Utah UPSTART Program Evaluation
Kindergarten Outcomes:
Program Impacts on Reading Proficiency

Cohort 2 Results

Technical Report

March 2012

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 educational technology in a home-based approach to develop the school readiness skills of preschool children. A majority (60%) of the 1,018 preschool children enrolled in the second year of UPSTART were from low income families. The evaluation of UPSTART's second 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 program's completion or "graduation" rate.

The Preschool Analysis

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, the parent's marital status, and the child's reported comfort level with computers. 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 model 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). 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 second year of UPSTART (i.e., Cohort 2). The descriptors included student demographics, the equipment that Cohort 2 (abbreviated as C2) students received, hours of UPSTART curriculum usage, and the graduation status of C2 students.

Findings: UPSTART Implementation

Most of the second year UPSTART participants (71%) received a computer drive with the UPSTART curriculum loaded on it. Approximately 12% of the second year participants were

loaned a computer and given free Internet access to help them access the UPSTART curriculum. Another 8% of the second year participants were loaned a personal computer to use at home while participating in UPSTART. The remaining 9% of the second year participants were provided with various combinations of educational technology – including cellular and wireless devices -- to enable them to access the UPSTART curriculum.

- The preschool test sample had a mean of 51 hours of participation in the UPSTART curriculum over the second year of the project. This compares with an average of 57 hours of instruction for program "graduates" in the test sample and an average of 49 hours of instruction for all students enrolled in UPSTART in the second year.

- Hours of instruction logged in the UPSTART curriculum was significantly and positively correlated with literacy skills measured by total posttest scores on the Brigance (r=.44) and the Bader (r=.22).

- Literacy skills measured by total posttest scores on the Brigance were shown to improve with increasing levels of UPSTART curriculum usage. This analysis controlled for initial levels of literacy skill development as measured by total pretest scores on the Brigance.

- The change in total Bader posttest scores as a function of instructional hours was not statistically significant. This analysis controlled for initial levels of literacy skill development as measured by total pretest scores on the Bader.

- The UPSTART graduation rate in the second year of the program was 76%. UPSTART graduation status was not significantly correlated with literacy development as measured by total posttest scores on the Brigance (r=.16) or the Bader (r=.15) of kindergarten.

Findings: UPSTART Impact on Literacy Development in Preschool

- UPSTART participation had a moderately strong impact on improving the literacy skills of UPSTART participants measured by the Brigance compared to nonparticipants. This impact could be accounted for by the treatment students' significantly better knowledge of lowercase letters and lowercase letter sounds. The statistical model controlled for initial levels of literacy development in knowledge of lowercase letters and sounds of lowercase letters.

- UPSTART participation also had a relatively small impact on improving the literacy skills of UPSTART participants measured by the Bader compared to nonparticipants. This impact could be accounted for by significantly better performance in the ability of treatment students to blend phonemes. The statistical model controlled for initial levels of literacy development measured by the Bader.

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

Introduction

UPSTART is a pilot project established by the Utah state legislature that uses a home-based education technology approach to develop the school readiness skills of preschool children. In its second year of operation during the 2010-11 school year, the project's implementation contractor – the Waterford Institute – enrolled 1,018 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 evaluation of UPSTART's second 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 program's completion or "graduation" rate.

A majority (60%) of the 1,018 preschool children that enrolled in the second year of UPSTART were from low income families, according to data provided by the Waterford Institute. Slightly more boys (51%) were enrolled than girls (49%). In terms of ethnicity, the vast majority (77%) of the enrollment was Caucasian, 14% were Hispanic, 3% were of Asian descent, 1% were Black, and 1% were Native American. Ethnicity for approximately 4% of the second year enrollment was unknown.

Background

Limitations in the first year evaluation of UPSTART included not being able to use a baseline or pretest measure of reading readiness at the preschool level. This resulted in evaluating the impact of the first year of UPSTART during kindergarten after the children had left the UPSTART preschool program, using a variation of the Posttest-Only Design with Nonequivalent Groups. First year results indicated that UPSTART students in at the beginning of kindergarten scored significantly higher on the DIBELS Next reading test compared to a matched control group, statistically controlling for ELL status, special education status, and gender. Follow-up results showed that the UPSTART students had maintained their achievement gains over their control group peers through the middle of kindergarten.

Uncertainties introduced by the posttest-only comparison group design included the following two concerns:

- **Selection bias**: We could not rule out whether the observed posttest differences in reading readiness were due to pre-existing differences.

- **Mortality**: We could not rule out whether the observed posttest differences were due to differential dropout rates of families from the study groups that might have changed the composition of the groups over time.

Thus, design limitations included not knowing whether selection effects or differential dropout rates accounted for the reading achievement differences we observed in kindergarten. In evaluating the second year of UPSTART, these problems were rectified by two design improvements: (1) using a pretest in fall 2010 with both treatment and control groups, and (2) conducting the testing during the preschool year.

Evaluation Design

The Cohort 2 evaluation implemented a quasi-experimental research design variant of the nonequivalent comparison group design. Specifically, the design implemented 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 study recruited 190 four year-old children; 95 treatment group children who had enrolled in UPSTART for Year 2 of the program and 95 control group children who had not enrolled in the UPSTART program.

In the diagram below, T stands for 4 year-old children who received the UPSTART preschool program during its second 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 measurements taken in the fall of 2010 and O2 indicates measurements taken in the fall of 2011.

NR	T	O1	X	O2
NR	C	O1		O2

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.[1] Because the study is not a randomized control trial, the groups are nonequivalent by definition, and consequently selection bias can be assumed to operate to some degree in some manner. The pretest allows us

[1] See Shadish, Cook, and Campbell (2002). *Experimental and Quasi-Experimental Designs for Generalized Causal*

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 the growth rates from pretest to posttest. If UPSTART shows stronger literacy growth rates, then the treatment group would be expected to show greater gain scores (posttest score minus pretest score) relative to the comparison group on the various subtests and total test scores.

Our research questions for the school readiness component of the evaluation study are as follows:

RQ1: D*o UPSTART students have better literacy skills at kindergarten than comparison 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 than comparison 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 are measures of early literacy skills relevant to emerging readers such as early phonemic awareness, letter recognition, letter sound knowledge and vocabulary development.

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 minutes of exposure per participant per week?*

RQ4: *What percent of 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 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 outcomes of interest.

Outcome Measures

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

- Phonological Awareness: rhyming, initial sound recognition, phonemic segmenting and blending.
- Phonics: letter name knowledge, sound knowledge, and word reading.
- Comprehension and Vocabulary: vocabulary knowledge and comprehension strategy development.
- Language Concepts: print concepts and basic oral language skills.

The Brigance. The Brigance IED 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.

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)

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

- *Expressive Grammar subtest*: the child is assessed on the ability to use plural *s*, *ing*, prepositions, and interpret and talk about an illustration. (Total possible subtest score = 12)

The Brigance academic and cognitive literacy scales included the:

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

The Bader. The Bader was selected as a measure of phonological awareness. Three subtests were administered from the Bader, 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 Total Bader score = 26)

Data Collection

As previously stated, 190 four year-old children were recruited for the C2 study; 95 treatment group children who had enrolled in UPSTART for Year 2 of the program and 95 control group children who had not enrolled in the UPSTART program. The children's parents were

administered an intake questionnaire (see Appendix A) in the fall of 2010 at the time their children were pretested on the Brigance and Bader. The children were posttested on the Brigance and Bader in the fall of 2011. Complete pretest and posttest data were obtained for 159 children on the Brigance (77 treatment children and 82 control children) and for 158 children on the Bader (76 treatment children and 82 control children).

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

The number of treatment and control children pretested and posttested with the Brigance is shown in Table 1 whereas the number of treatment and control children pretested and posttested with the Bader is shown in Table 2.

Table 1
Brigance Testing

Indicator	Treatment Group	Control Group
Pretested	94	95
Posttested	77	82
Difference	17	13
Attrition Rate	18%	14%

Table 2
Bader Testing

Indicator	Treatment Group	Control Group
Pretested	93	95
Posttested	76	82
Difference	17	13
Attrition Rate	18%	14%

As shown in Table 1, 77 treatment group children had both Brigance pretests and posttests as did 82 of the control group children. As shown in Table 2, 76 treatment group children had both Bader pretests and posttests as did 82 of the control group children. The final C2 study analysis examined the data collected for those children with both a pretest and posttest on the Brigance

and Bader. The principal effect of the observed study attrition was to reduce statistical power for the preschool analysis.

Analysis Strategy: Research Question 1

To determine whether UPSTART children have better literacy skills at kindergarten compared to control group children, the equivalence of the treatment and control groups in the final analysis sample was examined on the basis of the Brigance and Bader pretest scores and on the basis of those demographic characteristics that were significantly related to the posttests. Group equivalence on the pretests was examined using independent sample t-tests. Relationships between the demographics and the posttest scores were examined using correlation analyses.[3] 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.

Analysis Strategy: Research Question 2

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.[4]

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 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 were factored into four levels corresponding to

---

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

[4] To guard against Type I error (falsely rejecting the null hypothesis) in conducting tests of statistical inference (e.g., t-tests and multiple regression analysis) the criterion for statistical significance was based on the error rate for the collection of comparison required by the Brigance and the Bader. For example, comparisons among ten subtest means at the .01 level can result in at least one of the statistical tests being significant by chance: 10 (.01) = .10 which suggests that the effective significance level for the collection of comparisons is .10, not .01 in such a case (e.g., the Brigance comparions). This consideration resulted in a decision rule to set the confidence level at 99% 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.

quartiles. The ANCOVA was run separately for the Brigance and the Bader with Usage Group as the independent variable and respective pretest as a covariate.

Results

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

UPSTART Implementation

Findings reviewed under UPSTART implementation include enrollment in the second year, equipment provided to enrolled families by UPSTART, usage of the UPSTART curriculum, UPSTART graduates, and the relationship between levels of UPSTART curriculum usage and literacy outcomes.

UPSTART Enrollment

The Waterford Institute provided documentation for a second-year UPSTART enrollment of 1,018 children.

UPSTART Equipment Provided

The kind of education technology supports provided to children enrolled in UPSTART is shown in Table 3 for all 1,018 students enrolled in the second year and for the C2 preschool sample. The majority of the second year UPSTART students (approximately 70%) 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 C2 preschool sample most often (50%) also received a computer drive with the curriculum loaded on it.

Next most often, UPSTART loaned personal computers to almost 12% of the enrolled students and gave them free access to the Internet while they used the equipment. A slightly higher percentage of the preschool C2 sample (about 18%) received a free computer loan with free Internet access in the second year of the program. Another 8% of the program participants were given access to a home computer for free while they participated in the program. Among the C2 preschool sample, 18% were given access to a home computer for free while they participated in the program. The remaining 10% of the enrolled children received various combinations of computer technology to enable them to access the UPSTART curriculum.

Table 3
Percent of Students Provided Equipment by UPSTART

Equipment Provided	All UPSTART	Preschool Sample
Drive	70.6	50.0
Computer & Internet	11.8	18.1
Computer	8.1	18.1
Computer & Cellular	2.8	4.3
Computer & Wireless	1.6	--
Internet & Drive	1.6	1.1
Computer with Wireless & Internet	0.1	--
Cellular & Drive	0.7	2.1
Other	2.9	6.4
Sample Size	N = 1,018	N = 94

UPSTART Graduates

Of the 1,018 children documented as enrolled in UPSTART in the second year of the program, the Waterford Institute classified 776 as graduates of the program. This converts to a *graduation rate of 76%* for Cohort 2 (i.e., 776/1018 = .76, or 76%). UPSTART graduate status in year 2 of the program was significantly correlated with hours of instruction ($r=.34$, $p < .01$) but not with total posttest scores on the Brigance or the Bader.

UPSTART Usage

The average level of usage for all students enrolled in the second year of UPSTART was 49 hours of instruction, which converts to an expectation for 39 weeks of UPSTART implementation on the average, assuming 1.25 hours of instruction per week as called for by the UPSTART program design. The students in the C2 preschool sample used the UPSTART curriculum for 51 hours of instruction on the average, or an expectation of approximately 41 weeks of instruction over the course of the second program year. Students considered to be UPSTART graduates by the Waterford Institute used the UPSTART curriculum for 57 hours of instruction on the average, or an expectation for approximately 46 weeks of instruction over the course of the second program year.

The hours of instruction observed for all students documented to be enrolled in the second year of UPSTART are summarized in Table 4 compared to "graduates" and the students in the C2 preschool analysis sample.

Table 4
Hours of UPSTART Instruction

Group	N	Mean	SD	Range
All UPSTART	1,017	49.07	22.58	<1 – 175.81
UPSTART "Graduates"	776	56.96	19.08	26.39 - 175.81
C2 Preschool Sample	92	51.40	22.09	6.40 – 142.60

Examination of the data in Table 4 and the histograms in Figures 1-3 showing the distributions of hours of instruction for the three groups suggests that the preschool analysis sample is more representative of the second year UPSTART program population than the "graduates" group.

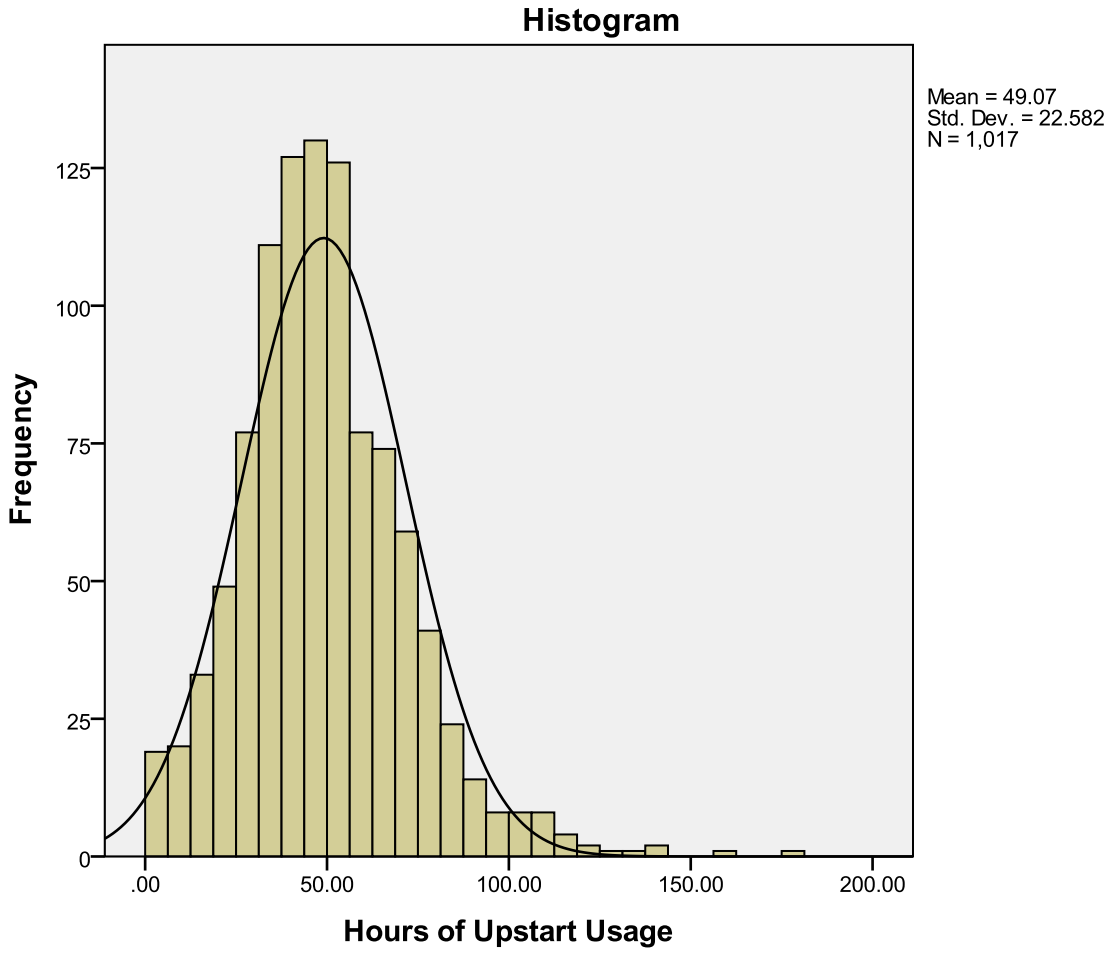


Figure 1. Hours of Instruction for All Students Enrolled in UPSTART in Year 2

Usage of the UPSTART curriculum for all students enrolled in the second year of the program (see Figure 1) is slightly right-skewed with a mean of approximately 49 hours of instruction and a standard deviation of approximately 23 hours. Because of the slight positive skew, the median

– 47 hours of instruction -- is a more accurate representation of the average usage of the UPSTART curriculum. Approximately 1.5% of the enrollment completed less than five hours of instruction. At the other end of the usage distribution, the top 1% of the enrollment completed 118 or more hours of instruction.

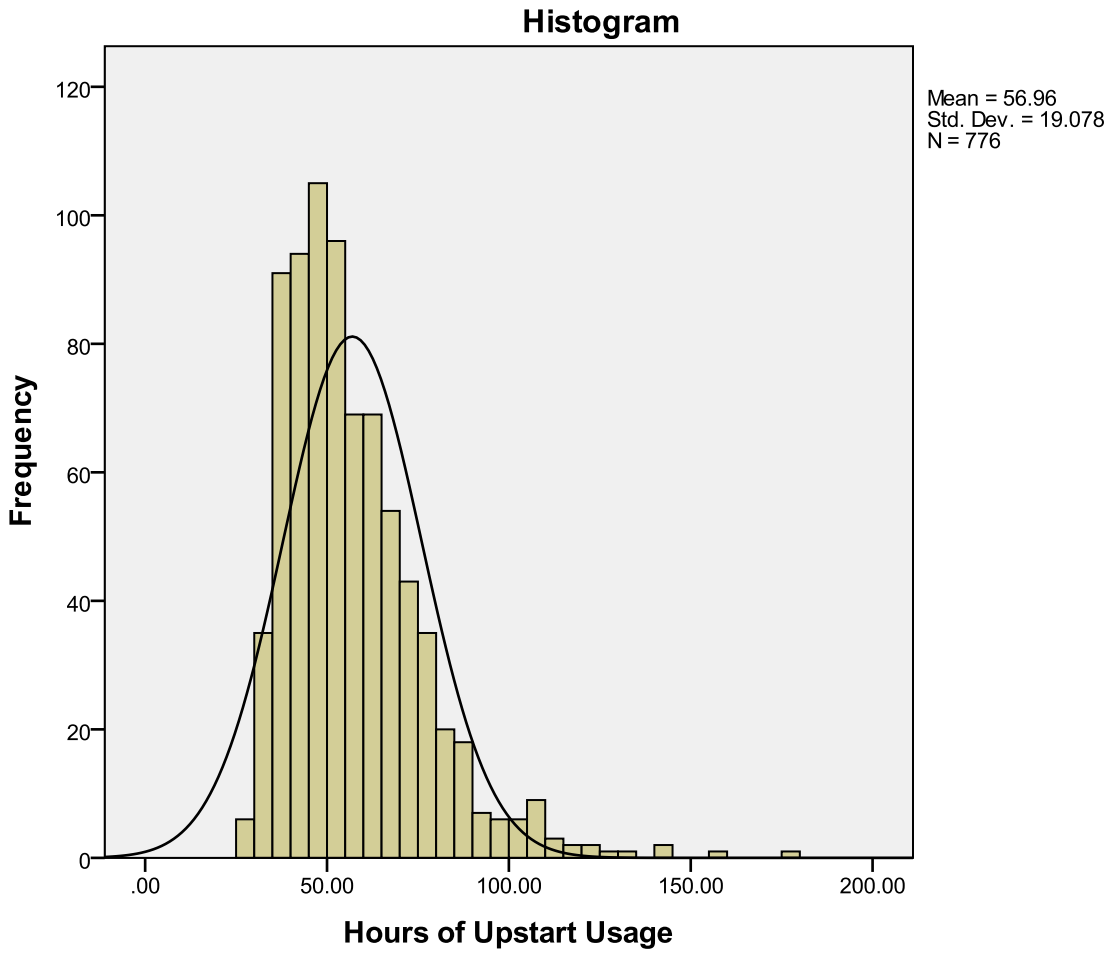


Figure 2. Hours of Instruction for UPSTART Graduates in Year 2

Lop off the bottom quintile from Figure 1 and you have Figure 2. The usage distribution for the C2 "graduates" subgroup starts with 23.4 hours of instruction and runs to 175.8 hours of instruction, as with the full program group. The graduates' distribution is right-skewed as well, which makes its median value of approximately 53 hours of instruction the more accurate representation of central tendency for this group. Because the bottom 20% of the graduates' distribution of usage hours has been removed, its skewness is actually more pronounced than that of the total program enrollment (i.e., a skew statistic of 1.547407 compared to 0.794).

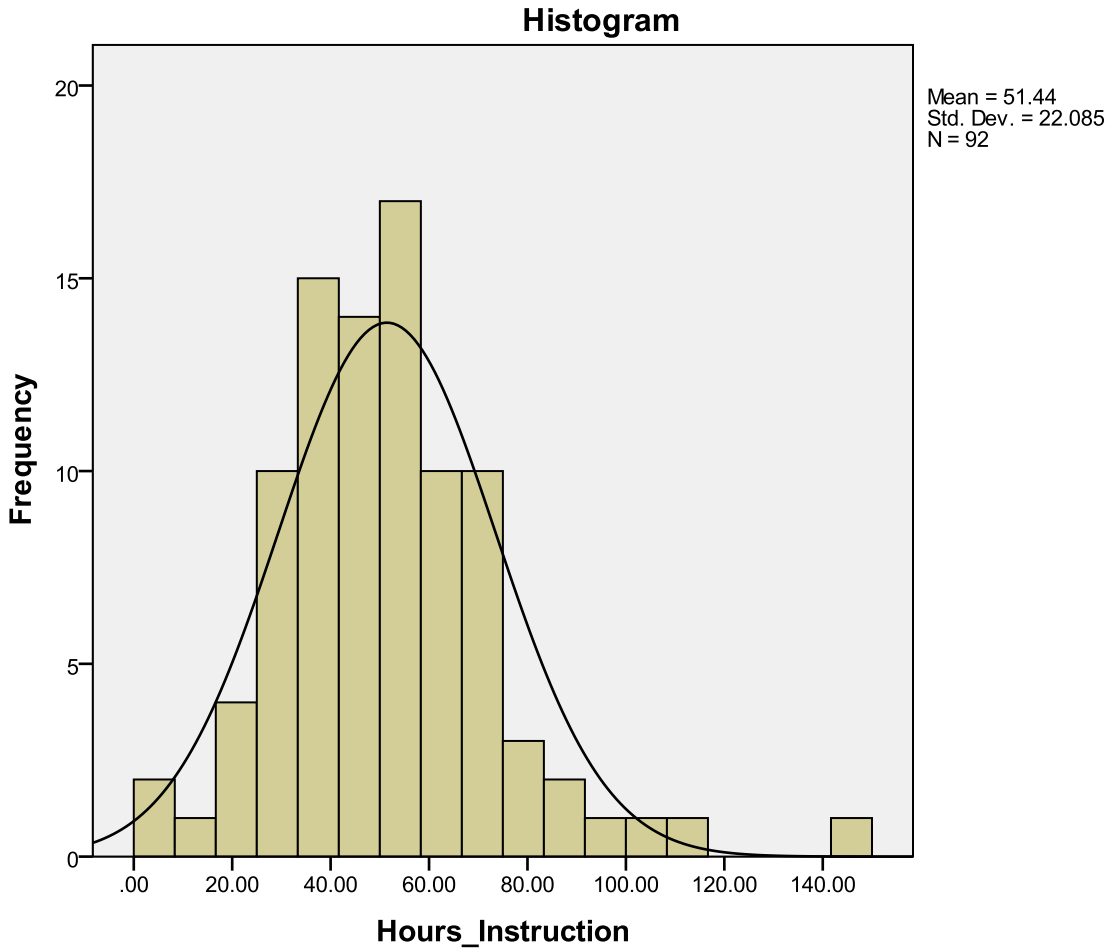


Figure 3. Hours of Instruction for Year 2 UPSTART Preschool Analysis Sample

UPSTART usage for the C2 preschool analysis sample (see Figure 3) is right-skewed (skew statistic = 1.036) with a mean of approximately 51 hours of instruction and a standard deviation of 22 hours. The analysis sample's median is 50 hours of instruction. For the analysis sample, hours of instruction are distributed as follows by quartile of usage:

- 1st Quartile: 6.4 hours to 37.2 hours
- 2nd Quartile: 37.3 hours to 49.9 hours
- 3rd Quartile: 50 hours to 63.92 hours
- 4th Quartile: 64 hours to 142.6 hours

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=.44, p<.01) whereas the relationship between UPSTART usage and literacy outcomes measured by the Bader Total Posttest were not quite so robust (r=.22, p=.056). 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.

Table 5 shows that UPSTART usage is significantly and positively related to posttest literacy outcomes measured by the Brigance Total Posttest, statistically controlling for initial levels of literacy[5] as measured by total scores on the Brigance pretest.[6] Sample size for the Brigance usage effects analysis was 75 and the observed statistical power was less than optimal.[7]

Table 5
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	60350.652[a]	4	15087.663	16.402	.000	.484
Intercept	87612.902	1	87612.902	95.243	.000	.576
Brigance_1	42485.424	1	42485.424	46.185	.000	.398
Usage Group	9891.127	3	3297.042	3.584	.018	.133
Error	64392.335	70	919.890			
Total	2465223.000	75				
Corrected Total	124742.987	74				

Adjusted R Squared = .45

In Table 6 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 preschool analysis sample and displays the difference in Brigance total posttest scores in the column labeled B – expressed as a regression coefficient. The parameter estimates for Usage Group indicate that participants in the first quartile of usage (37 hours of instruction or

[5] Exploratory analyses showed that the treatment and control groups differed on parental marital status (married vs. otherwise and the child's degree of computer comfort. However, these covariates proved to be statistically non-significant as control measures when entered into the regression model in conjunction with pretest scores.

[6] The Partial Eta Square statistic for Usage Group in Table 5 indicates that increasing exposure to the UPSTART curriculum accounts for about 13% of the literacy skills measured by the Brigance at the beginning of kindergarten. This indicates a moderately strong effect of UPSTART at the beginning of kindergarten.

[7] The observed power for the analysis was .77; the desired power for a given analysis is .80 or higher.

less in the preschool analysis sample) score approximately 31 points lower on the Brigance Total Posttest at the beginning of kindergarten than participants in the fourth quartile of usage (64 or more hours of instruction in the preschool analysis sample). The parameter estimates for the second and third quartiles of usage indicate that those who used the UPSTART curriculum between 37 and 63 hours scored approximately 21 points lower on the average than fourth quartile UPSTART users (64 or more hours). These estimates suggest a linear trend such that the development of early literacy skills at kindergarten entry tends to increase with increasing levels of UPSTART curriculum usage in preschool.

Table 5

Parameter Estimates: Brigance Total Posttest Score - Beginning K-

Parameter	B	Std. Error	t	Sig.	Partial Eta Squared
Intercept	123.597	12.777	9.674	.000	.572
Brigance_1	.595	.088	6.796	.000	.398
[Usage_Group=1.00]	-31.118	10.369	-3.001	.004	.114
[Usage_Group=2.00]	-21.855	9.351	-2.337	.022	.072
[Usage_Group=3.00]	-20.682	9.615	-2.151	.035	.062
[Usage_Group=4.00]	0[b]	.	.	.	.

[a.] Parameter set to zero because it is redundant

This positive linear trend in the development of early literacy skills measured by the Brigance as associated with UPSTART usage quartile is apparent in Figure 4 below. The line graph displays adjusted Brigance means at each quartile of instruction time, using pretest scores as the covariate. The adjusted mean estimates are shown in Table 6 below. Note: 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 = .84) but not for quartiles 2 and 3 (power = .64 and .56 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 1 point 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, and usage quartiles 2 and 3 versus quartile 4, because of the possibility that this lack of significance can be explained by low statistical power.

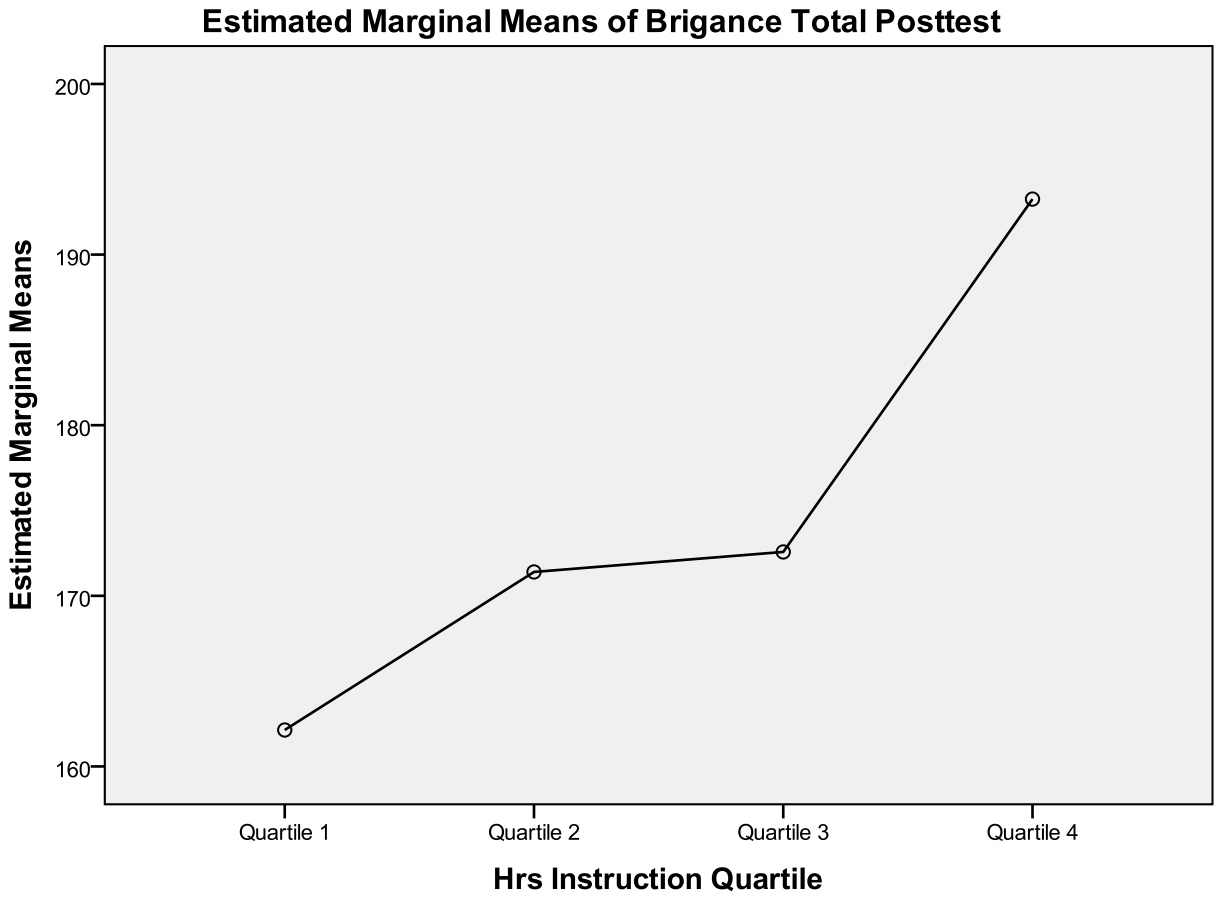
Table 6
Parameter Estimates: Brigance Total Posttest Score by Usage Quartile

Estimates

Dependent Variable: Brigance Total Posttest

Hrs Instruction Quartile	Mean	Std. Error	95% Confidence Interval	
			Lower Bound	Upper Bound
Quartile 1	162.136[a]	8.121	145.939	178.332
Quartile 2	171.399[a]	6.791	157.854	184.943
Quartile 3	172.572[a]	7.156	158.301	186.844
Quartile 4	193.254[a]	6.381	180.526	205.981

a. Covariates appearing in the model are evaluated at the following values: Brigance Total Pretest = 117.13.



Covariates appearing in the model are evaluated at the following values: Brigance Total Pretest = 117.13

Figure 4. Mean Brigance Total Postest Scores by UPSTART Usage Quartile

The results are somewhat different for the Bader at the beginning of kindergarten – see Table 7 -- which shows that increases in literacy skills measured by the Bader (phonological awareness) do not increase significantly with usage of the UPSTART curriculum (p=.50). As with the Brigance analysis, the Bader ANCOVA uses pretest Bader scores as a covariate to control for initial levels of phonological awareness. Sample size for the Bader usage effects analysis was 74 and the observed statistical power was low (power = .21).

Table 7
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	776.932[a]	4	194.233	3.360	.014	.163
Intercept	3223.278	1	3223.278	55.751	.000	.447
Bader_1	654.779	1	654.779	11.325	.001	.141
Usage Group	138.345	3	46.115	.798	.499	.034
Error	3989.284	69	57.816			
Total	19950.000	74				
Corrected Total	4766.216	73				

Adjusted R Squared = .11

The parameter estimates in Table 8 indicate that none of the usage group quartiles are significantly different from each other.

Table 8
Parameter Estimates: Bader Total Posttest Score - Beginning K-

Parameter	B	Std. Error	t	Sig.	Partial Eta Squared
Intercept	12.158	1.939	6.271	.000	.363
Bader_1	.536	.159	3.365	.001	.141
[Usage_Group=1.00]	-3.136	2.585	-1.213	.229	.021
[Usage_Group=2.00]	-2.788	2.327	-1.198	.235	.020
[Usage_Group=3.00]	-.517	2.439	-.212	.833	.001
[Usage_Group=4.00]	0[b]	.	.	.	.

The lack of statistically significant differences across usage quartile for the Bader is more evident from the results shown in Table 9 for adjusted Bader Total Posttest means by usage quartile.

Table 9
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	12.654[a]	2.045	8.574	16.734
Quartile 2	13.002[a]	1.702	9.607	16.397
Quartile 3	15.273[a]	1.850	11.581	18.964
Quartile 4	15.790[a]	1.586	12.626	18.954

a. Covariates appearing in the model are evaluated at the following values: Bader Total Pretest = 6.7703.

Again, the possibility exists that statistical power is an explanatory factor underlying the lack of statistical significance for differences in UPSTART usage as measured by Bader literacy outcomes. This is particularly compelling when consideration is given to the fact that the Total Bader scale range is relatively narrow (i.e., scores from zero to 26).

UPSTART Outcomes

In this section, the main research questions of interest to the C2 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

- Treatment-Control Group Differences
- Identifying significant posttest predictors

- Posttest Analysis
- Multiple Regression Analysis

Brigance Pretest and Posttest Results

The performance of the treatment and control group children was essentially equivalent on the overall Brigance at the pretest, although the control group children scored significantly higher on the auditory discrimination pretest (see Appendix B for details). There were several demographic differences between the treatment and control children in the Brigance analysis sample but only the child's comfort level with a computer and the parent's marital status (being married or not) were significantly related to Brigance posttest outcomes. Both of these covariates favored the treatment group over the control group. See Appendix C and Appendix D for details.

Posttest results showed that the treatment group performed significantly better than the control children on the *overall Brigance* (by an average difference of 7.9 points overall). In terms of subtest performance, this difference showed up as significantly higher ($p<.01$) posttest performance for the treatment group children on the *Lowercase Letter Knowledge* test and the *Sounds of Lowercase Letters* test (an average difference of 2.6 points on each of the two subtests). See Appendix E for details.

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 almost 25 points on the average. The final Brigance regression model is shown in Tables 10 and 11.

Table 10

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

	Sum of Squares	Df	Mean Square	F	Sig.
Regression	180595.811	2	90297.906	110.726	.000
Residual	127219.484	156	815.510		
Total	307815.296	158			

Table 11
OLS Regression Coefficients for Total Brigance Posttest – Beginning of Kindergarten

	Unstandardized Coefficients		Standardized Coefficients			Partial
	B	Std. Error	Beta	t	Sig.	Correlation
(Constant)	59.574	7.555		7.886		
TOTAL PRETEST	.787	.055	.740	14.302	.000	.753
STUDY GROUP	24.882	4.553	.283	5.465	.000	.401

Adjusted R Square = .58

The observed effect size for the UPSTART treatment group is in the medium effects size range (partial R square for Study Group = .16). Preliminary regression models showed that the effect of the two demographic covariates (child's level of comfort with a computer and parent's marital status) were statistically non-significant when entered into the regression equation with the total Brigance pretest.

UPSTART impact as measured by the Brigance is attributable to the treatment group's superior posttest performance on the *Lowercase Letter Knowledge* subtest and the *Sounds of Lowercase Letters* subtest. This conclusion is based on the t-test results shown in Appendix E and on a series of regression models that examined treatment group effects holding subtest performance constant. The only subtests that proved to be statistically significant in predicting study group differences on the Brigance were the Lowercase Letter Knowledge subtest and the Sounds of Lowercase Letters subtest.

Bader Pretest and Posttest Results

The pretest performance of the treatment and control group children was essentially equivalent on the overall Bader and on each of the Bader subtests; see Appendix F[8]. Posttest results showed a statistically significant treatment group effect on the overall Bader as well as on the Phoneme Blending subtest; see Appendix G. While there were several demographic differences between the treatment and control children in the Bader analysis sample (see Appendix G), only the parent's marital status (being married or not) was significantly related to Bader posttest outcomes (see Appendix D). This initial difference favored the treatment group (93% of the treatment group parents were married vs. 80% for the control group).

[8] The control group scored higher than the UPSTART treatment group on all of the Bader pretests, but these differences were not statistically significant. None of the subtest differences were statistically significant at either the .05 level or the .01 level. The latter was the criterion of significance set for this evaluation. Differences on the total Bader pretest were statistically significant at p = .02.

Adjusting for the initial differences using multiple regression analysis, it was found that the treatment group outscored the control group on the Bader overall posttest by approximately 4 points on the average. The final Bader regression model is shown in Tables 12 and 13.

Table 12
OLS ANOVA Summary Table for Total Bader Posttest – Beginning of Kindergarten

	Sum of Squares	Df	Mean Square	F	Sig.
Regression	2876.142	2	1438.071	32.202	.000
Residual	6921.858	155	44.657		
Total	9798.000	157			

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)	5.855	1.071		5.468		
TOTAL PRETEST	.667	.087	.527	7.685	.000	.525
STUDY GROUP	3.954	1.082	.251	3.656	.000	.282

Adjusted R Square = .28

The observed effect size for the UPSTART treatment group on the Bader is in the small effects size range (partial R square for Study Group = .08). Preliminary regression models showed that the effect of the demographic covariate (parent's marital status) was statistically non-significant when entered into the regression equation with the total Bader pretest.

UPSTART impact as measured by the Bader is attributable to the treatment group's superior posttest performance on the *Phoneme Blending* subtest. This conclusion is based on the t-test results shown in Appendix G and on a series of regression models that examined treatment group effects holding subtest performance constant. The only subtest that proved to be statistically significant at the 99% confidence interval in predicting study group differences on the Bader was the *Phoneme Blending* test.

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=77) on the Total Brigance and on nine of the ten subtests. Treatment group growth on the Receptive Objects subtest was not statistically significant. For the matched Brigance control group (N=82), there was statistically significant growth on the Total Brigance and on eight of the ten subtests. Control group growth on the Receptive Objects subtest and the Auditory Discrimination subtest was not statistically significant.

Differences in growth rates between the treatment and control group were significantly different at the 99% CI for the *overall Brigance* and the *Sounds of Lower Case Letters s*ubtest, both of which favored the UPSTART treatment group. These results are shown in Table 14.

Table 14
Treatment-Control Group Differences in Growth Rates on the Brigance

Brigance Test	Treatment Group		Control Group		Significance p<.01
	Mean Growth	99% Confidence Interval Growth Rate	Mean Growth	99% Confidence Interval Growth Rate	
Expressive Objects	0.545	0.15 – 0.94	0.915	0.47 – 1.36	NS
Receptive Objects	0.234	0.01 – 0.46	0.305	-0.10 - 0.71	NS
Expressive Grammar	1.208	0.75 1.66	1.04	0.51 – 1.57	NS
Visual Discrimination	5.260	3.62 – 6.90	3.67	2.37 – 4.95	NS
Recites Alphabet	9.740	6.46 -13.03	4.00	1.15 – 6.85	NS
Lowercase Letter Knowledge	19.948	14.46 -25.44	11.54	7.30 – 15.77	NS
Lowercase Letter Sounds	10.442	7.75 – 13.13	5.43	3.21 – 7.64	**
Auditory Discrimination	3.558	2.27 – 4.85	1.00	-0.31 – 2.31	NS
Survival Sight Words	2.247	1.28 -3.22	1.31	0.69 – 1.92	NS
Basic Preprimer Vocabulary	6.403	3.94 -8.87	3.83	2.11 – 5.55	NS
Total Brigance	59.84	49.18 - 69.99	33.01	25.88 - 40.15	**

Figure 5 shows the growth measured by the Total Brigance from pretest to posttest for the matched samples. The overall initial difference between the two groups is statistically non-significant. At posttest, the line graph reveals that the UPSTART treatment group has pulled away from the control group, demonstrating greater overall growth in phonics skills – particularly in the ability to produce sounds of lower case letters – compared to the control group.

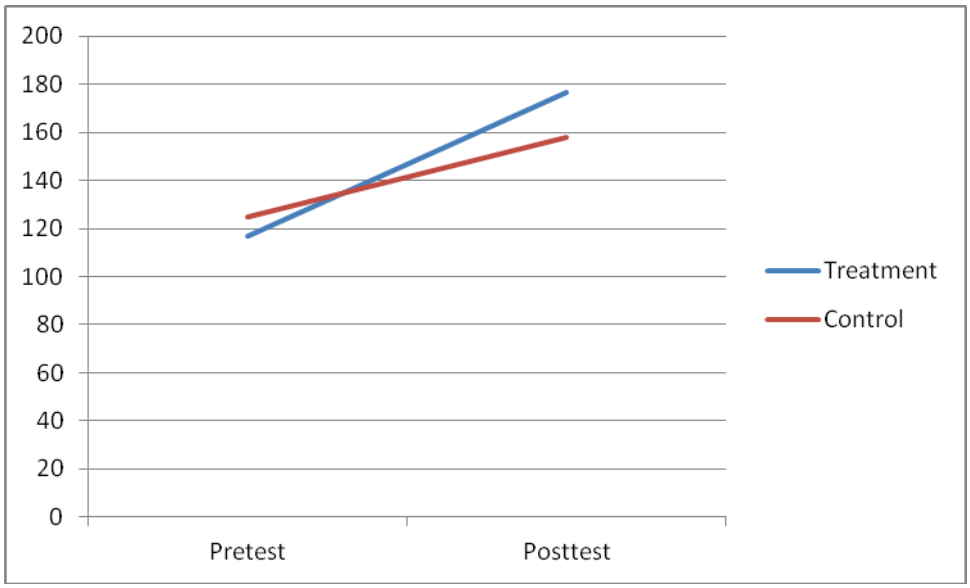


Figure 5. Treatment and Control Group Growth on the Brigance from Pretest to Posttest

Bader Growth Score Results

There was statistically significant growth from pretest to posttest for the matched Bader treatment group sample (N=76) on the Total Bader and all subtests. For the matched Bader control group (N=82), there was statistically significant growth on the Total Bader and on two of the three subtests. Control group growth on the Rhyme Recognition subtest was not statistically significant.

Differences in growth rates between the treatment and control group were significantly different at the 99% CI for the *overall Bader* and the *Phoneme Blending* subtest. These results are shown in Table 15.

Table 15

Treatment-Control Group Differences in Growth Rates on the Bader

Bader Test	Treatment Group		Control Group		Significance p<.01
	Mean Growth	99% Confidence Interval Growth Rate	Mean Growth	99% Confidence Interval Growth Rate	
Rhyme Recognition	2.18	1.02 – 3.35	0.87	-0.28 – 2.01	NS
Phoneme Blending	3.21	2.26 – 4.16	1.06	0.29 - 1.83	**
Phoneme Segmenting	2.18	1.08 - 3.29	0.95	0.24 – 1.67	NS
Total Bader	7.58	5.18 - 9.97	2.88	1.13 - 4.62	**

Figure 6 shows the growth measured by the Total Bader from pretest to posttest for the matched samples. The initial pretest difference between the two groups is statistically non-significant. The posttest difference between the treatment and control group is statistically significant and reveals that the UPSTART children have developed their phonemic awareness skills – specifically phoneme blending skills – significantly more than have the control group children.

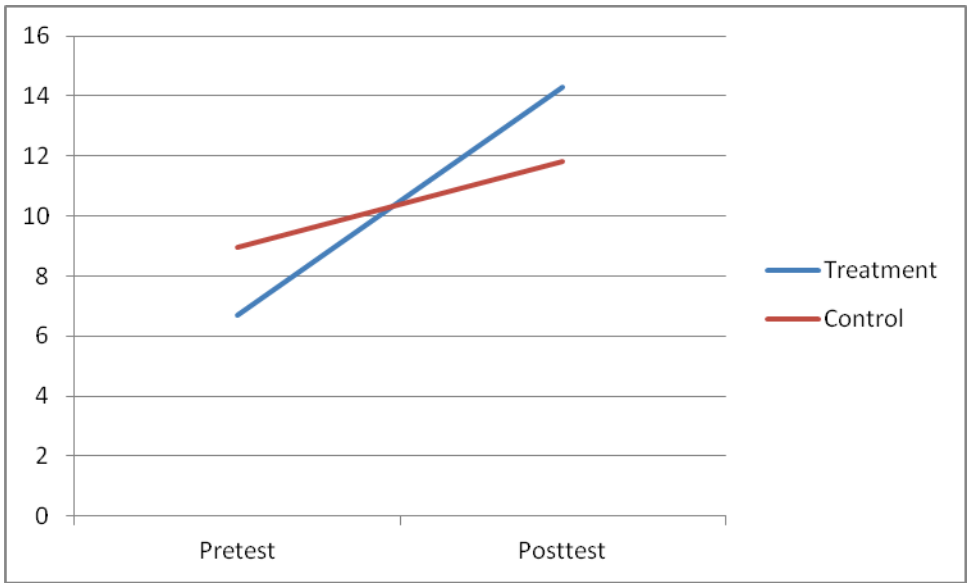


Figure 6. Treatment and Control Group Growth on the Bader from Pretest to Posttest

Summary and Conclusions

This final section of the Year Two evaluation report summarizes:

- The data that were collected and analyzed;
- The analysis methods employed;
- Findings regarding UPSTART implementation; and
- Findings on UPSTART's impact on literacy outcomes as measured in preschool and at the beginning of kindergarten.

Test Data Collected and Analyzed

190 four year-old children were recruited for the C2 study; 95 treatment group children who had enrolled in UPSTART for Year 2 of the program and 95 control group children who had not enrolled in the UPSTART program. The children's parents were administered an intake questionnaire (see Appendix A) in the fall of 2010 at the time their children were pretested on the Brigance and Bader. The children were posttested on the Brigance and Bader in the fall of 2011. Complete pretest and posttest data were obtained and analyzed for 159 children on the Brigance (77 treatment children and 82 control children) and for 158 children on the Bader (76 treatment children and 82 control children).

The 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.[9] 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

[9] 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.

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 in estimating the effect of UPSTART usage on literacy skills at the beginning of kindergarten as measured by the posttest on each respective measure. In the ANCOVA analyses, the effects of usage at quartiles one through three were compared with usage at the fourth quartile.

Findings: UPSTART Implementation

The Waterford Institute provided documentation for a second-year UPSTART enrollment of 1,018 children. A majority (60%) of the 1,018 preschool children that enrolled in the second year of UPSTART were from low income families, according to data provided by the Waterford Institute. Slightly more boys (51%) were enrolled than girls (49%). In terms of ethnicity, the vast majority (77%) of the enrollment was Caucasian, 14% were Hispanic, 3% were of Asian descent, 1% were Black, and 1% were Native American. Ethnicity was unknown for 4% of the second year enrollment.

Most of the second year participants (70%) received a computer drive with the UPSTART curriculum loaded on it. Approximately 12% of the second year participants received a computer loan and free Internet access to help them access the UPSTART curriculum. Another 8% of the second year participants were loaned a personal computer to use at home while participating in UPSTART. The remaining 10% of the second 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 C2 preschool test sample had a mean of 51 hours of UPSTART curriculum usage over the second year of the project. This compares with an average of 57 hours of instruction for program "graduates" and an average of 49 hours of instruction for all students enrolled in UPSTART in the second year.

- 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=.44) and the Bader (r=.22).

- Literacy skills measured by the Brigance improved with increasing levels of UPSTART curriculum usage. This was not the case with literacy skills measured by the Bader, which did not improve significantly with increasing levels of UPSTART curriculum usage

- The UPSTART graduation rate in the second year of the program was 76%. UPSTART graduation status in the second year of the program was significantly correlated with hours of instruction (r=.34) but not with literacy outcomes measured at the beginning of kindergarten by the Brigance (r=.16) or the Bader (r=.15).

Findings: UPSTART Impact on Literacy Skills at the beginning of Kindergarten

- UPSTART participation had a moderately strong impact on improving the phonics skills of UPSTART participants compared to nonparticipants at the beginning of kindergarten as measured by the Brigance. The observed effects were mostly due to improvements in the UPSTART children's knowledge of lowercase letters and their ability to produce sounds of lower case letters. This analysis controlled for initial literacy levels as measured by the Brigance pretest.

- UPSTART participation had a relatively small impact on improving the phonological awareness of UPSTART participants compared to nonparticipants at the beginning of kindergarten, as measured by the Bader. The observed effects were mostly due to improvements in the UPSTART children's ability to blend phonemes. This analysis controlled for initial literacy levels as measured by the Bader pretest.

- UPSTART participants showed significantly stronger growth rates on the *overall Brigance* and the *Sounds of Lower Case Letters* subtest compared to control group children.

- UPSTART participants showed significantly stronger growth rates on the *overall Bader* and the *Phoneme Blending* subtest compared to control group children.

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?
   □$_1$ Yes □$_2$ No

   1a. If yes, did your 4-year-old also use the program?
   □$_1$ Yes □$_2$ No

Child Information

2. What is your child's birthday? _____

3. What is your child's gender?
   □$_1$ Male □$_2$ Female

4. What year will your child be entering Kindergarten?
   □$_1$ 2011 □$_2$ 2012

5. What is your child's ethnicity?
   □$_1$ Hispanic □$_2$ Native American/Alaskan Native □$_3$ Asian/Pacific Islander
   □$_4$ Caucasian □$_5$ African American □$_6$ Other:_____

6. What is your child's primary language?
   □$_1$ English □$_2$ Spanish □$_3$ Portuguese □$_4$ Chinese
   □$_5$ German □$_6$ Japanese □$_7$ Other:_____

7. Is your child currently attending a daycare/preschool?
   □$_1$ Yes □$_2$ No

   7a. **If yes**, approximately how many hours a week does your child attend a daycare/preschool?
   □$_1$ less than 10 hours □$_2$ 10-19 hours □$_4$ 20-24 hours
   □$_5$ 25-29 hours □$_6$ 30-34 hours □$_7$ 35 or more hours

8. Does your child have access to a computer in your house?
   □$_1$ Yes □$_2$ No

9. Does your child use a computer in her/his day care or preschool?
   □$_1$ Yes □$_2$ No □$_3$ Not Applicable (not in day care or preschool)

10. How comfortable is your child using a computer?
    □$_1$ Very comfortable □$_2$ Somewhat comfortable □$_3$ Somewhat uncomfortable
    □$_4$ Not comfortable □$_5$ Very uncomfortable

Caregiver Information

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

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

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

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

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

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

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

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

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

Thank you for participating in the Utah UPSTART Evaluation!

Appendix B
Brigance Pretest Analysis of Treatment-Control Group Differences

Brigance PreTest	*Group*	*N*	*Mean*	*SEM*	*t*	*Significance*
Expressive Objects	Treatment	77	25.40	0.17	1.31	NS
	Control	82	25.05	0.21		
Receptive Objects	Treatment	77	26.74	0.08	0.46	NS
	Control	82	26.66	0.15		
Expressive Grammar	Treatment	77	8.90	0.17	0.13	NS
	Control	82	8.87	0.17		
Visual Discrimination	Treatment	77	11.92	0.66	-1.40	NS
	Control	82	13.10	0.52		
Recites Alphabet	Treatment	77	9.03	1.09	-2.23	NS
	Control	82	12.45	1.08		
Lowercase Letter Knowledge	Treatment	77	21.36	2.14	-0.42	NS
	Control	82	22.67	2.24		
Sounds of Lowercase Letters	Treatment	77	6.25	0.95	-0.66	NS
	Control	82	7.18	1.05		
Auditory Discrimination	Treatment	77	4.18	0.44	-3.33	**
	Control	82	6.15	0.40		
Survival Sight Words	Treatment	77	1.49	0.25	-0.16	NS
	Control	82	1.55	0.24		
Basic Preprimer Vocabulary	Treatment	77	1.53	0.59	0.63	NS
	Control	82	1.07	0.44		
Total Brigance	Treatment	77	116.81	4.65	-1.21	NS
	Control	82	124.74	4.64		

**p<.01

Appendix C
Brigance Sample: Treatment – Control Differences on Demographics

Covariate	*Group*	*N*	*Mean*	*SEM*	*t*	*Significance*
% Male	Treatment	77	48	5.80		
	Control	81	44	5.60	0.53	NS
% Caucasian	Treatment	76	80	4.65		
	Control	81	78	4.65	0.34	NS
% Hispanic	Treatment	76	09	3.38		
	Control	81	16	4.10	-1.26	NS
% Primary language is English	Treatment	77	95	2.58		
	Control	81	93	2.93	0.55	NS
% Attend preschool 10+ hours per week	Treatment	36	47	8.44		
	Control	64	36	6.05	1.10	NS
%Currently attending preschool or daycare	Treatment	76	49	5.80		
	Control	81	79	4.60	-4.02	**
% Child has access to a computer	Treatment	76	91	3.40		
	Control	81	88	3.70	0.60	NS
Child uses PC in preschool or daycare	Treatment	48	1.79	0.06		
	Control	66	1.65	0.06	1.68	NS
Child comfort level with computers	Treatment	75	4.25	0.09		
	Control	79	3.87	0.12	2.56	**
% Caregiver is mother	Treatment	76	89	3.59		
	Control	81	91	3.14	-0.43	NS
% Parent is Caucasian	Treatment	76	87	3.95		
	Control	81	86	3.83	0.05	NS
% Parent is Hispanic	Treatment	76	07	2.90		
	Control	81	11	3.51	-0.98	NS
% Parent's primary language is English	Treatment	76	93	2.90		
	Control	81	90	3.51	0.98	NS
Parent Educational Attainment (recoded)[10]	Treatment	76	3.36	0.10		
	Control	81	3.26	0.10	0.72	NS
% Parent is married	Treatment	76	93	2.90		
	Control	80	80	4.50	2.49	**
Parent employment status	Treatment	76	1.61	0.10		
	Control	81	1.67	0.09	-0.41	NS
Spouse employment status	Treatment	72	2.80	0.07		
	Control	70	2.74	0.08	0.59	NS
Household size	Treatment	76	5.04	0.10		
	Control	81	4.68	0.12	2.27	NS
Household income category	Treatment	76	3.71	0.12		
	Control	80	3.69	0.13	0.11	NS

**p<.01

[10] 1 = HS Dropout; 2 – HS Graduate; 3= Some College; 4 = College Graduate; 5 = Graduate Degree

Appendix D
Pearson Correlations with Total Posttest Scores

Variable	Brigance	Bader
Study Group[11]	.21**	.16
Pretest	.71**	.48**
Male	-.08	-.13
Caucasian	.20	.19
Hispanic	-.23**	-.22**
Attends Daycare/Preschool	-.02	.04
Hours per week daycare/preschool	-.17	-.20
Computer access	.22**	.21**
Uses computer at daycare/preschool	.21	.19
Computer comfort	.21	..10
Parent employment status	-.03	-.04
Spouse employment status	-.07	-.11
Household size	.11	-.03
Household income	.28**	.31**
Primary language is English	.13	.19
Attends daycare/preschool 10+ hours per week	-.25	-.30**
Caregiver is mother	.01	.02
Parent is Caucasian	.18	.14
Parent is Hispanic	-.18	-.19
Parent's primary language is English	.13	.18
Parent Educational Attainment (recoded)	.30**	.22**
Parent is married	.30**	.22**

**p<.01

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

Appendix E
Brigance Posttest Analysis of Treatment-Control Group Differences

Brigance PreTest	*Group*	*N*	*Mean*	*SEM*	*t*	*Significance*
Expressive Objects	Treatment	77	25.95	0.12	-0.09	NS
	Control	82	25.96	0.12		
Receptive Objects	Treatment	77	26.97	0.08	0.38	NS
	Control	82	26.96	0.02		
Expressive Grammar	Treatment	77	08.90	0.02	0.97	NS
	Control	82	10.10	0.17		
Visual Discrimination	Treatment	77	17.18	0.40	0.77	NS
	Control	82	16.76	0.38		
Recites Alphabet	Treatment	77	18.77	1.06	1.51	NS
	Control	82	16.45	1.10		
Lowercase Letter Knowledge	*Treatment*	*77*	*41.31*	*1.67*	*2.63*	**
	Control	*82*	*34.21*	*2.13*		
Sounds of Lowercase Letters	*Treatment*	*77*	*16.69*	*1.09*	*2.56*	**
	Control	*82*	*12.61*	1.17		
Auditory Discrimination	Treatment	77	7.74	0.31	1.24	NS
	Control	82	7.15	0.36		
Survival Sight Words	Treatment	77	3.74	0.45	1.61	NS
	Control	82	2.85	0.32		
Basic Preprimer Vocabulary	Treatment	77	7.94	1.04	2.28	NS
	Control	82	4.90	0.83		
Total Brigance	Treatment	77	176.39	4.72	2.71	**
	Control	82	157.76	4.96		

**p<.01

Appendix F
Bader Pretest Analysis of Treatment-Control Group Differences

Bader PreTest	*Group*	*N*	*Mean*	*SEM*	*t*	*Significance*
Rhyme Recognition	Treatment	76	4.76	0.40	-1.86	NS
	Control	82	5.77	0.36		
Phoneme Blending	Treatment	76	1.41	0.29	-1.78	NS
	Control	82	2.20	0.34		
Phoneme Segmenting	Treatment	76	0.53	0.19	-1.39	NS
	Control	82	0.98	0.25		
Total Bader	Treatment	76	6.70	0.65	-2.28	NS
	Control	82	8.94	0.73		

** p<.01

Appendix G
Bader Posttest Analysis of Treatment-Control Group Differences

Bader Posttest	*Group*	*N*	*Mean*	*SEM*	*t*	*Significance*
Rhyme Recognition	Treatment	76	6.95	0.40	0.59	NS
	Control	82	6.63	0.36		
Phoneme Blending	Treatment	76	4.62	0.38	2.55	**
	Control	82	3.26	0.40		
Phoneme Segmenting	Treatment	76	2.71	0.37	1.58	NS
	Control	82	1.93	0.33		
Total Bader	Treatment	76	14.28	0.93	1.97	NS
	Control	82	11.82	0.84		

** p<.01

Appendix H
Bader Sample: Treatment – Control Differences on Demographics

Covariate	*Group*	*N*	*Mean*	*SEM*	*t*	*Significance*
% Male	Treatment	76	49	5.80	0.53	NS
	Control	81	44	5.60		
% Caucasian	Treatment	75	80	4.65	0.34	NS
	Control	81	78	4.65		
% Hispanic	Treatment	75	09	3.38	-1.26	NS
	Control	81	16	4.10		
% Primary language is English	Treatment	76	95	2.58	0.55	NS
	Control	81	93	2.93		
% Attend preschool 10+ hours per week	Treatment	36	47	8.44	1.10	NS
	Control	64	36	6.05		
%Currently attending preschool or daycare	Treatment	76	49	5.80	-4.02	**
	Control	81	79	4.60		
% Child has access to a computer	Treatment	75	91	3.40	0.60	NS
	Control	81	88	3.70		
Child uses PC in preschool or daycare	Treatment	48	1.79	0.06	1.68	NS
	Control	66	1.65	0.06		
Child comfort level with computers	Treatment	75	4.25	0.09	2.56	**
	Control	79	3.87	0.12		
% Caregiver is mother	Treatment	75	89	3.59	-0.43	NS
	Control	81	91	3.14		
% Parent is Caucasian	Treatment	75	87	3.95	0.05	NS
	Control	81	86	3.83		
% Parent is Hispanic	Treatment	75	07	2.90	-0.98	NS
	Control	81	11	3.51		
% Parent's primary language is English	Treatment	75	93	2.90	0.98	NS
	Control	81	90	3.51		
Parent Educational Attainment (recoded)[12]	Treatment	75	3.36	0.10	0.72	NS
	Control	81	3.26	0.10		
% Parent is married	Treatment	75	93	2.90	2.49	**
	Control	80	80	4.50		
Parent employment status	Treatment	75	1.61	0.10	-0.41	NS
	Control	81	1.67	0.09		
Spouse employment status	Treatment	71	2.80	0.07	0.59	NS
	Control	70	2.74	0.08		
Household size	Treatment	75	5.04	0.10	2.27	NS
	Control	81	4.68	0.12		
Household income category	Treatment	75	3.71	0.12	0.11	NS
	Control	80	3.69	0.13		

**p<.01

[12] 1 = HS Dropout; 2 – HS Graduate; 3= Some College; 4 = College Graduate; 5 = Graduate Degree