacy achievement difference between usage quartiles 1 and 2 is a little more than 12 points and is not statistically significant whereas the difference between quartiles 1 and 3 is 16 points and approaches statistical significance (p=.076). *Using quartile one as a baseline, a statistically significant usage effect is not observed until participants reach the 4th quartile of usage -- approximately 84 hours of instruction.*
The covariance adjusted Brigance posttest means presented in Table 8 suggest a positive linear trend. This positive linear trend in the development of early literacy skills associated with

UPSTART usage quartile is apparent in Figure 5 below. The line graph displays adjusted Brigance means at each quartile of instruction time, using Brigance pretest scores and whether the child had prior access to a computer at home as covariates.

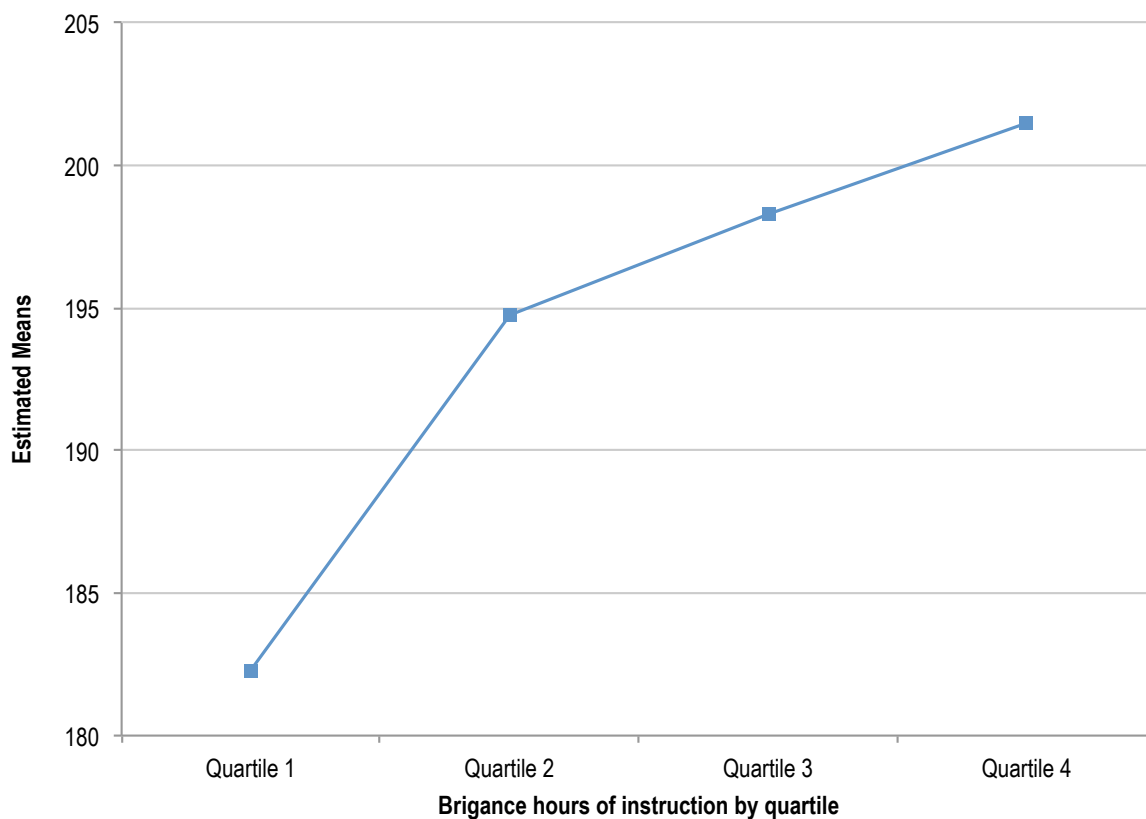*Figure 5. Mean Brigance Total Posttest Scores by UPSTART Usage Quartile*

Covariates appearing in the model are evaluated at the following values: Total Brigance Pretest = 130.8217, Child has access to a computer at home = .96

Only the difference between quartile 1 and quartile 4 is statistically significant. This result is partially explained by the observed statistical power available for analyzing the four levels (quartiles) of usage, which is adequate for quartile 1 (power = .82) but not for quartiles 2 and 3 (power = .20 and .08 respectively). It is clear that there is a big difference in the literacy effects associated with the amount of instruction at quartile 1 vs. quartile 4. It is also apparent that there is not much difference in the literacy effects between usage quartiles 2 and 3 (i.e., a difference of 4 points on the Total Brigance). However, it is not clear whether there is a real difference in literacy effects between usage quartile 1 and usage quartiles 2 and 3 because of the possibility that this lack of significance can be explained by low statistical power. Yet the linear trend observed in Figure 5 coupled with the fact that usage and literacy achievement are positively correlated, suggests that the more UPSTART participants use the curriculum the more they will tend to develop their early phonics and language- related cognitive skills.

UPSTART Usage Effects Measured by the Bader. The results are somewhat different for the Bader at the beginning of kindergarten – see Table 9 -- which shows that increases in literacy skills measured by the Bader (phonological awareness) do not increase significantly with usage of the UPSTART curriculum (p=.26). The Bader ANCOVA used pretest Bader scores, prior access to a computer at home, and child ethnicity (if Hispanic) as covariates to adjust Bader posttest scores for the influence of variables other than hours of instruction.[9] Sample size for the Bader usage effects analysis was 112 and the observed statistical power was low (power = .35).

Table 9
Tests of Between Subjects Effects: Bader Total Posttest - Beginning K-

Source	Type III Sum of Squares	df	Mean Square	F	Sig.	Partial Eta Squared
Corrected Model	1636.869	6	272.811	11.203	.000	.390
Intercept	305.346	1	305.346	12.540	.001	.107
Bader_Pre	752.565	1	752.565	30.905	.000	.227
Computer_Home	226.065	1	226.065	9.284	.003	.081
Hispanic	163.190	1	163.190	6.702	.011	.060
Usage Group	98.487	3	32.829	1.348	.263	.037
Error	2556.810	105	24.351			
Total	19950.000	112				
Corrected Total	4766.216	111				

Adjusted R Squared = .36;   Computed using alpha = .05

The parameter estimates in Table 10 indicate that none of the usage group quartiles are significantly different from each other. The Table 10 results also suggest that the ANCOVA's lack of usage group effects is related to low statistical power in the Bader sample.

---

[9] Exploratory analyses that controlled for hours of instruction indicated that the following variables were significantly related to Total Bader posttest scores: Bader pretest, prior access to a computer at home, child ethnicity (if Hispanic or Caucasian), whether the parent's primary language was English, and household size. The final ANCOVA included three covariates: Bader pretest, prior access to a computer at home, and the child's ethnicity (Hispanic vs. otherwise).

Table 10

Parameter Estimates: Bader Total Posttest Score - Beginning K-

Parameter	B	Std. Error	t	Sig.	Partial Eta Squared	Observed Power
Intercept	9.611	2.782	3.454	.001	.102	.928
Bader_Pre	.447	.080	5.559	.000	.227	1.000
Computer_Home	7.423	2.436	3.047	.003	.081	.855
Hispanic	-9.808	3.789	-2.589	.011	.060	.727
[Usage_Group=1.00]	-2.247	1.370	-1.640	.104	.025	.369
[Usage_Group=2.00]	-0.057	1.345	-0.042	.966	.000	.050
[Usage_Group=3.00]	0.014	1.340	0.010	.992	.000	.050
[Usage_Group=4.00]	0	.	.	.		.

The lack of statistically significant differences across usage quartile for the Bader is more evident from the results shown in Table 11 for adjusted Bader Total Posttest means by usage quartile. *Taken at face value, these results suggest that phonological awareness did not improve with increasing UPSTART usage among the C3 participants.*

Table 11

Parameter Estimates: Bader Total Posttest Score by Usage Quartile

Estimates

Dependent Variable: Bader Total Posttest

Hrs Instruction Quartile	Mean	Std. Error	95% Confidence Interval	
			Lower Bound	Upper Bound
Quartile 1	18.272	.955	16.377	20.166
Quartile 2	20.462	.946	18.586	22.238
Quartile 3	20.533	.935	18.679	22.387
Quartile 4	20.519	.953	18.620	22.409

a. Covariates appearing in the model are evaluated at the following values: Bader Total
Pretest = 8.9196, Child has prior acccess.to computer at home = .96, Hispanic = .0179.

Again, the possibility exists that statistical power is an explanatory factor underlying the lack of statistical significance for differences in UPSTART usage associated with Bader literacy outcomes. The other explanatory factor is that the Total Bader scale range is relatively narrow (i.e., scores from 1 to 26) and the scale metric may be too coarse to discriminate usage effects.

UPSTART Outcomes

In this section, the main research questions of interest to the C3 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 performance of the treatment and control group children was essentially equivalent on the overall Brigance at the pretest. However, the control group children scored significantly higher on the pretests measuring a preschool child's ability to recite the alphabet and to discriminate word sounds (see the T-Test results in Table 12 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 (82% vs. 47%). 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 12
Brigance Pretest Analysis of Treatment-Control Group Differences

Brigance Pretest	Group	N	Mean	SD	t	Significance
	Treatment	129	25.48	1.26		
Expressive Objects	Control	130	25.53	1.46	-0.295	NS
	Treatment	129	26.90	0.37		
Receptive Objects	Control	130	26.86	0.43	0.757	NS
	Treatment	129	9.16	1.19		
Expressive Grammar	Control	130	9.08	1.33	0.498	NS
	Treatment	129	14.10	4.50		
Visual Discrimination	Control	130	14.10	4.54	0.001	NS
	Treatment	129	10.66	9.21		
Recites Alphabet	Control	130	13.98	9.98	-2.786	**
Lowercase Letter	Treatment	129	27.26	20.07		
Knowledge	Control	130	25.80	20.28	0.584	NS
Sounds of Lowercase	Treatment	129	9.02	9.44		
Letters	Control	130	8.89	9.92	0.109	NS
Auditory	Treatment	129	4.67	3.61		
Discrimination	Control	130	6.08	3.44	-3.218	**
	Treatment	129	1.72	2.21		
Survival Sight Words	Control	130	1.76	2.22	-0.148	NS
Basic Preprimer	Treatment	129	1.85	5.25		
Vocabulary	Control	130	1.45	4.48	0.671	NS
	Treatment	129	130.82	42.03		
Total Brigance	Control	130	133.53	43.66	-0.509	NS

**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), whether or not the parent's primary language was English, the parent's educational attainment, and whether or not the parent was married. These four 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 and the parent's marital status (being married or not) were retained as statistical control variables in the final regression analysis. Marital status favored the treatment group over the control group (95% to 83%). 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 *all but one of the Brigance tests* (i.e., the Receptive Objects subtest). These results are shown below in Table 13.

Table 13
Brigance Posttest Analysis of Treatment-Control Group Differences

Brigance Posttest	Group	N	Mean	SD	t	Significance
Expressive Objects	Treatment	129	26.35	.79		
	Control	130	26.01	1.02	3.006	**
Receptive Objects	Treatment	129	26.99	.09		
	Control	130	26.98	.15	1.000	NS
Expressive Grammar	Treatment	129	10.20	1.38		
	Control	130	9.62	1.33	3.475	**
Visual Discrimination	Treatment	129	18.44	2.05		
	Control	130	16.61	2.82	5.985	**
Recites Alphabet	Treatment	129	21.19	8.04		
	Control	130	17.29	9.16	3.635	**
Lowercase Letter Knowledge	Treatment	129	44.60	12.33		
	Control	130	38.12	17.19	3.488	**
Sounds of Lowercase Letters	Treatment	129	20.12	7.99		
	Control	130	14.71	9.77	4.886	**
Auditory Discrimination	Treatment	129	8.22	2.58		
	Control	130	7.06	3.04	3.295	**
Survival Sight Words	Treatment	129	4.88	4.39		
	Control	130	3.29	3.06	3.369	**
Basic Preprimer Vocabulary	Treatment	129	13.25	9.19		
	Control	130	6.37	8.16	6.369	**
Total Brigance	Treatment	129	194.23	36.81		
	Control	130	166.05	41.32	6.369	**

**p≤.01

Using the data from Table 13, effect sizes[10] 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 Figure 6 below and show the magnitude of the average performance difference in standard deviation units between the C3 treatment group and the control group on each of the Brigance assessments administered in the C3 evaluation. Because the effect sizes are all positive, they favor the UPSTART treatment group. Statistically significant ES estimates shown in Figure 6 range from .33 (Expressive Objects) to .84 (Vocabulary), indicating a mixture of small, medium, and large effects. The Brigance ES estimates can be categorized as follows:[11]

- Small Effects (ES = .20-.50): Expressive Objects, Expressive Grammar, Alphabet, Letter Knowledge, and Auditory Discrimination.
- Medium Effects (ES = .50-.80): Visual Discrimination, Letter Sounds, Sight Words, and Total Brigance.
- Large Effects (ES >.80): Vocabulary

---

[10] 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.

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

The total Brigance is estimated to have an effect size of .68, indicating that the overall difference between the treatment and control group at posttest was approximately two-thirds of a standard deviation, considered a medium effect by Cohen. On average, treatment group children scored 28 points higher on the total Brigance at posttest compared to control group children.

Figure 6: UPSTART's Impact on the Brigance in Effect Size Units
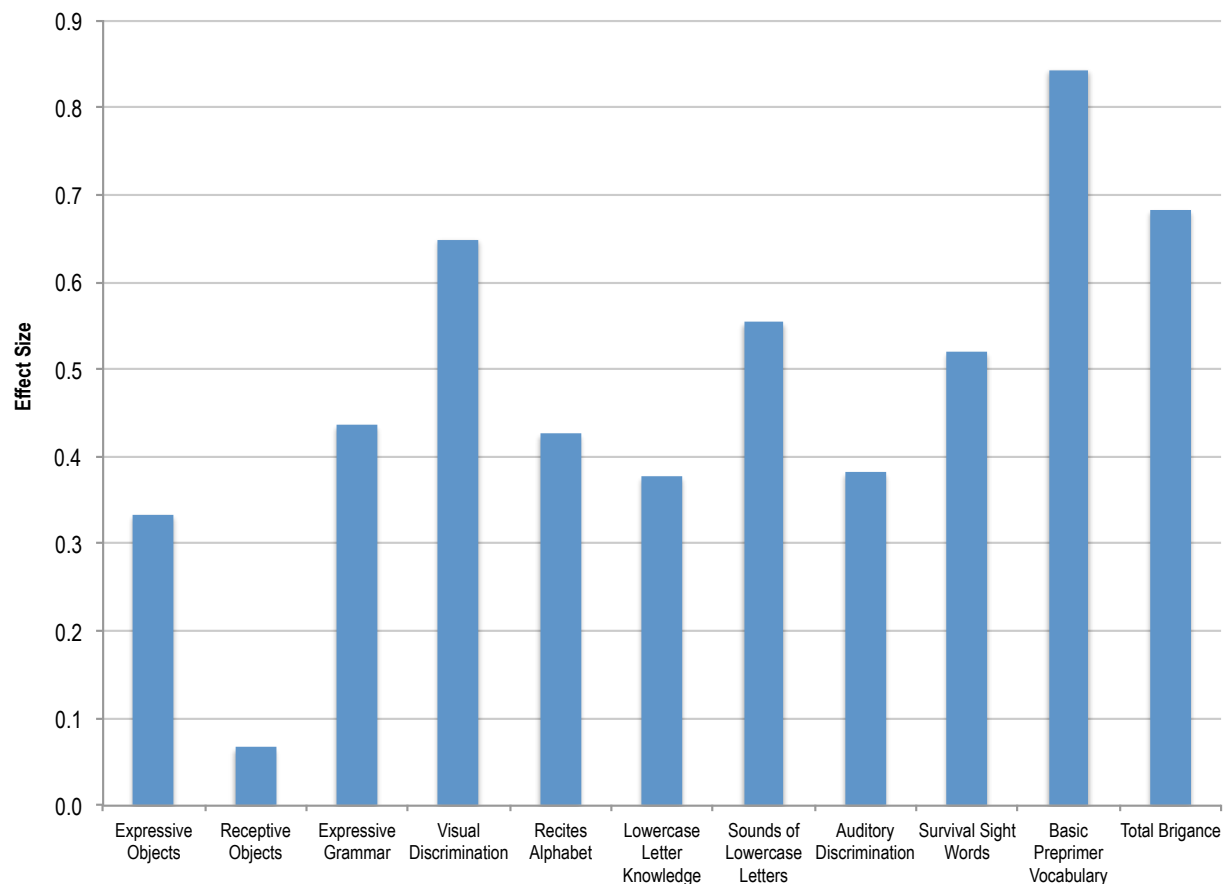
Brigance Posttest Regression Results. Adjusting for the initial differences between the treatment and control groups through the use of multiple regression analysis, it was found that the treatment group children outscored the control group children on the *overall Brigance* posttest by 27.98 points. This is almost identical to the t-test results reviewed above (i.e., a difference of 28 points on the average). The final Brigance regression model[12] is shown in Tables 14 and 15.

---

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

Table 14

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

Model	Sum of Squares	df	Mean Square	F	Sig.
Regression	297431.370	3	99143.790	171.154	.000
Residual	147712.761	255	579.266		
Total	445144.131	258			

Table 15

OLS Regression Coefficients for Total Brigance Posttest – Beginning of Kindergarten

Model	Unstandardized Coefficients		Standardized Coefficients			Partial
	B	Std. Error	Beta	t	Sig.	Correlation
(Constant)	59.779	6.100		9.800	.000	
Total Pretest	.690	.036	.711	19.391	.000	.77
Married	16.931	4.994	.127	3.391	.001	.21
Study Group	*27.979*	3.058	.337	9.149	.000	.50

Adjusted R Square = .66;   Partial R Square = .25

The observed effect size for the Brigance treatment group (partial R square for Study Group = .25) is in the medium effects size range (see Cohen, 1988; Chapter 9). The results presented in Table 13 and Figure 6 indicate that UPSTART's impact as measured by the Brigance was across the board, with the lone exception of the Receptive Objects subtest.   On average, UPSTART achieved medium size literacy impacts as measured by the Brigance.

Bader Pretest Results. The pretest performance of the treatment and control group children was essentially equivalent on the Total Bader but two of the Bader subtests favored the control group. Table 16 shows that the control group children scored slightly higher on the Bader Rhyme Recognition pretest and the Bader Phoneme Blending pretest.[13]

---

[13] These differences were not statistically significant at the .01 level.

Table 16
Bader Pretest Analysis of Treatment-Control Group Differences

Bader Pretest	Group	N	Mean	SD	t	Significance
Rhyme Recognition	Treatment	112	5.83	3.23		
	Control	120	6.62	2.89	-1.96	*
Phoneme Blending	Treatment	112	1.97	2.86		
	Control	120	2.79	3.25	-2.04	*
Phoneme Segmenting	Treatment	112	1.12	2.41		
	Control	120	1.00	2.19	0.38	NS
Total Bader	Treatment	112	8.92	6.01		
	Control	120	10.41	6.14	-1.86	NS

* $p \leq .05$

While there were a number of demographic differences between the treatment and control children in the Bader analysis sample (see Appendix D), only the parent's educational attainment and the child's ethnicity (if Hispanic) were significantly related to Bader posttest outcomes (see Appendix C). These initial differences both favored the treatment group:

- Only 2% of the treatment group children were Hispanics vs. 11% for the control group.
- The treatment group parents were more highly educated than the control parents.

These two demographics and 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. Posttest results showed a statistically significant treatment group effect on each of the Bader subtests as well as the Total Bader; see Table 17 below. The mean observed (unadjusted) difference between the treatment and control group on the Total Bader posttest was 5.95 points.

Table 17
Bader Posttest Analysis of Treatment-Control Group Differences

Bader Posttest	Group	N	Mean	SD	t	Significance
Rhyme Recognition	Treatment	112	8.71	2.14		
	Control	120	7.50	2.83	3.71	**
Phoneme Blending	Treatment	112	6.14	2.61		
	Control	120	3.90	3.19	5.88	**
Phoneme Segmenting	Treatment	112	5.09	3.10		
	Control	120	2.59	2.95	6.29	**
Total Bader	Treatment	112	19.95	6.15		
	Control	120	13.99	7.01	6.89	**

** $p < .01$

Effect size estimates for the Bader posttest results are presented below in Figure 7. Using Cohen's standardized difference score conventions, the Bader effect size estimates can be categorized as follows:

- Small Effect (ES = .20-.50): Rhyme Recognition (ES=.43)
- Medium Effects (ES = .50-.80): Phoneme Blending (ES=.70)
- Large Effects (ES >.80): Phoneme Segmentation and the Total Bader (ES=.85)

Figure 7: UPSTART's Impact on the Bader in Effect Size Units
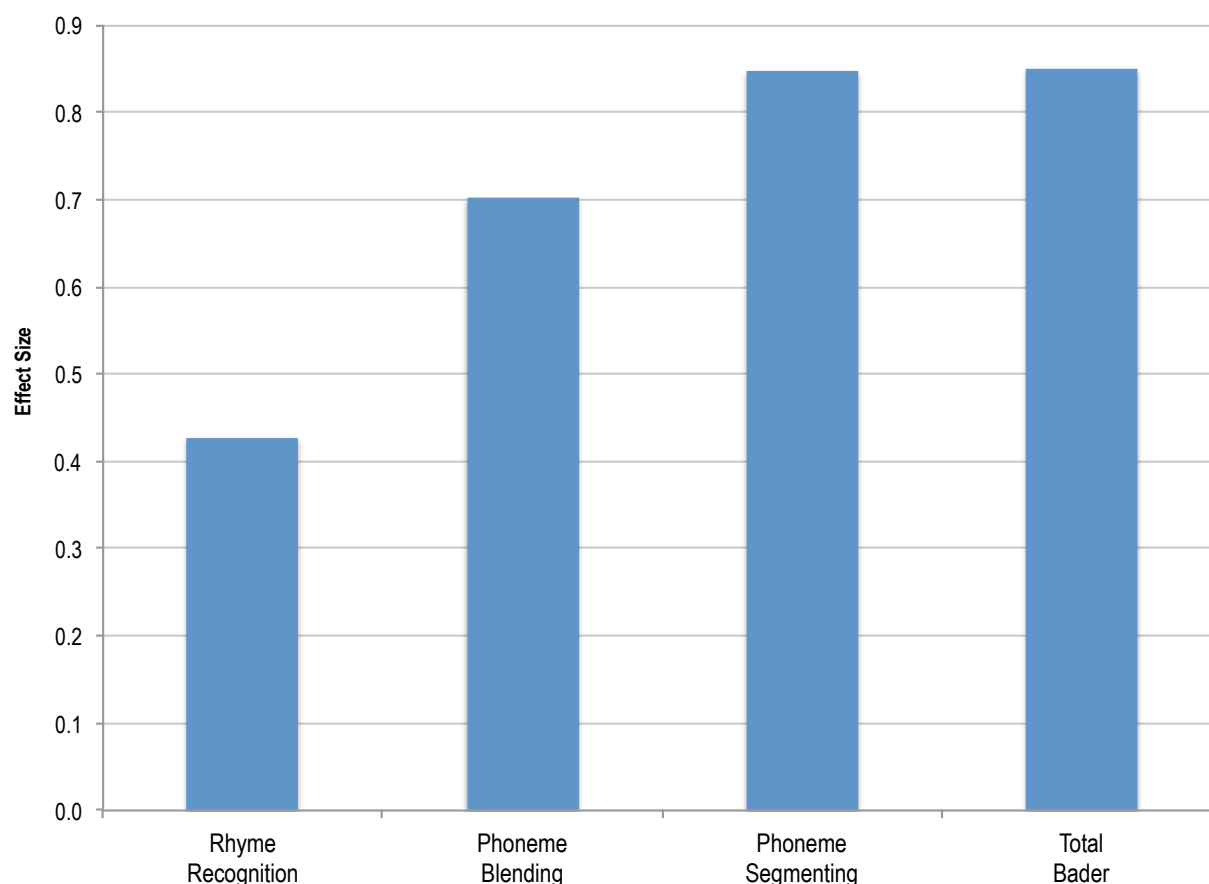
Bader Posttest Regression Results. Adjusting for the initial between group differences using multiple regression analysis, it was found that the treatment group outscored the control group on the Total Bader by 6.87 points on the average. The final Bader regression model[14] is shown below in Tables 12 and 13.

---

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

Table 12

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

Model	Sum of Squares	df	Mean Square	F	Sig.
Regression	5234.324	2	2617.162	87.436	.000
Residual	6854.534	229	29.932		
Total	12088.858	231			

Predictors: Bader Pretest and Study Group

Table 13

OLS Regression Coefficients for Total Bader Posttest – Beginning of Kindergarten

	Unstandardized Coefficients		Standardized Coefficients			Partial
	B	Std. Error	Beta	t	Sig.	Correlation
(Constant)	7.625	.794		9.599		
Total Pretest	.612	.059	.517	10.307	.000	.563
Study Group	*6.865*	.724	.475	9.480	.000	.531

Adjusted R Square = .43;   Partial R Square = .28

The partial R square for Study Group in the Bader regression analysis is .28, which suggests that UPSTART had a relatively large impact on the C3 participants as measured by the overall Bader (see Cohen, 1988; Chapter 9). That is, the adjusted treatment effect of 6.87 points on the Total Bader should be considered a large effect. It should also be noted that the statistically adjusted overall treatment effect of 6.87 is larger than the raw difference score reported in the discussion of the t-test findings (i.e., 5.95 points). Finally, it should be noted that the C3 UPSTART impact as measured by the Bader was substantial, not only in size but in breadth, as the gains in phonological awareness were observed for UPSTART participants across the board for all three subtests as well as for the Total Bader.

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=129) on the Total Brigance and on all ten subtests. Similar results were observed for the matched Brigance control group (N=130).

Growth rates were significantly different at the 99% CI between the treatment and control group for the *overall Brigance* and five subtests: *Vocabulary, Sight Words, Auditory Discrimination, Letter Sounds, and Recites Alphabet.* All of these differences in growth rates favored the UPSTART treatment group. These results are shown in Table 14 below.

Table 14
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	.476	.196 – .756	.868	.571 – 1.165	NS
Receptive Objects	.115	.018 – .212	.093	.014 - .171	NS
Expressive Grammar	.538	.157 - .919	1.046	.713 – 1.379	NS
Visual Discrimination	2.507	1.555 – 3.459	4.341	3.407 – 5.274	NS
Recites Alphabet	3.307	1.129 -5.486	10.572	8.287 – 12.776	**
Letter Knowledge	12.315	8.720 -15.910	17.333	13.441 – 21.228	NS
Letter Sounds	5.815	4.021 – 7.609	11.100	9.137 – 13.063	**
Auditory Discrimination	.984	.008 – 1.960	3.550	2.664 – 4.435	**
Survival Sight Words	1.530	.951 -2.110	3.155	2.355 – 3.954	**
Basic Vocabulary	4.923	3.369 -6.476	11.395	9.406 – 13.384	**
Total Brigance	32.515	26.661 - 38.369	63.410	56.680 - 70.141	**

Figure 8 uses bar charts to compare the growth rates of the treatment and control group as measured by the Total Brigance and each of its subtests from pretest to posttest for the matched samples.

Figure 8. Growth Rate Comparisons on the Brigance
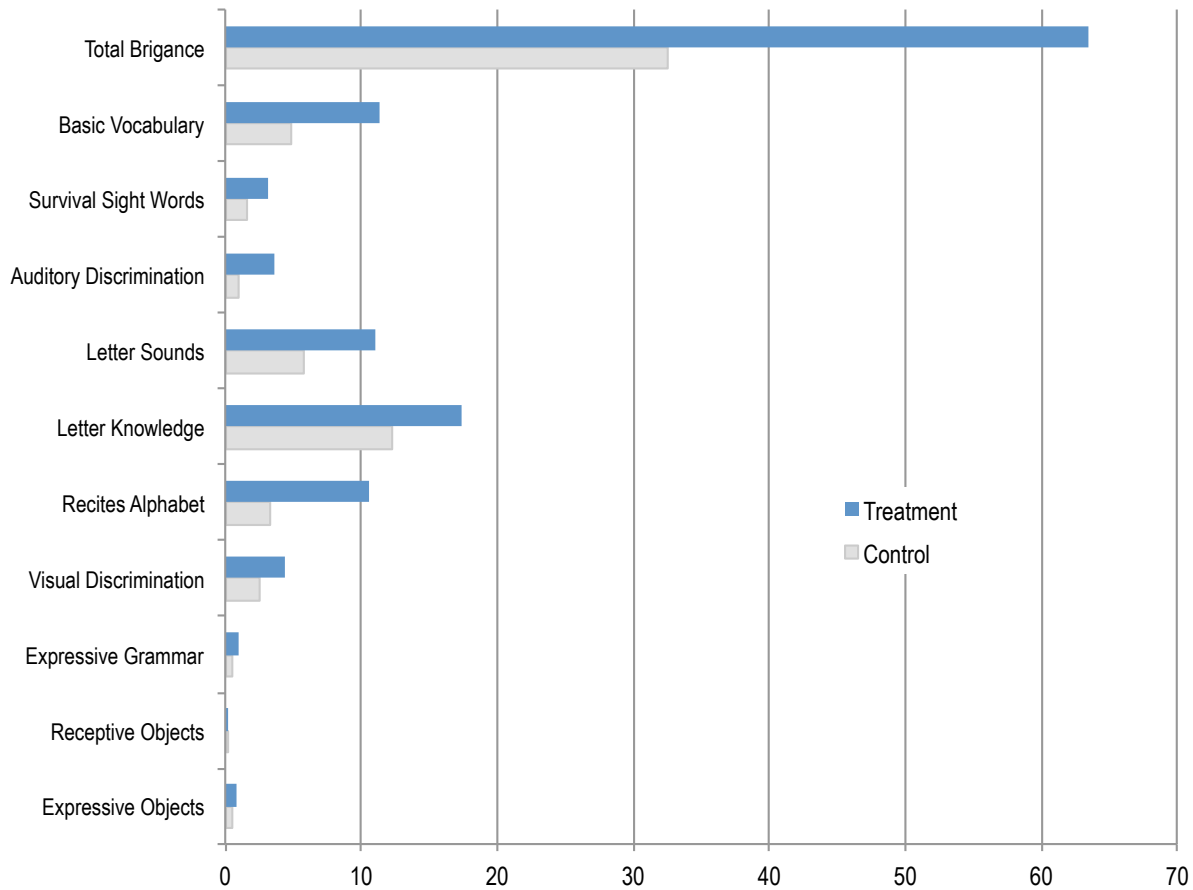
Bader Growth Score Results. There was statistically significant growth from pretest to posttest for the matched Bader treatment group sample (N=112) and for the matched Bader control group sample (N=120) on the Total Bader and all of the Bader subtests. Additionally, the treatment group showed significantly stronger growth rates (statistically significant at the 99% CI) relative to the control group on the Total *Bader* and all of the Bader subtests as well. These results are shown in Table 15.

Table 15
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	.883	.081 – 1.684	2.883	2.048 – 3.719	**
Phoneme Blending	1.108	.452 – 1.763	4.169	3.406 - 4.932	**
Phoneme Segmenting	1.591	.945 - 2.238	3.973	3.168 – 4.777	**
Total Bader	3.583	2.239 - 4.927	11.026	9.472 - 12.581	**

Figure 6 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.
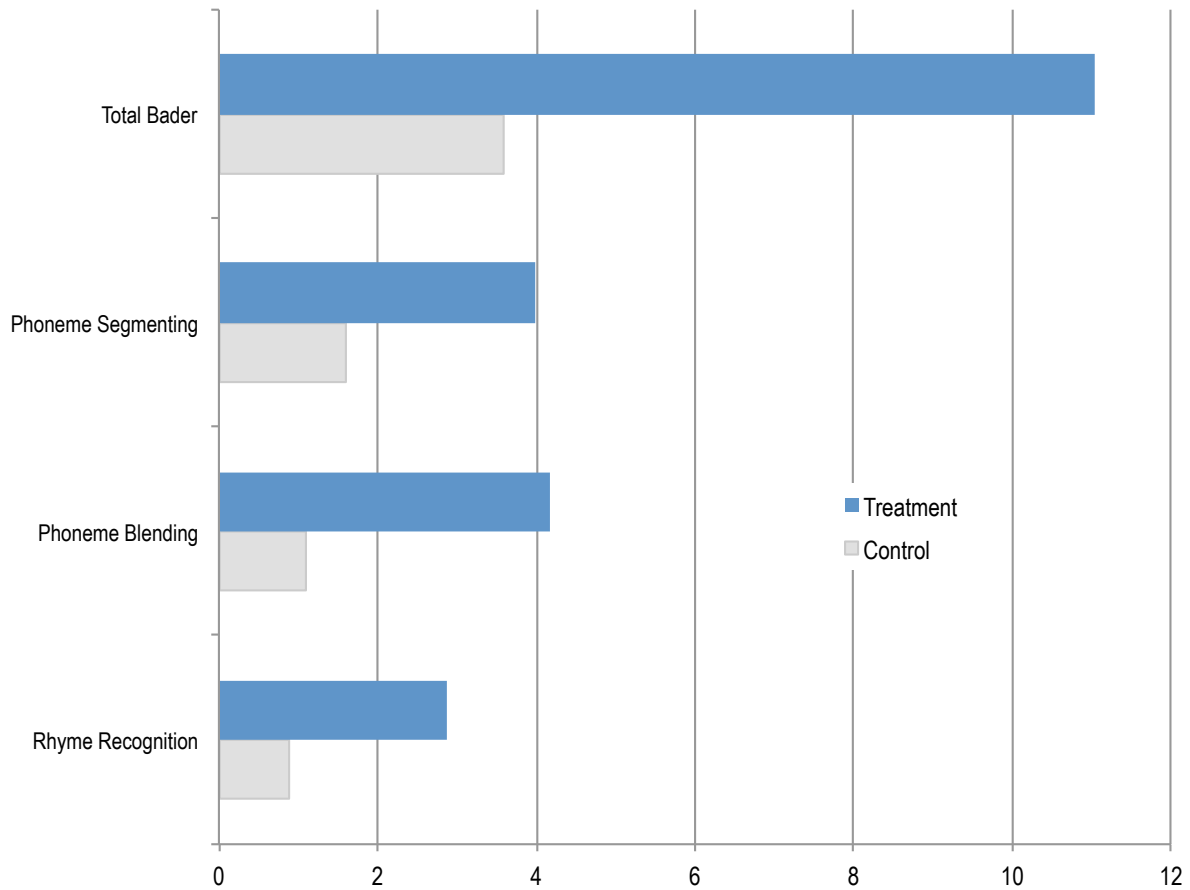
Figure 9. Growth Rate Comparisons on the Bader

Summary and Conclusions

This final section of the Year Three evaluation report summarizes:
- The data that were collected and analyzed;
- The analysis methods employed;
- C3 findings and trends in UPSTART implementation; and
- C3 findings and trends in UPSTART's impact on the development of early literacy in young children as measured in preschool and at the beginning of kindergarten.

Data Collection

305 four year-old children were recruited for the C3 study; 151 treatment group children who had enrolled in UPSTART for Year 3of the program and 154 control group children who had not enrolled in the UPSTART program. The children's parents were administered 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 approximately one year after their pretest. Complete pretest and posttest data were obtained and analyzed for 259 children on the Brigance (129 treatment group children and 130 control children) and for 232 children on the Bader (112 treatment children and 120 control children).

The Brigance sample experienced a 15% attrition rate whereas the Bader sample experienced a 24% attrition rate. Electing not to return for posttesting was equally evident among the treatment and control group families. However, the higher attrition rate on the Bader was also caused by test floor effects: children not being able to score on the test. Since the testing sequence was Brigance followed by Bader, it could be that testing fatigue contributed to the poorer performance of the children on the Bader. It could also be that if the testing experience with the Bader was rather negative for the child, then that may also have contributed to families deciding not to return for additional testing.

Data Analysis

To determine whether UPSTART children have better literacy skills at kindergarten compared to control group children, group equivalence on the pretests was examined using independent sample t-tests. Relationships between the demographics and the posttest scores were then examined using correlation analyses.[15] 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

---

[15] 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.

which the pretest was entered first, followed by a set of demographic covariates, followed by the treatment-control group.

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 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 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 and selected demographic characteristics (e.g., access to a computer at home). The ANCOVA analyses estimated the effects of usage at quartiles one through three compared with usage at the fourth quartile, controlling for covariates entered in the final model. Importantly, it was determined that the ANCOVA analyses were statistically underpowered with the size of the final analysis samples obtained from the Bader (N=112) and the Brigance (N=129).

UPSTART Implementation Results

The Waterford Institute provided documentation for a third-year UPSTART enrollment of 1, 168 children. A majority (64%) of the 1,168 preschool children that enrolled in the third year of UPSTART were from low income families, according to data provided by the Waterford Institute. Slightly more boys (53%) were enrolled than girls (47%). In terms of ethnicity, the vast majority of the enrollment was Caucasian (79%), 16% were Hispanic, 3% were of Asian descent, 1% were Black, and 1% were Native American. Ethnicity was unknown for 1% of the third year enrollment.

As in previous years, most of the C3 participants (71%) received a computer drive with the UPSTART curriculum loaded on it. Approximately 12% of the third year participants received a computer loan and free internet access to help them access the UPSTART curriculum. Another 7% of the C3 participants were loaned a personal computer to use at home while participating in UPSTART. The remaining 10% of the third 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 C3 preschool analysis samples had a mean of approximately 75 hours of UPSTART curriculum usage over the third year of the project. This compares with an average of approximately 70 hours of instruction for program "graduates" and approximately 67 hours of instruction for all students enrolled in UPSTART in the third year of the program.

- Length of participation in the UPSTART curriculum was significantly and positively correlated with literacy skills at the beginning of kindergarten as measured by the Brigance (r=.40) and the Bader (r=.20).

- Early language and cognitive literacy skills measured by the Brigance generally improved with increasing levels of UPSTART curriculum usage. This was not the case with the phonological awareness skills measured by the Bader, which did not improve significantly with increasing levels of UPSTART curriculum usage. Both of these findings replicate the C2 usage results.

- The UPSTART graduation rate has continued to rise each year: from 59% in Year 1, to 76% in Year 2, to 94% in Year 3. This indicates that UPSTART is making very good progress in achieving the curriculum usage goals set for program implementation.

- UPSTART graduate status in the third year of the program was significantly correlated with hours of instruction but not with literacy outcomes measured by the Brigance or the Bader at the beginning of kindergarten. Both of these findings replicate the C2 usage results.

UPSTART Impact Results

- UPSTART participation generally continues to show moderately strong impacts on improving the phonics skills of young preschool children. With one exception, these effects were observed across all domains assessed by the Brigance with the C3 group at the beginning of Kindergarten. The largest impact observed in UPSTART's third year of operation was in the development of vocabulary. Overall, treatment group children scored an average of 28 points higher on the Brigance posttest compared to control group children.

- Overall, UPSTART achieved large impacts on improving the phonological awareness skills of participants in Year 3 of the program. Treatment group effects were observed across all three of the phonological awareness domains measured by the Bader: the ability to recognize rhymes; the ability to blend phonemes; and the ability to segment phonemes.

- UPSTART participants showed significantly stronger growth rates compared to control group children on the *Total Brigance* and five of the ten subtests, including: *Vocabulary, Sight Words, Auditory Discrimination, Letter Sounds and Reciting the Alphabet*.

- UPSTART participants showed significantly stronger growth rates compared to control children on the *Total Bader* and all of the Bader subtests, including *Rhyme Recognition, Phoneme Blending* and *Phoneme Segmenting*.

Recommendations

Programmatically, the one recommendation that should help the participants improve their phonics and beginning reader skills is to offer incentives for reaching usage milestones that appear correlated with literacy achievement. These might include rewarding the attainment of three successive levels of curriculum usage, awarded at the following milestones:

- 65 hours
- 75 hours
- 85 hours

Methodologically, the evaluators should continue efforts to increase sample size. Specifically, efforts should focus on expanding the treatment group so that it is more representative of the program population and so that greater statistical power is available for conducting the usage analysis. Determining the desired sample size, or options in sample size improvement, should be approached by having a statistical power analysis study conducted. The other way in which the effective sample size can be increased is by reducing attrition. The key to reducing sample attrition might be to provide sufficient breaks between testing sessions so that testing fatigue on the part of the assessors and the children and families being assessed is minimized.

Appendix A

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?
    $\square_1$ Yes $\square_2$ No

    1a. If yes, did your 4-year-old also use the program?
       $\square_1$ Yes $\square_2$ No

Child Information

2. What is your child's birthday? _____

3. What is your child's gender?
    $\square_1$ Male $\square_2$ Female

4. What year will your child be entering Kindergarten?
    $\square_1$ 2011 $\square_2$ 2012

5. What is your child's ethnicity?
    $\square_1$ Hispanic $\square_2$ Native American/Alaskan Native $\square_3$ Asian/Pacific Islander
    $\square_4$ Caucasian $\square_5$ African American $\square_6$ Other:_____

6. What is your child's primary language?
    $\square_1$ English $\square_2$ Spanish $\square_3$ Portuguese $\square_4$ Chinese
    $\square_5$ German $\square_6$ Japanese $\square_7$ Other:_____

7. Is your child currently attending a daycare/preschool?
    $\square_1$ Yes $\square_2$ No

    7a.**If yes**, approximately how many hours a week does your child attend a daycare/preschool?
       $\square_1$ less than 10 hours $\square_2$ 10-19 hours $\square_4$ 20-24 hours
       $\square_5$ 25-29 hours $\square_6$ 30-34 hours $\square_7$ 35 or more hours

8. Does your child have access to a computer in your house?
    $\square_1$ Yes $\square_2$ No

9. Does your child use a computer in her/his day care or preschool?
    $\square_1$ Yes $\square_2$ No $\square_3$ Not Applicable (not in day care or preschool)

10. How comfortable is your child using a computer?
    $\square_1$ Very comfortable $\square_2$ Somewhat comfortable $\square_3$ Somewhat uncomfortable
    $\square_4$ Not comfortable $\square_5$ 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
Brigance Sample: Treatment – Control Differences on Demographics

Covariate	Group	N	Mean	SD	t	Significance
Child is male	Treatment	129	.57	.50		
	Control	130	.44	.50	2.06	*
Child is Caucasian	Treatment	128	80	.30		
	Control	130	.90	.36	1.26	NS
Child is Hispanic	Treatment	129	.03	.17		
	Control	130	.12	.32	-2.63	**
Child's primary language is English	Treatment	129	.99	.09		
	Control	130	.95	.23	2.16	*
Child attends preschool <10 hours per week	Treatment	129	.35	.48		
	Control	130	.51	.50	-2.61	**
Child currently attends preschool or daycare	Treatment	129	.47	.50		
	Control	129	.82	.38	-6.42	**
Child has access to a computer at home	Treatment	129	.96	.19		
	Control	129	.91	.28	1.55	NS
Child uses PC in preschool or daycare	Treatment	127	2.36	.64		
	Control	128	1.91	.68	5.44	**
Child comfort level with computers	Treatment	128	1.85	.83		
	Control	127	1.98	.99	-1.16	NS
Caregiver is mother	Treatment	129	.92	.27		
	Control	130	.92	.27	-.018	NS
Parent is Caucasian	Treatment	129	.95	.02		
	Control	130	.89	.03	1.58	NS
Parent is Hispanic	Treatment	129	.02	.15		
	Control	130	.08	.28	-2.20	*
Parent's primary language is English	Treatment	129	.99	.00		
	Control	130	.95	.23	2.37	*
Parent Educational Attainment (recoded)[16]	Treatment	129	2.83	.79		
	Control	129	2.38	.71	4.80	**
Parent is married	Treatment	129	.95	.21		
	Control	130	.83	.38	3.23	**
Parent employment status	Treatment	129	2.46	.71		
	Control	129	2.40	.73	.60	NS
Spouse employment status	Treatment	128	1.36	.81		
	Control	129	1.50	1.02	-1.26	NS
Household size	Treatment	129	4.86	.94		
	Control	129	4.74	1.07	.98	NS
Household annual income category	Treatment	128	4.01	1.28		
	Control	128	3.85	1.27	.98	NS

$**p<.01$
$*p<.05$

---

[16] 1 = High School; 2= Some College; 3 = College Graduate; 4 = Graduate Degree

Appendix C
Correlations with Total Posttest Scores

Variable	Brigance	Bader
Study Group[17]	.34**	.41**
Pretest	.72**	.46**
Child is Male	-.07	-.06
Child is Caucasian	.26**	.20**
Child is Hispanic	-.29**	-.24**
Hours per week attends daycare/preschool (S)	-.17	-.07
Child attends daycare/preschool	-.04	-.04
Child has computer access at home	.18**	.16**
Child uses computer at daycare/preschool	-.08	-.13
Child's computer comfort	-.16**	-.05
Parent employment status (S)	.03	.02
Spouse employment status (S)	-.16**	-.05
Household size	-.05	-.08
Household income	.25**	.17**
Child's primary language is English	.13	.17**
Attends daycare/preschool <10 hours per week	.11	.00
Caregiver is mother	.02	.00
Parent is Caucasian	.16	.12
Parent is Hispanic	-.18**	-.17**
Parent's primary language is English	.20**	.20**
Parent Educational Attainment	.30**	.28**
Parent is married	.31**	.15
	**p≤.01	**p≤.01

Note: Variables with (S) indicates correlation is Spearman's rho; otherwise, correlations are Pearson.

---

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

Bader Sample: Treatment – Control Differences on Demographics

Covariate	Group	N	Mean	SD	t	Significance
Child is Male	Treatment	112	.56	.50		
	Control	120	.42	.50	-5.89	**
Child is Caucasian	Treatment	111	.91	.29		
	Control	120	.86	.35	1.22	NS
Child is Hispanic	Treatment	112	.02	.13		
	Control	120	.11	.31	-2.91	**
Child's primary language is English	Treatment	112	.99	.09		
	Control	120	.99	.20	1.61	NS
Attend preschool <10 hours per week	Treatment	112	.34	.48		
	Control	120	.50	.53	-2.50	**
Currently attending preschool or daycare	Treatment	112	.46	.50		
	Control	119	.81	.40	-5.89	**
Child has access to a computer at home	Treatment	112	.96	.21		
	Control	119	.92	.28	1.22	NS
Child uses PC in preschool or daycare	Treatment	110	2.38	.65		
	Control	119	1.94	.68	5.01	**
Child comfort level with computers	Treatment	111	1.86	.87		
	Control	118	1.96	.99	0.82	NS
Caregiver is mother	Treatment	112	.92	.27		
	Control	120	.92	.28	0.82	NS
Parent is Caucasian	Treatment	112	.95	.23		
	Control	120	.89	.31	1.54	NS
Parent is Hispanic	Treatment	112	.02	.13		
	Control	120	.08	.28	-2.32	NS
Parent's primary language is English	Treatment	112	.98	.13		
	Control	120	.93	.26	2.09	NS
Parent Educational Attainment (recoded)[18]	Treatment	112	2.79	.80		
	Control	119	2.40	.68	4.02	**
Parent is married	Treatment	112	.96	.21		
	Control	120	.84	.37	2.93	**
Parent employment status	Treatment	112	2.45	.72		
	Control	119	2.38	.74	0.71	NS
Spouse employment status	Treatment	111	1.40	.86		
	Control	119	1.48	.99	-0.67	NS
Household size	Treatment	112	4.86	.94		
	Control	119	4.69	1.07	1.27	NS
Household annual income category	Treatment	111	3.92	1.27		
	Control	118	3.87	1.25	0.28	NS

$**p<.01$

---

[18] 1 = High School; 2= Some College; 3 = College Graduate; 4 = Graduate Degree