

2012 Report of the UPSTART Program

Background

Waterford Research Institute, a nonprofit founded in 1976 with a mission to use technology to help provide educational equity and excellence for all children, was selected to administer the UPSTART Program in March 2009.

As planned and administered by Waterford, UPSTART (Utah Preparing Students Today for A Rewarding Tomorrow) provides a unique response to the need for preschool training in Utah, providing the State's youngest students with access in their homes to outstanding instruction in reading, math, and science. The overarching goal of the UPSTART Program is to provide truly individualized instruction, serving children with the greatest needs, while, at the same time, challenging gifted children to reach their full potential.

UPSTART is comprised of three software programs:

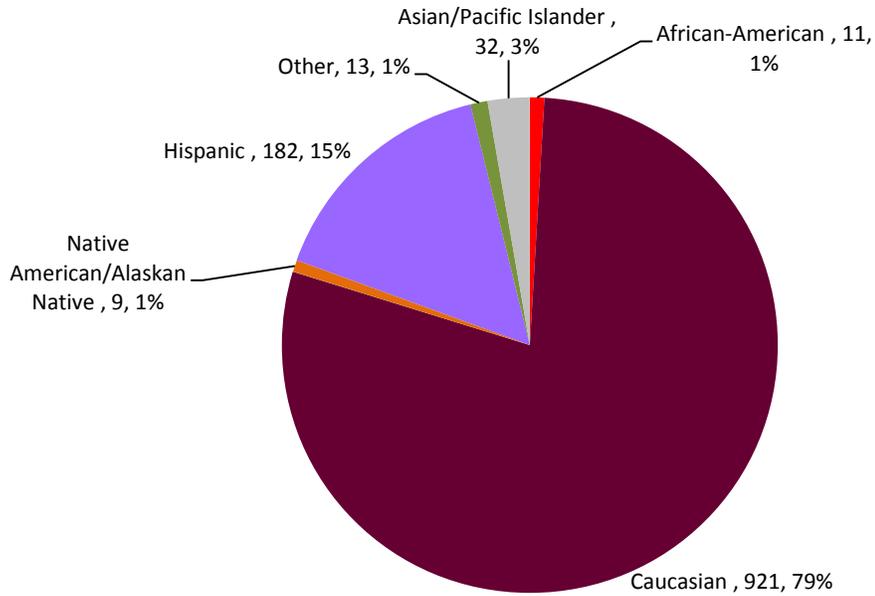
- *Rusty and Rosy Learn with Me™* (RRLWM) delivers individualized instruction in reading, math, and science that adapts to each child's learning trajectory. The program includes: 360 digital books; 330 animated songs; more than 7,000 total activities; 2,500 lessons; and 450 instructional hours that meet national, state, and professional standards and guidelines.
- The *Waterford Assessments of Core Skills™* is a fundamental testing breakthrough for assessing very young children who do not know how to read. It is computer adaptive, offers immediate scoring and reports, and is completed in one to two brief sessions.
- *Camp Consonant™* offers additional help with reading. Set in a fun camp setting, it features 3,700 unique activities, including songs, games, reading passages, and a built-in motivational system. It provides more than 150 hours of instruction and features a brain-research-based, multi-sensory approach.

The UPSTART Program recognizes the home and parents and caregivers as key educational resources. The home provides the benefit that education can take place seven days a week without the need to travel for access to instruction. Parents and caregivers can provide the motivation for children to ensure that they spend the necessary time on program materials. To capitalize on these key resources, UPSTART provides a special support organization for participating children and their parents or caregivers. Unlike a typical support structure which is passive in relation to users except when there are problems and questions to be solved, the UPSTART support team maintains frequent contact through written materials, in-person and online training, emails, and telephone calls. The strategy is to provide a steady stream of data on children's usage, performance, and needs, as well as to introduce motivational strategies for maintaining their interest.

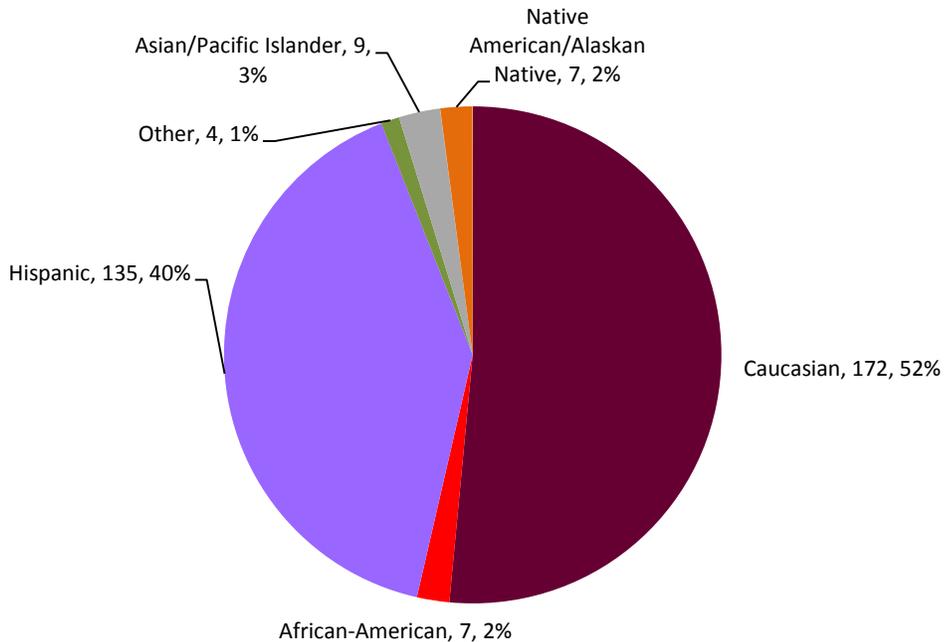
Year 3

In Year 3, Waterford was able to serve all families interested in participating in UPSTART. A total of 1,151 families participated with 1,168 UPSTART children. Of those, 331 families with 334 children received computers, Internet, or both. The following charts show diversity, SES, income level, "other" preschool attendance, and population density information for all participants and the subset receiving equipment.

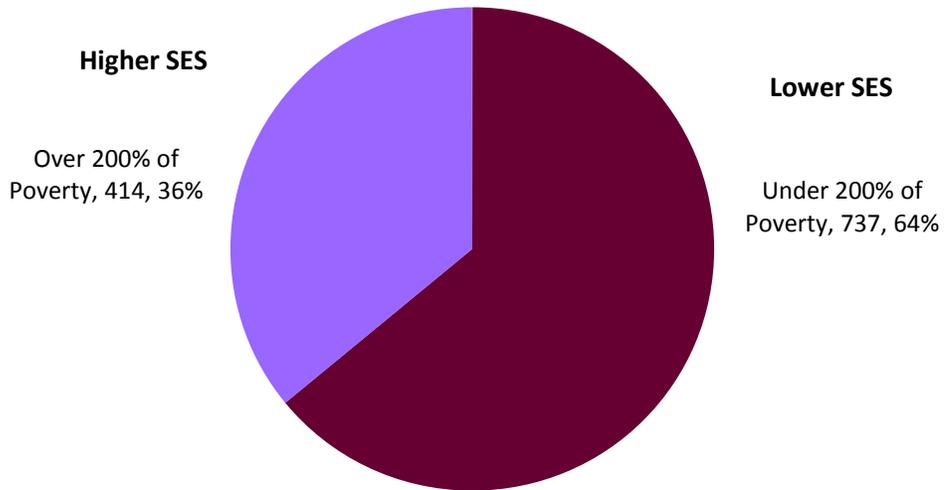
**Year 3 All Cohorts
Participant Children Grouped by Ethnicity - 1168 Total**



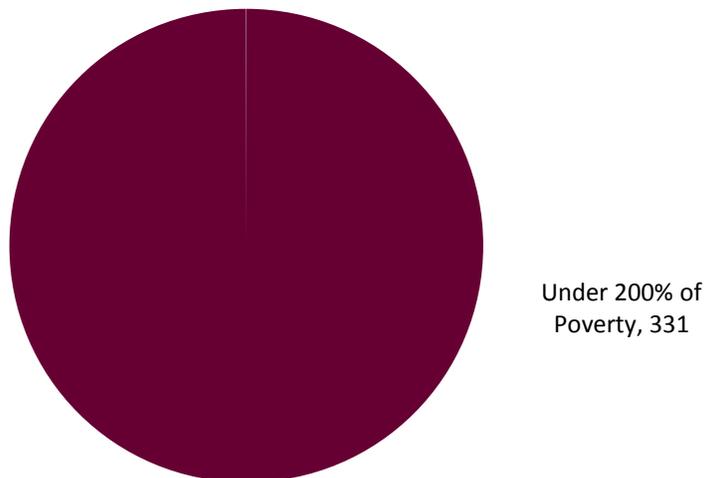
**Year 3 All Cohorts
Participant Children Receiving Equipment Grouped by
Ethnicity - 334 Total**



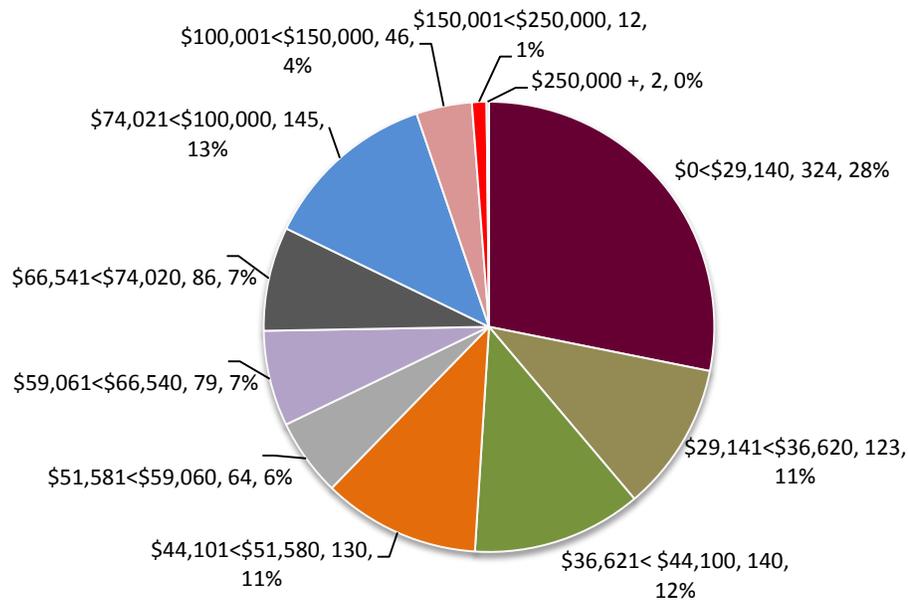
**Year 3 All Cohorts
Participant Families Grouped by SES
1151 Total**



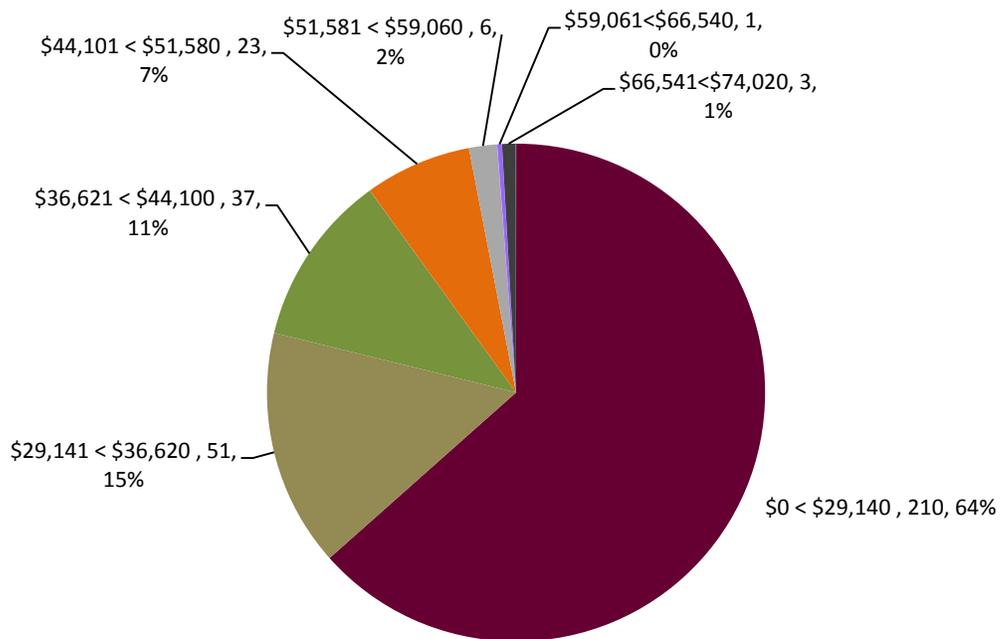
**Year 3 All Cohorts
Participant Families Receiving Equipment Grouped
by SES - 331 Total**



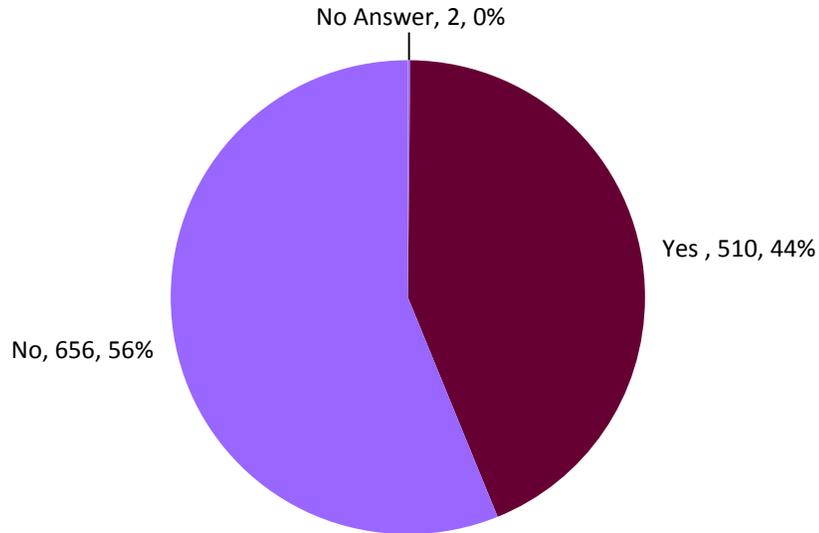
**Year 3 All Cohorts
Participant Families by Income Level - 1151 Total**



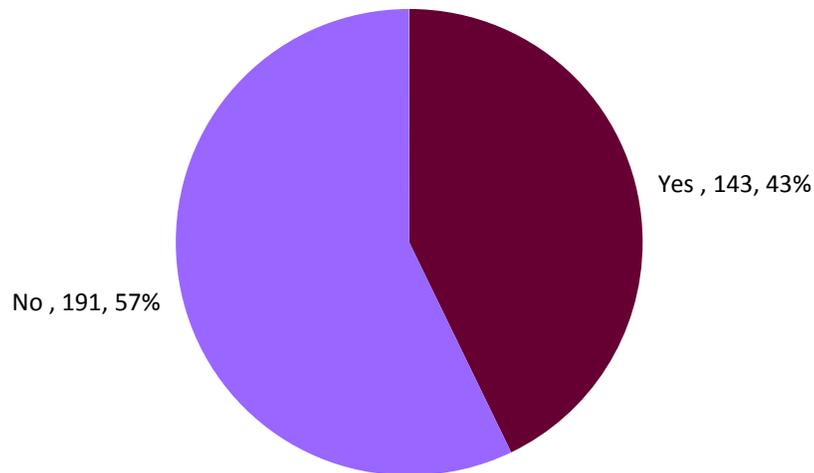
**Year 3 All Cohorts
Income Level for Participants Receiving Equipment - 331 Total**



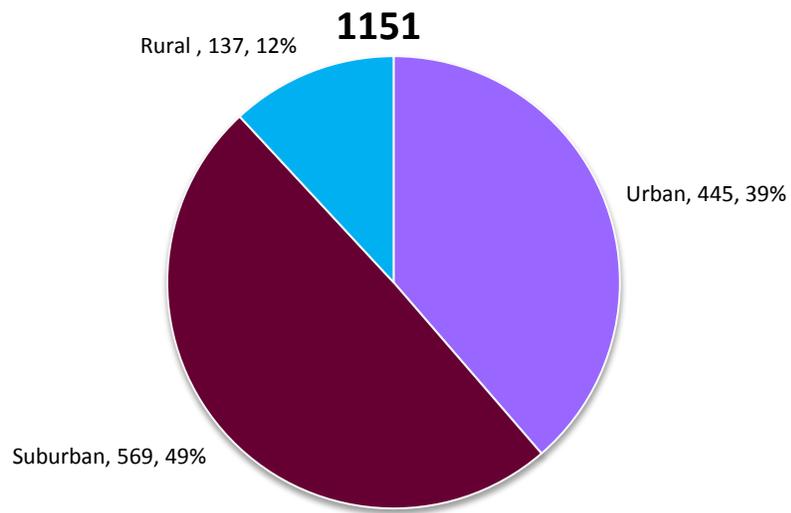
**Year 3 All Cohorts
Participant Children Grouped by Additional
Preschool - 1168 Total**



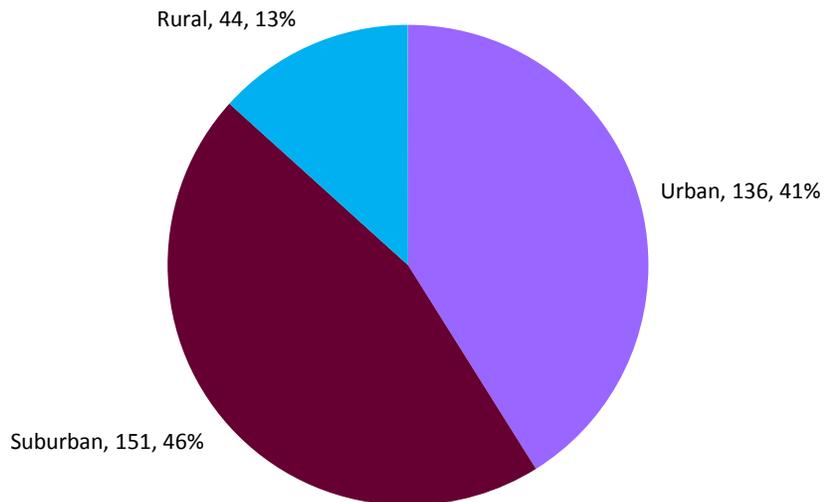
**Year 3 All Cohorts
Participant Children Grouped by Additional
Preschool and Receiving Equipment - 334 Total**



**Year 3 All Cohorts
Participant Families by Population Density - Total**



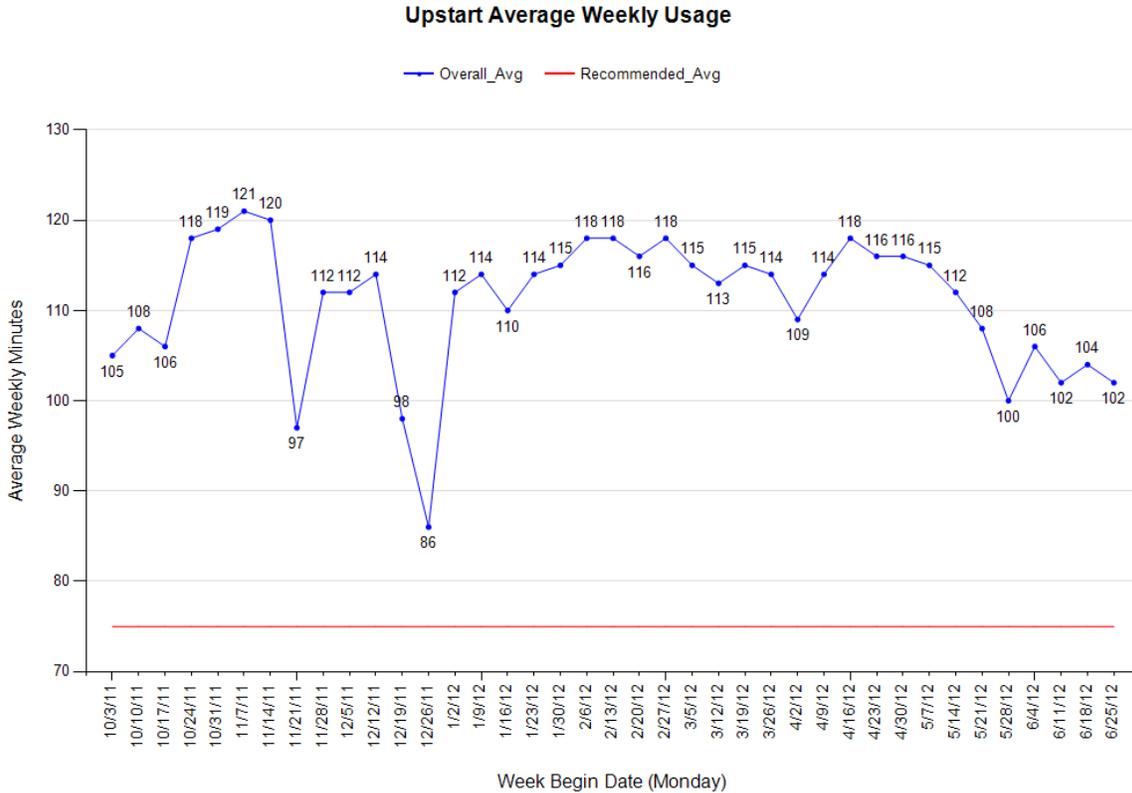
**Year 3 All Cohorts
Participant Families Grouped by Population Density
and Receiving Equipment - 331 Total**



Year 3 Results: Usage

Usage is the key to success for the UPSTART Program. When parents/caregivers enroll children in the program, they commit to 15 minutes of use on the reading program, five days a week, for a total of 75 minutes each week.

The average total Year 3 reading usage was 4,027. The average weekly reading usage was 103 minutes. No significant differences in usage were shown based on demographic data. Importantly, however, for the first time in the program, weekly usage *never* fell below the 75- minute minimum.



Year 3 Results: Assessment Data

Waterford uses the *Waterford Assessments of Core Skills™* (WACS) to assess children in the UPSTART Program. WACS is a computerized adaptive test of early literacy for students in pre-kindergarten through 2nd grade. Initial content validity for WACS was established against state and national standards for the 11 subtests:

- letter recognition
- letter sound
- initial sound recognition
- blending
- segmenting (an advanced skill not included in the UPSTART WACS test)
- reading real words
- reading nonsense words
- sight words

- reading comprehension
- listening comprehension
- vocabulary

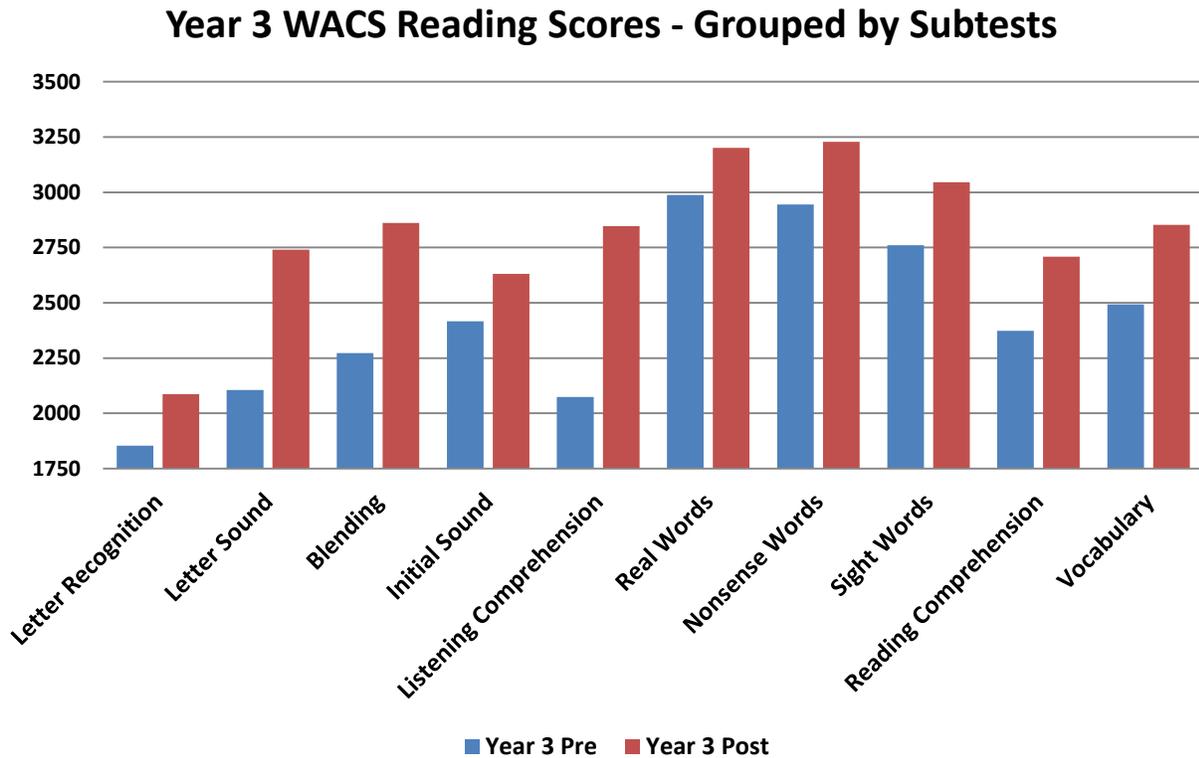
All items were calibrated for item response theory to determine item difficulty. To establish concurrent validity and predictive validity student performance on WACS was compared to performance on five commonly-used standardized tests also measuring early reading skills [DIBELS, Texas Primary Reading Indicator (TPRI), Idaho Reading Indicator (IRI); Iowa Tests of Basic Skills (ITBS), and Stanford Achievement Test Series (SAT 10)]. All correlations between tests are highly significant. Additional analyses indicate that WACS is internally coherent and has strong test-retest reliability.

WACS is an adaptive test. Pre-kindergarten students such as those in the UPSTART Program, only see the basic skills unless they perform well, in which case they also see advanced skills. When students take WACS for the first time, most of them receive basic skills only. As students use RRLWM and advance in their reading abilities, they perform better at the end of the program when taking WACS, and are able to successfully complete basic as well as advanced skills. Therefore, the number of students receiving advanced skills increases from the beginning of the program to the end of the program. Additionally, as more students complete reading comprehension successfully at the end of the program, fewer students complete listening comprehension at that time.

The following is the question difficulty ranges for WACS by grade. Note each grade is divided into thirds, for example, kindergarten beginning, kindergarten intermediate, and kindergarten advanced.

Grade	Beginning	Intermediate	Advanced
Preschool	1001 - 1333	1334 - 1666	1667 - 2000
K	2001 - 2333	2334 - 2666	2667 - 3000
1	3001 - 3333	3334 - 3666	3667 - 4000
2	4001 - 4333	4334 - 4666	4667 - 5000
3	5001 - 5333	5334 - 5666	5667 - 6000
4	6001 - 6333	6334 - 6666	6667 - 7000

Year 3 pre- and post-test scores for the ten subtests are shown in the following graph.

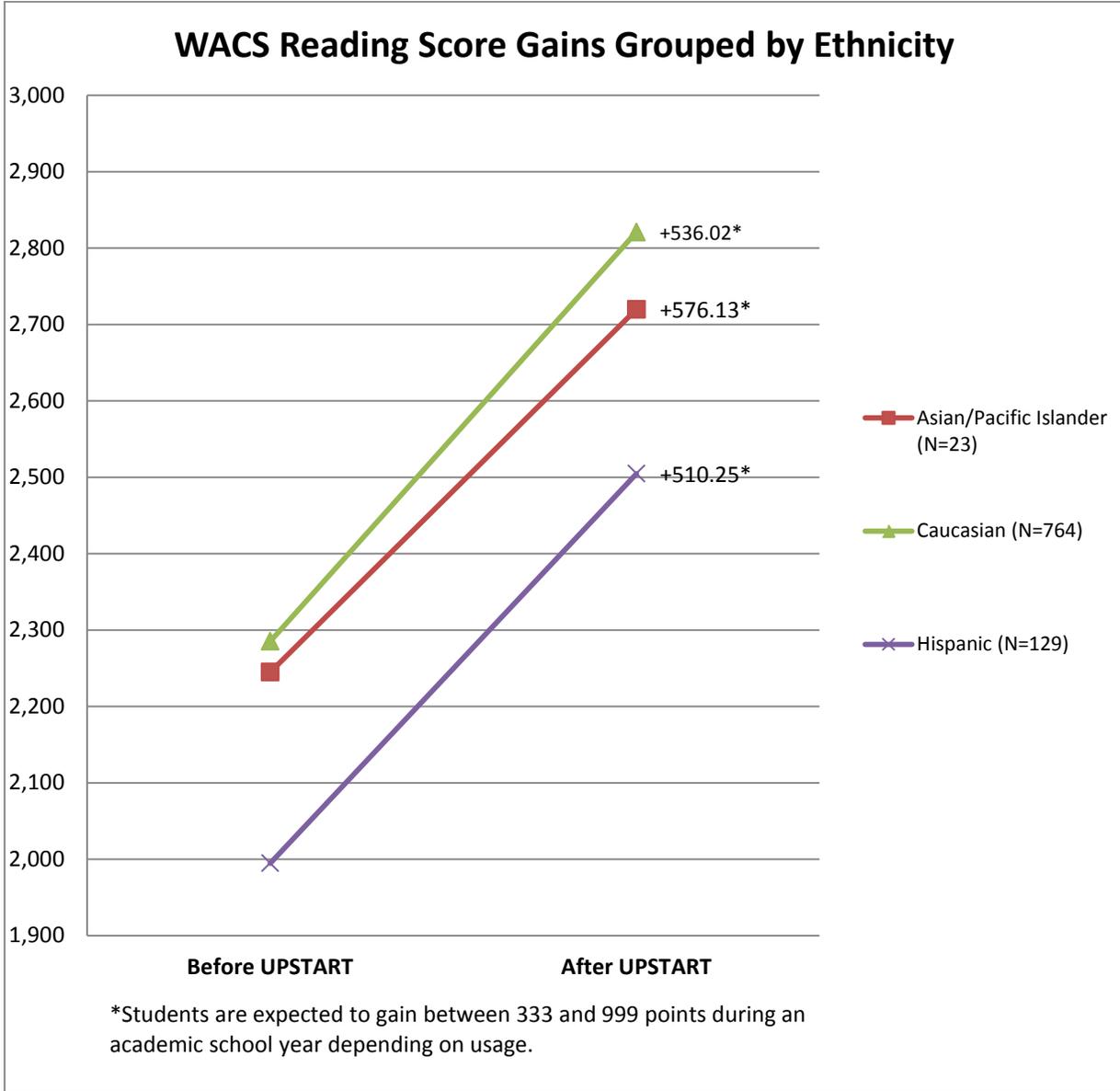


The following table gives both the numerical score and a grade-level equivalent to indicate participants' end-of-program performance.

Measured Individual Skill	Final WACS Score	Grade Equivalent
Letter Recognition	2087	K – Beginning (1)
Letter Sound	2740	K – Advanced
Blending	2861	K – Advanced
Initial Sound	2631	K – Intermediate
Listening Comprehension	2847	K – Advanced
Real Words	3200	1 – Beginning
Nonsense Words	3228	1 – Intermediate
Sight Words	3045	1 – Beginning
Reading Comprehension	2709	K – Advanced
Vocabulary	2852	K – Advanced
Overall	2773	K – Advanced

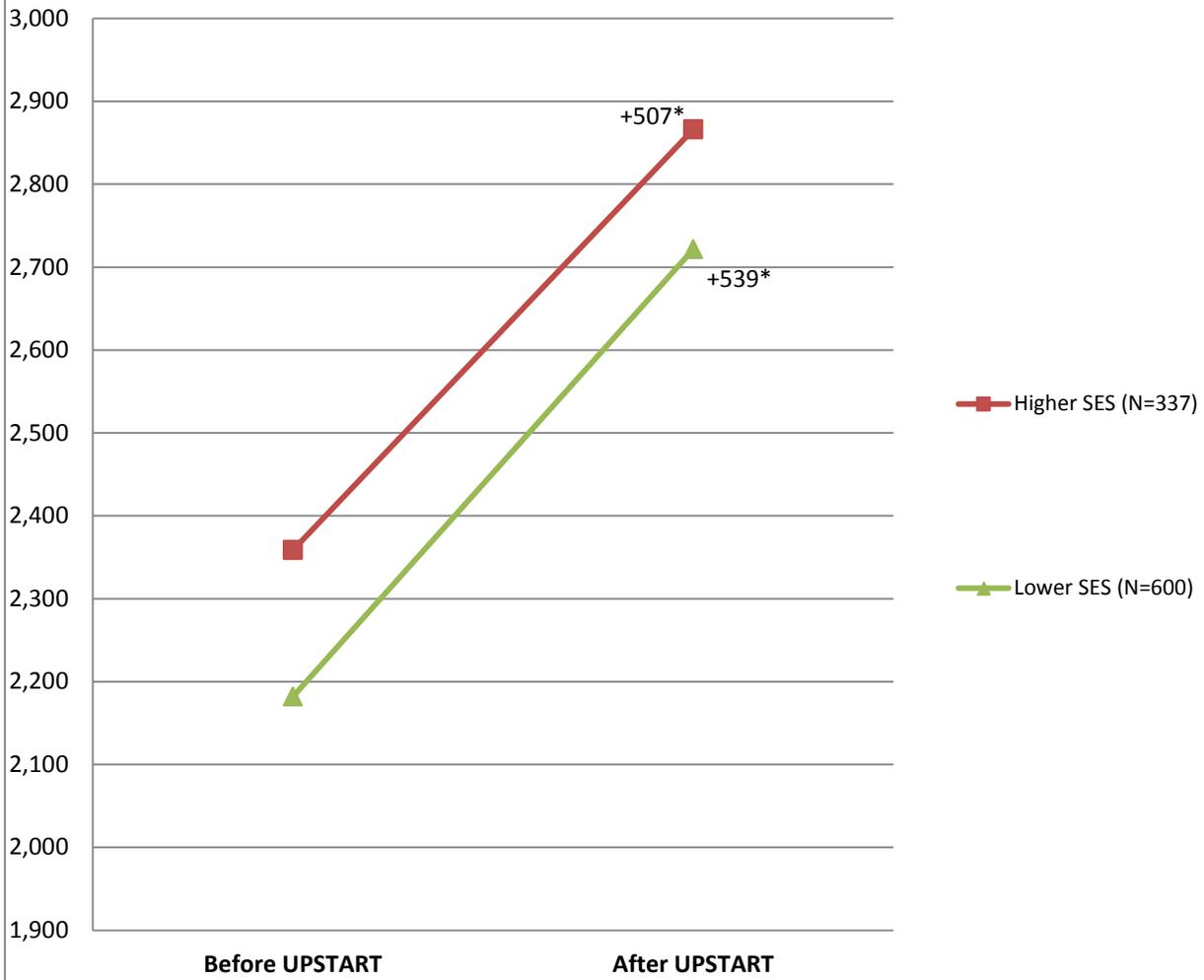
(1) Kindergarten Beginning is the ceiling for Letter Recognition

The following graphs show breakdowns of WACS data by ethnicity, SES, and preschool attendance.



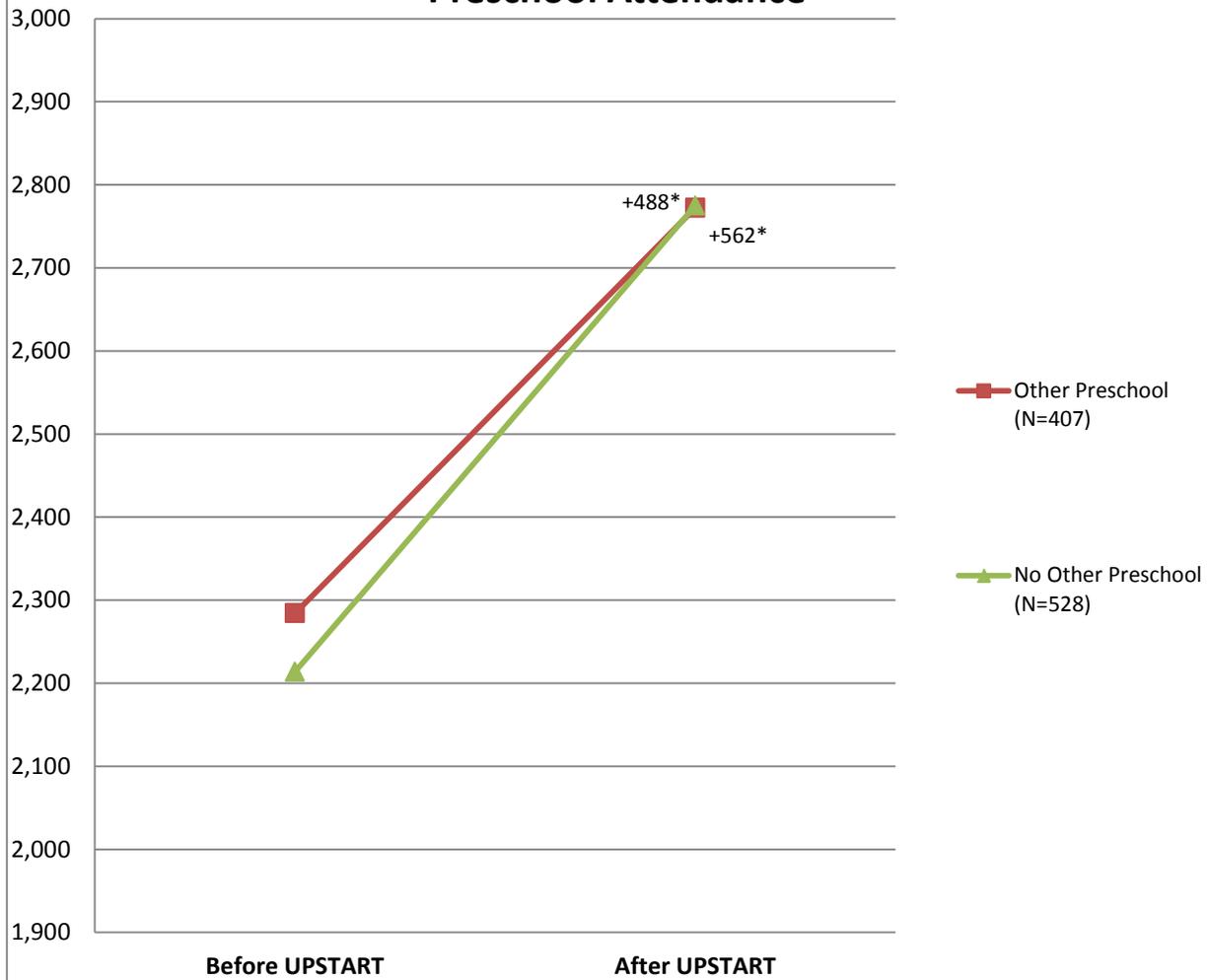
Note: Due to low N (participant count for this category) some ethnicities are not represented in the above graph.

WACS Reading Score Gains Grouped by SES



*Students are expected to gain between 333 and 999 points during an academic school year depending on usage.

WACS Reading Score Gains Grouped by Other Preschool Attendance



*Students are expected to gain between 333 and 999 points during an academic school year depending on usage.

The following table shows WACS data by household income.

WACS Reading Score Gains Grouped by Household Income						
Group	Pretest		Posttest		Gain	Final Grade Level
	N	Score	N	Score		
Less Than \$29,140	276	2123.04	233	2624.76	501.72	K Intermediate
\$29,140 to \$36,619	121	2184.86	109	2771.62	586.76	K Advanced
\$36,620 to \$44,099	136	2238.14	126	2807.92	569.78	K Advanced
\$44,100 to \$51,579	130	2255.58	115	2822.44	566.86	K Advanced
\$51,580 to \$59,059	61	2279.93	52	2814.33	534.4	K Advanced
\$59,060 to \$66,539	74	2264.09	61	2773.54	509.45	K Advanced
\$66,540 to \$74,019	79	2378.43	69	2831.22	452.79	K Advanced
\$74,020 to \$99,999	129	2349.57	122	2854.97	505.4	K Advanced
\$100,000 to \$149,999	42	2475.86	40	2915.78	439.92	K Advanced
\$150,000 to \$249,999	10	2471.1	9	3147.67	676.57	1 Beginning

Math, Science, and Reading Intervention

By design and direction from the UPSTART Advisory Committee, UPSTART focuses first on reading, as described above in the usage and assessment data sections. However, UPSTART participants also have access to RRLWM’s math and science curriculum. Year 3 participating children averaged 1,484 minutes on math and science. The math and science minutes are *in addition* to the usage minutes reported above for reading. Science and math are not assessed as part of the UPSTART program.

UPSTART also includes the multi-sensory intervention program *Camp Consonant*[™]. During Year 3, 136 UPSTART children used the program. Based on usage and progress in the core curriculum, Waterford staff members consulted individually with parents or caregivers to discuss their observations of their child’s learning styles and needs. Waterford staff members then described the multi-sensory approach used in *Camp Consonant*[™] and gave parents and caregivers the choice to move the children to *Camp Consonant*[™] or continue in RRLWM, the UPSTART core curriculum.

Training and Testing

During Year 3 Waterford presented 82 training and testing sessions across the state (Brigham City, Logan, Centerville/Layton, Cedar City, Provo/Orem, Ogden, Vernal, St. George, Salt Lake City, Blanding/Montezuma Creek, Centro de la Familia Providence, Centro de la Familia Honeyville, and Ibapa) to parents/caregivers and participants. Sixty four “graduations” (which included final testing) occurred during July and August 2012.

Survey of Participating Parents and Caregivers

Waterford surveyed parents and caregivers at the end of Year 3, and responses are detailed below.

Overall, do you feel participating in UPSTART was beneficial?			
		Response Percent	Response Count
Yes		99.8%	950
No		0.2%	2
Please explain:			815
answered question			952
skipped question			2

Do you feel UPSTART helped prepare your child for kindergarten?			
		Response Percent	Response Count
Yes		99.8%	939
No		0.4%	4
Please explain:			799
answered question			941
skipped question			13

Did you find it difficult to get your child to use the program?

		Response Percent	Response Count
Yes		44.3%	393
No		64.6%	573

How did you motivate your child to use the program? 891

answered question	887
skipped question	67

Were the representatives friendly?

		Response Percent	Response Count
Yes		99.9%	940
No		0.1%	1

answered question	941
skipped question	13

Were the representatives knowledgeable?

		Response Percent	Response Count
Yes		99.1%	933
No		1.0%	9

answered question	941
skipped question	13

Would you recommend the program to family member or friends?

		Response Percent	Response Count
Yes		99.6%	935
No		0.4%	4
answered question			939
skipped question			15

If you had any issues, did you feel that they were resolved well?

		Response Percent	Response Count
Yes		96.0%	881
No		4.8%	44
answered question			918
skipped question			36

Did you feel like the representatives wanted to help you succeed?

		Response Percent	Response Count
Yes		99.7%	928
No		0.5%	5
answered question			931
skipped question			23

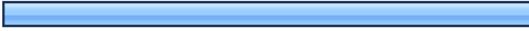
In general, did your child like / enjoy the software?

		Response Percent	Response Count
Yes		98.6%	911
No		1.9%	18
answered question			924
skipped question			30

In general, did you like the software?

		Response Percent	Response Count
Yes		99.6%	923
No		0.5%	5
answered question			927
skipped question			27

Overall, was the software too easy, too hard, or just right?

		Response Percent	Response Count
Too easy		8.9%	81
Too hard		3.0%	27
Just right		91.8%	833
answered question			907
skipped question			47

Was the content appropriate for your child's age?

		Response Percent	Response Count
Yes		99.0%	903
No		1.4%	13
answered question			912
skipped question			42

Do you feel the usage requirements are too strict?

		Response Percent	Response Count
Yes		12.6%	115
No		88.2%	805
answered question			913
skipped question			41

Was the purpose of consistent usage explained to you?

		Response Percent	Response Count
Yes		97.8%	888
No		2.2%	20
answered question			908
skipped question			46

Which of the following motivational tools enhanced your experience with the program? Check the appropriate boxes below:

	Very helpful	Somewhat helpful	Not helpful	Didn't know about it	Response Count
Kids Corner	22.4% (187)	33.1% (277)	10.5% (88)	34.1% (285)	836
Vocabulary word of the week	30.2% (261)	43.3% (374)	19.2% (166)	7.6% (66)	864
Weekly usage e-mail	77.2% (690)	18.7% (167)	2.8% (25)	1.6% (14)	894
Achievement chart	62.5% (554)	24.7% (219)	7.1% (63)	6.2% (55)	887
Contests and incentives	45.6% (405)	35.4% (314)	15.1% (134)	4.1% (36)	888
Conversations with user support representatives	54.8% (484)	36.0% (318)	6.0% (53)	3.6% (32)	883
answered question					918
skipped question					36

Cost

The cost-per-participant for UPSTART Year 3 was \$1,276. That number includes providing computers and Internet connectivity to qualifying families in the program without equipment and access.

At the end of Year 3, equipment inventory indicated that all but 14 participants had returned their equipment, for a loss rate of 1.2% (note, Waterford will be making a final effort to secure unreturned equipment in mid-September).

UPSTART Outreach Activities

The UPSTART Advisory Committee and Waterford are dedicated to reaching as many at-risk children as possible with the UPSTART Program. To that end, during the past year, Waterford was involved in the following outreach activities to advise parents and caregivers of the availability of UPSTART.

Online Advertising

City Grid

KSL.com

Supermedia.com

Newspapers (articles and purchased advertising space)

El Observador de Utah (article about UPSTART)

Deseret News (article about UPSTART)

Millard County Chronicle, Millard

Tintic Newsletter, Tintic

Nephi Times, Juab

Sanpete Messenger, North and South Sanpete

Wayne County Insider, Sevier and Wayne Counties

Emery County Progress, Emery

Meetings

Benjamin Wynn, Centro de la Familia, Honeyville

Boys and Girls Club of Salt Lake City

Catholic Archdiocese

Catholic Community Services

Chris Brent, RUCD Head Start, Blanding

Connexiones Familiares (Family Connections)

Family Counseling Service of Northern Utah

Flor Estrada, Centro de la Familia, Providence

Guadalupe Schools

Housing Authority of Salt Lake City

Intermountain Catholic

International Rescue Committee

Kathy Carnes, Centro de la Familia, Centerfield

Lynette Mitchell, Executive Director, RUCD Head Start

Mark Burge, Principal Montezuma Creek Elementary

Matt Bozada, Centro de la Familia, Genola

Ogden Head Start

Ogden Preparatory Academy
Saints Peter and Paul Catholic Church
Rancho Market
Rudy Anderson, Utah State Head Start Collaboration
Rural School District Association
Salt Lake CAP Head Start
Salt Lake CASA Program
Salt Lake City Mayor's Early Childhood Learning and Literacy Action Team
Smithfield City Council
Utah Hispanic Chamber of Commerce
Viva Market
WIC, North Salt Lake
William Greer, Workforce Development
YWCA

Mailings

Manufacturers business mailing (about 1,300 pieces)
Social welfare organizations, health clinics, churches (about 1,700 pieces)
All Utah school district superintendents (districts that asked for fliers are listed below)

Additional Follow Up with Interested School Districts

Alpine
Box Elder
Carbon
Duchesne
Emery
Iron
Kane
Logan
North Summit
San Juan
Tintic
Washington
Wayne

Events

Brigham City Peach Days
Charter School Convention
Cinco de Mayo
Salt Lake City Indian Walk-in Center, Utah Indian Child Welfare Conference
Junior League CARE Fair
Mom's Club Preschool Exhibit
People Helping People – Single Mother's Employment Fair
Utah Home Education Association

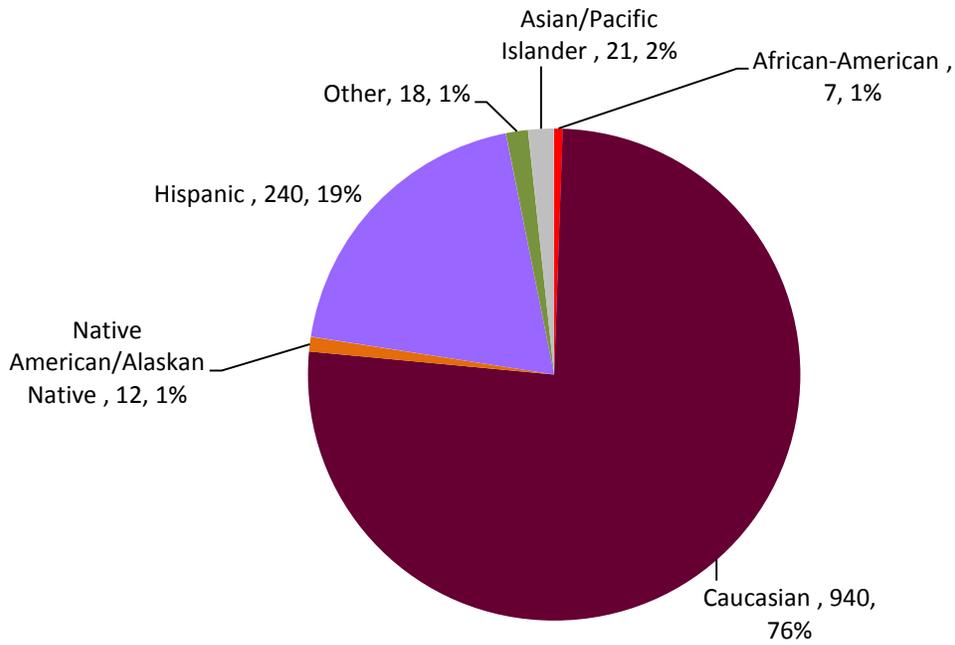
Year 4

Year 4 of the UPSTART Program is currently underway, with 1,212 families and 1,238 children participating. Unfortunately, an additional 586 families were wait-listed for the program and, in the end, could not be served. Another 230 families applied for the program after the wait list was identified. All of the wait-listed and other families were notified of their status and given the opportunity to purchase the program, pursuant to the original UPSTART enabling legislation.

Seventy training sessions were provided across the state (Ephraim was added as a training site) during August and the first week of September. Of participating families, 870 received program drives, and the remainder received Internet service, computers, or a combination of computers and Internet service.

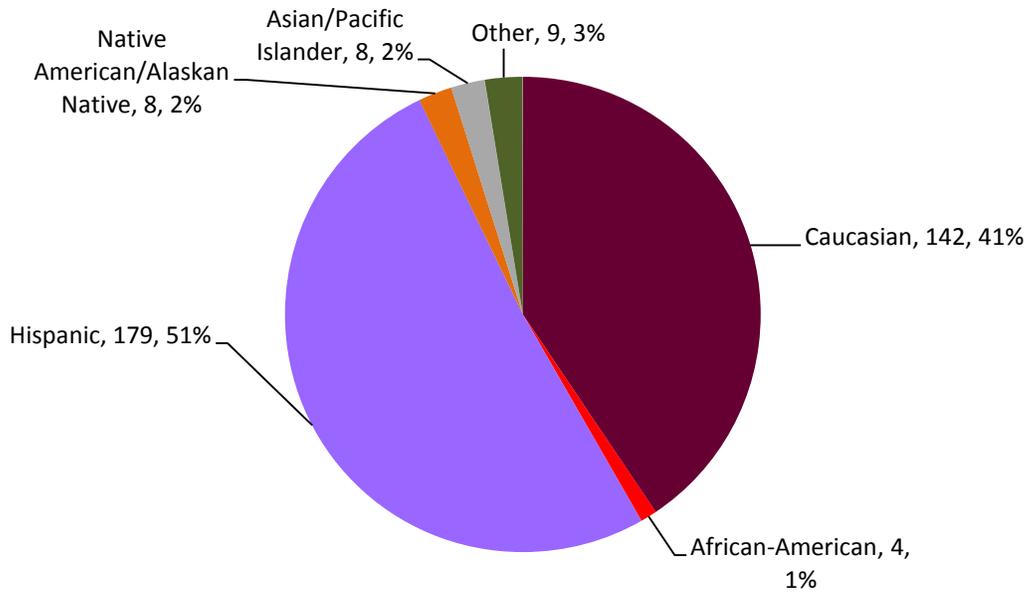
The following charts show diversity, SES, income level, “other” preschool, and population density information for all participants and the subset receiving equipment.

Year 4 - Cohorts 1A & 1B Ethnicity of Participant Children - 1238 Total

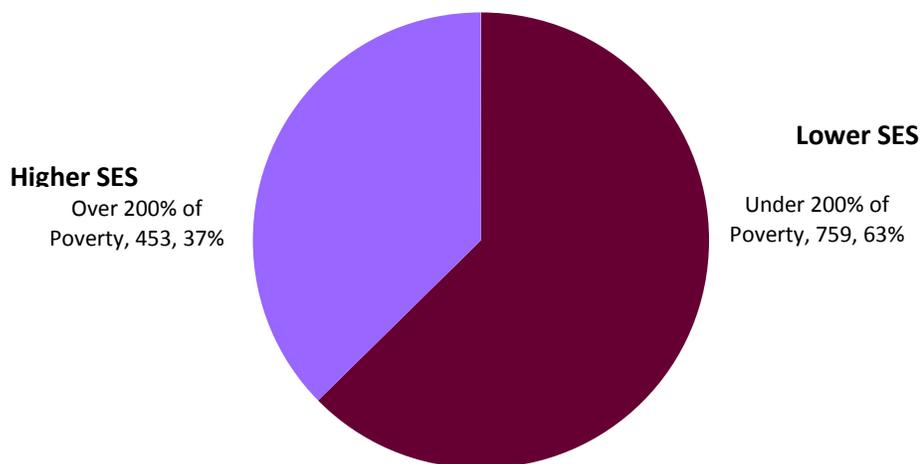


Year 4 - Cohorts 1A & 1B Ethnicity for Participant Children Receiving Equipment - 350 Total

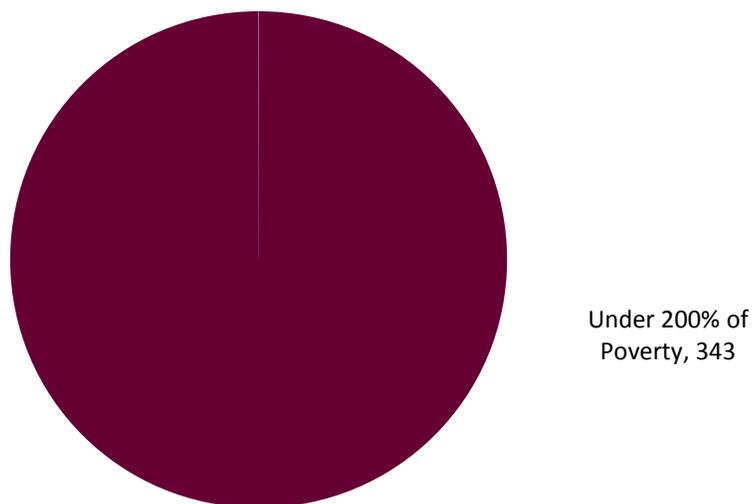
Receiving Equipment - 350 Total



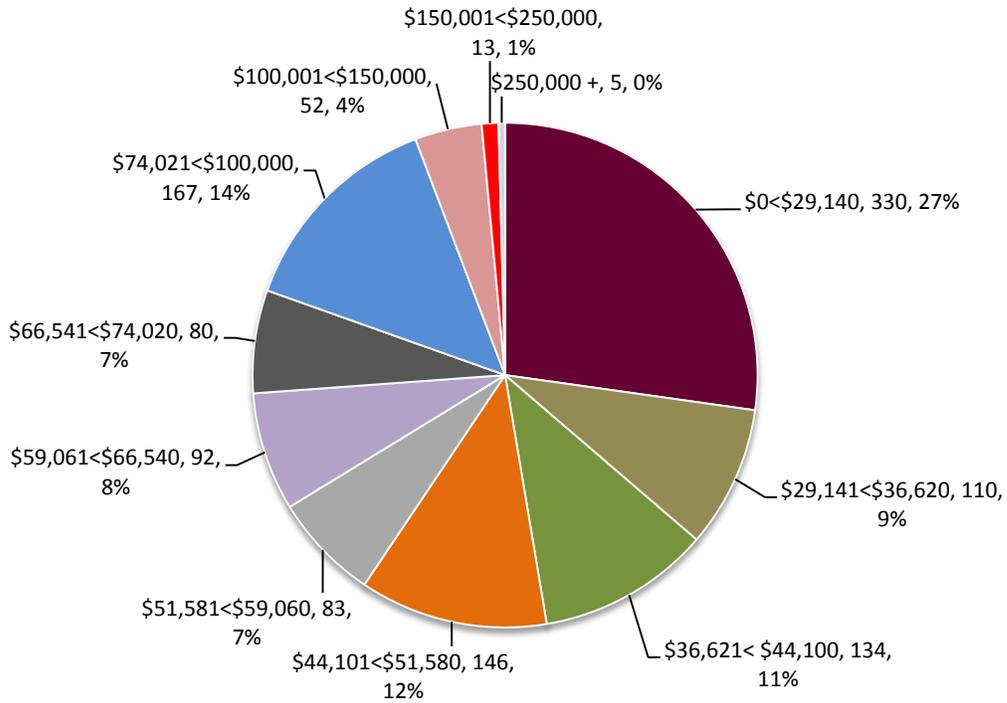
Year 4 - Cohorts 1A & 1B SES for Participant Families - 1212 Total



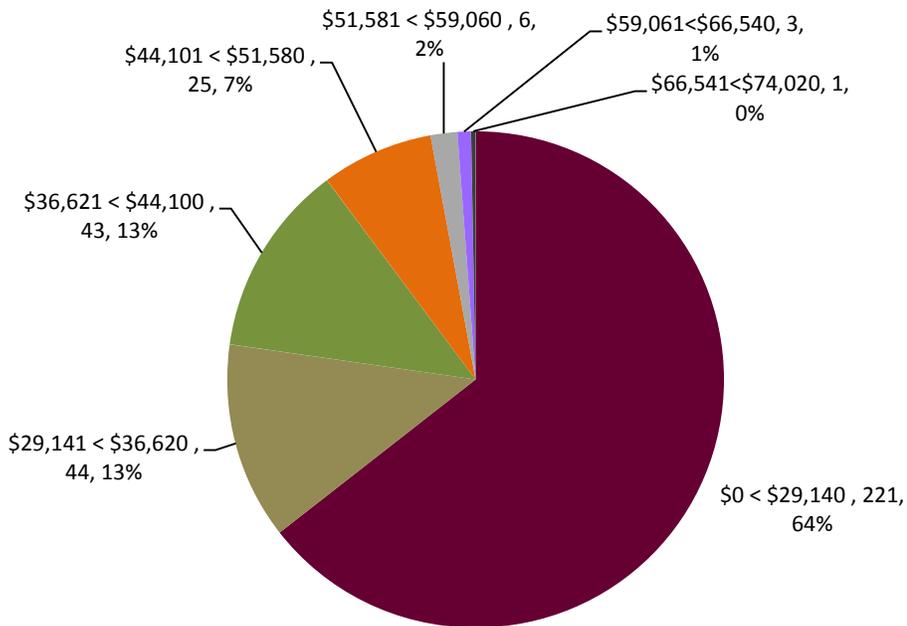
Year 4 All Cohorts Participant Families Receiving Equipment Grouped by SES - 343 Total



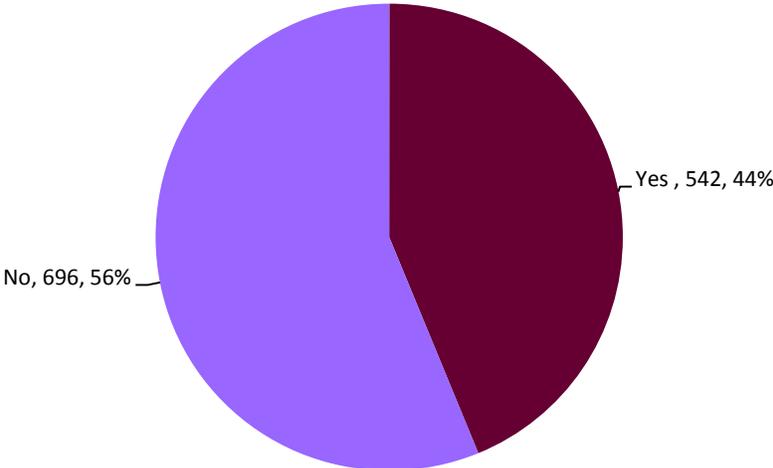
Year 4 - Cohorts 1A & 1B Participant Families Income Level - 1212 Total



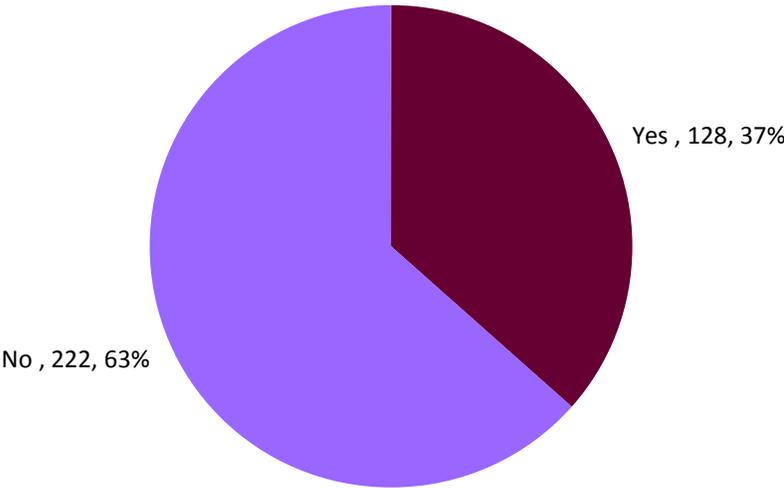
Year 4 - Cohorts 1A & 1B Income Level for Participant Families Receiving Equipment - 343 Total



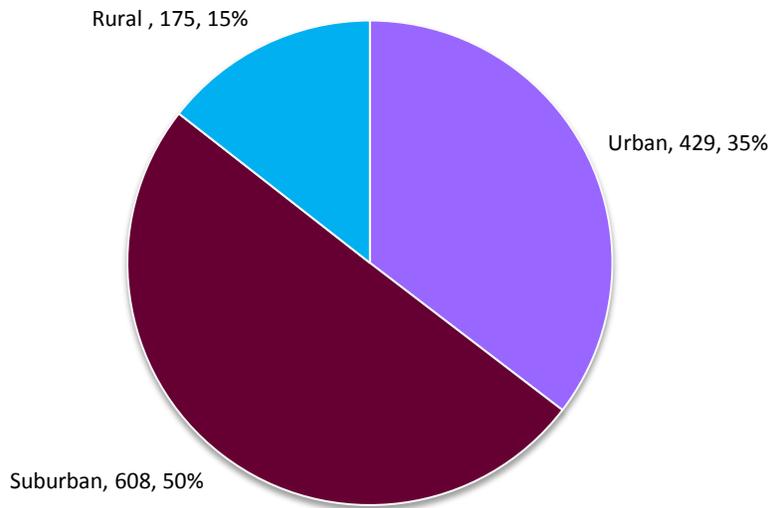
Year 4 - Cohorts 1A & 1B Additional Preschool for Participant Children - 1238 Total



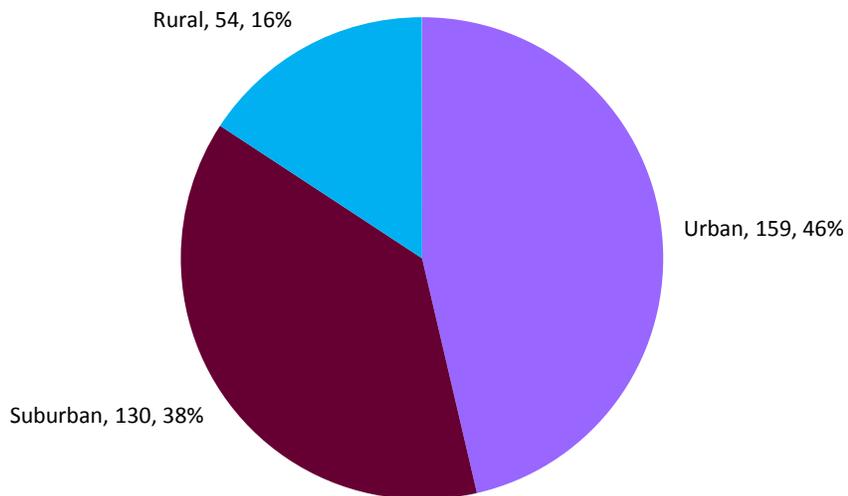
Year 4 - Cohorts 1A & 1B Additional Preschool for Participant Children Receiving Equipment - 350 Total



Year 4 - Cohorts 1A & 1B Population Density of Participant Families - 1212 Total



Year 4 - Cohorts 1A & 1B Participant Families receiving equipment by Population Density - 343 Total



UPSTART Advisory Committee

The UPSTART Program continues to benefit greatly from its outstanding Advisory Committee. The committee meets quarterly to receive information about the program and make recommendations. Note the inclusion of three new members: Kathleen Petersen, Title I Director, Washington County School District; Jeana Stuart Swapp, Education Manager, Centro de la Familia de Utah; and Lynette Mitchell, Rural Utah Child Development Head Start. In addition to being outstanding educational practitioners, all three bring important and improved access for UPSTART to at-risk children throughout the state.

- Brenda Hales, Associate Superintendent for Student Achievement and School Success, Utah State Office of Education
- Tiffany Hall, K-12 Literacy Coordinator, Utah State Office of Education
- Christine Kearn, Education Director for Governor Gary Herbert's Office
- Nancy Livingston, Literacy Specialist
- Joanne Milner, Education Partnership Coordinator for Salt Lake City Mayor Ralph Becker's Office
- Lynette Mitchell, Executive Director, Rural Utah Child Development Head Start
- Kathleen Petersen, Title I Director, Washington County School District
- Dale Smith, Department Chair, Family and Human Studies, Salt Lake Community College
- Jeana Stuart Swapp, Education Manager, Centro de la Familia de Utah
- Cheryl Wright, Associate Professor, Department of Family and Consumer Studies, University of Utah
- Claudia Miner, UPSTART Program Director, Waterford Institute
- Dustin Heuston, ex officio, Chairman, Waterford Institute

UPSTART Program External Evaluation

In January 2012, Waterford received the Year 1 UPSTART Program external evaluation conducted by the Los-Angeles-based Evaluation and Testing Institute. The report in its entirety is appended to this report. Below is a brief overview of the Year 1 report.

The evaluation of UPSTART's first year of implementation was specifically designed to assess the program's impact on developing the children's reading proficiency once they enrolled in kindergarten. Tests data were obtained from 7 Utah public school districts for 137 children who had participated in UPSTART during its first year of operation during the 2009-2010 school year. Control group data were provided by school districts for all non-participating students tested. 247

non-participating kindergarten students were randomly selected for the control group. The test data used for analysis came from DIBELS NEXT (DN).

The Evaluation and Training Institute produced a 29-page report, and summary highlights follow.

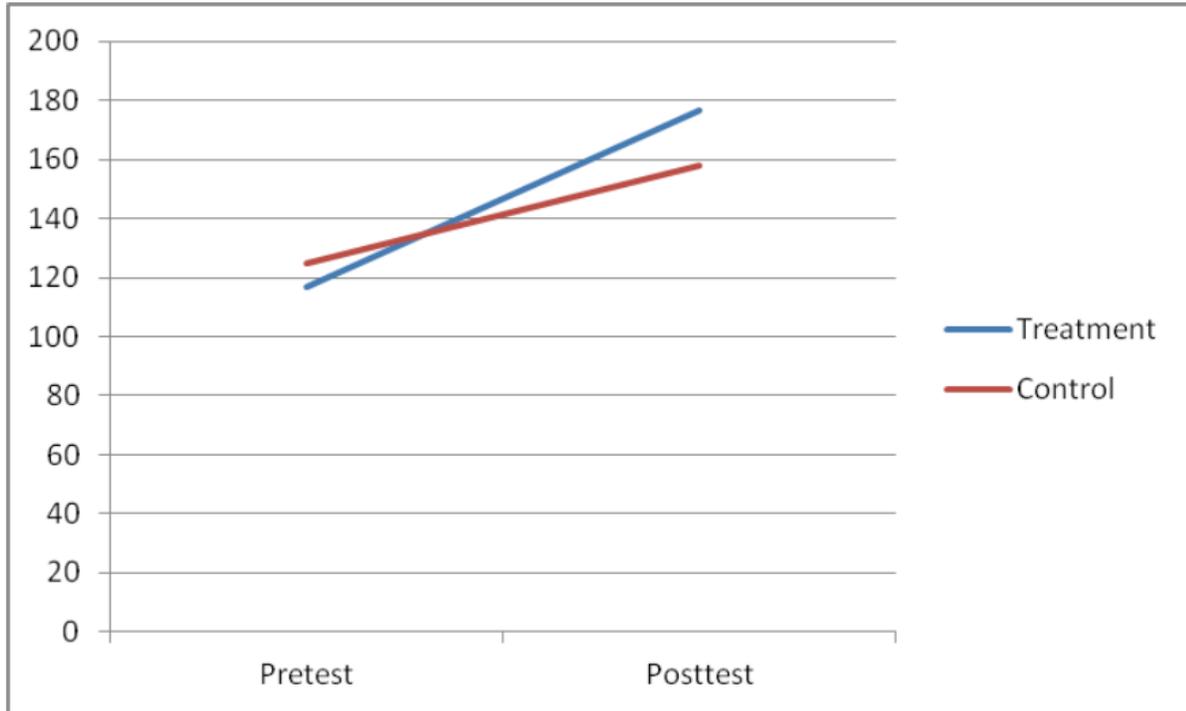
- We find that children who had participated in UPSTART during preschool scored almost 18 points higher in reading proficiency as measured by the DN Composite compared to beginning kindergarten children who did not participate in UPSTART prior to enrolling in public school.
 - *Based on these results, the evidence is that UPSTART appears to significantly improve reading readiness for beginning kindergarten students* (italics by ETI).
- We find that middle kindergarten children who had participated in the UPSTART preschool program scored approximately 19 points higher in reading proficiency as measured by the DN Composite for middle of kindergarten compared to middle kindergarten children who did not participate in UPSTART prior to enrolling in public school.
 - Based on this evidence, we can say that the *UPSTART participants appear to have sustained their gains in reading proficiency through the middle of kindergarten* (italics by ETI).
- Based on the first 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.

Waterford received the Year 2 external evaluation in March 2012. The report in its entirety is appended to this report. The following are highlights of the Year 2 report.

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. The children were measured on two tests of early literacy skills: the Brigance Inventory of Early Development and the Bader Reading and Language Assessment. The overall conclusion of the report is:

- “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.”
- **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. The following figure 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

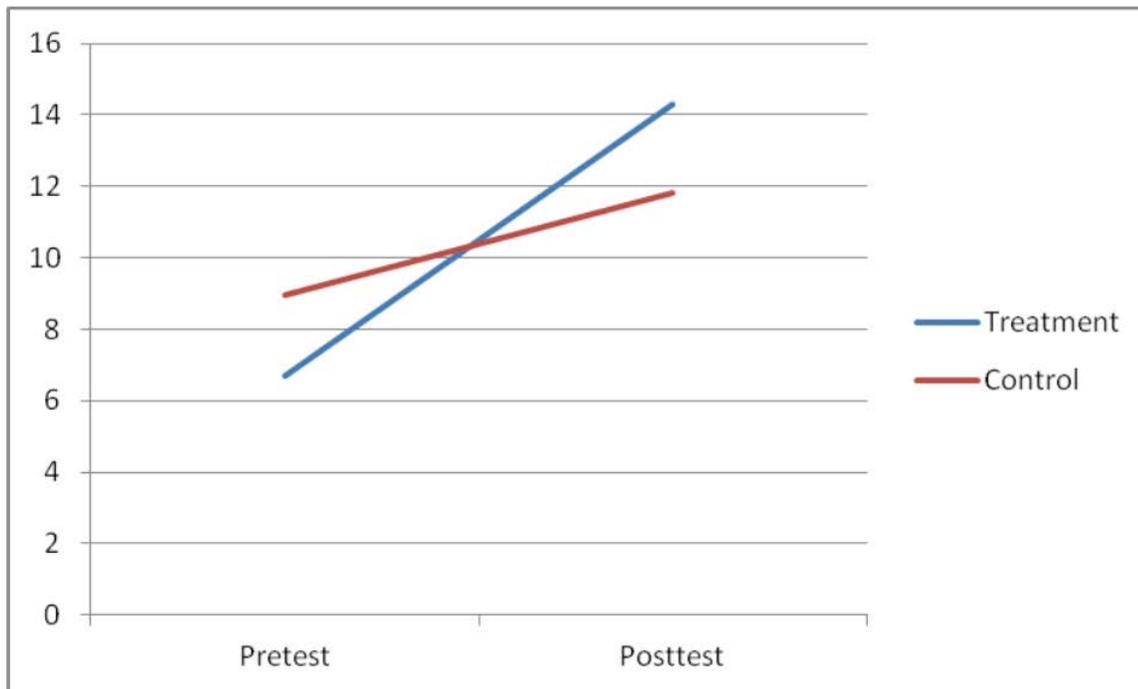
phonics skills—particularly in the ability to produce sounds of lower case letters—compared to the control group.



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. The figure below 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.



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

- 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 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 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 *Phoneme Blending* subtests compared to control group children.

DIBELS Scores

The Office of Legislative Research and General Counsel this year requested “student performance on pre-Kindergarten and post-kindergarten assessments conducted by school districts and charter schools for students who participated in the home-based educational technology program and those

who did not participate in the program.” Waterford does not receive that information from the Utah State Office of Education. However, the external evaluation for Year 1 did include pre-K and post-K DIBELS results as part of the study. Those results were reported as follows:

[W]e find that children who had participated in UPSTART during preschool scored almost 18 points higher in reading proficiency as measured by the DN [DIBELS Next] Composite compared to beginning kindergarten children who did not participate in UPSTART prior to enrolling in public school.

[W]e find that middle kindergarten children who had participated in the UPSTART preschool program scored approximately 19 points higher in reading proficiency as measured by the DN Composite for middle of kindergarten compared to middle kindergarten children who did not participate in UPSTART prior to enrolling in public school.

Contact Information for Questions

This report was prepared by the following Waterford Institute staff members:

- Dr. Claudia Miner, Vice President of Development and UPSTART Program Director (claudiaminer@waterford.org)
- Dr. Haya Shamir, Vice President of Applied Research and Learning and Chief Scientist (hayashamir@waterford.org)
- Ann Izzo, Director of User Support (annizzo@waterford.org)

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

Cohort 1/Year 1 Results Technical Report

August 2011

Submitted by:



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

All correspondence should be sent to:

Jon Hobbs, Ph.D.
jhobbs@eticonsulting.org

Contents

EXECUTIVE SUMMARY	1
INTRODUCTION	3
BACKGROUND	3
EVALUATION DESIGN	4
Research Questions	4
Outcome Measures	6
Test Data Collected	6
KINDERGARTEN ANALYSIS	8
Exploratory Analyses	9
Variable Transformations	9
OLS Regression Analysis	9
Diagnosing Heteroskedasticity	10
Weighted Least Squares Analysis	10
Analysis of Implementation Time	10
UPSTART IMPLEMENTATION	12
UPSTART Enrollment	12
UPSTART Equipment Provided	12
UPSTART Graduates	13
UPSTART Usage	14
How UPSTART Usage Relates to Reading Proficiency	17
UPSTART OUTCOMES	21
Does UPSTART improve reading readiness?	21
Does UPSTART sustain improvements in reading readiness?	22
SUMMARY AND CONCLUSIONS	25
Test Data Collected and Analyzed	25
The Analysis	25
Findings: UPSTART Implementation	26
Findings: UPSTART Impact on Reading Proficiency in Kindergarten	27

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 (56%) of the 1,248 preschool children enrolled in the first year of UPSTART were from low-income families. The evaluation of UPSTART's first year of implementation was specifically designed to assess the program's impact on developing the children's reading proficiency once they enrolled in kindergarten. Other objectives included documenting the extent to which participants used the computerized curriculum; establishing the relationship between curriculum usage and reading proficiency outcomes; and documenting the program's completion or "graduation" rate.

The Kindergarten Analysis

A weighted least squares regression approach was used to estimate reading proficiency at the beginning of kindergarten as measured by the DIBLES Next (DN) beginning kindergarten composite. Low-income status was used as a weight variable to correct for non-constant variance in estimating the regression coefficients. English Language Learner (ELL) status, special education status, and gender were used as control variables in the final model to estimate reading proficiency for children who had participated in UPSTART the previous school year compared to children from the general kindergarten population who had not participated in UPSTART.

An ordinary least squares regression approach was used to estimate reading proficiency in middle kindergarten as measured by the DIBLES Next (DN) middle kindergarten composite. A hierarchical block design was used in the final model, which controlled for race, ELL status, low-income status, and special education status in examining differences in reading proficiency between the UPSTART treatment group and the kindergarten control group.

The effect of UPSTART usage on reading proficiency 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 kindergarten analysis sample. The final model statistically controlled for ELL status and special education status in kindergarten in comparing usage at each quartile with usage at the fourth quartile in estimating the effect of UPSTART on reading proficiency at the beginning and middle of kindergarten using the DN Composites.

Findings: UPSTART Implementation

Most of the first year UPSTART participants (70%) received a computer drive with the UPSTART curriculum loaded on it. Approximately 11% of the first year participants were loaned a computer and given free Internet access to help them access the UPSTART curriculum. Another 7% of the first year participants were loaned a personal computer to use at home while

participating in UPSTART. The remaining 10% to 12% of the first year participants were provided with various combinations of educational technology to enable them to access the UPSTART curriculum, including wireless and cellular devices.

- The kindergarten test sample had a mean of 68 hours of participation in the UPSTART curriculum over the first year of the project. This compares with an average of 76 hours of instruction for program “graduates” and an average of 53 hours of instruction for all students enrolled in UPSTART in the first year.
- Length of participation in the UPSTART curriculum was significantly and positively correlated with reading proficiency at the beginning ($r=.30$) and middle ($r=.29$) of kindergarten.
- Reading proficiency was shown to improve with increasing levels of UPSTART curriculum usage for both beginning and middle kindergarten children who had participated in UPSTART in preschool. These analyses controlled for ELL and special education status in kindergarten, both of which negatively impact reading proficiency outcomes.
- The UPSTART graduation rate in the first year of the program was 59%. UPSTART graduation status was not significantly correlated with reading proficiency at either the beginning ($r=.13$) or middle ($r=.10$) of kindergarten.

Findings: UPSTART Impact on Reading Proficiency in Kindergarten

- UPSTART participation had a moderately strong impact on improving the reading proficiency of UPSTART participants compared to nonparticipants at the beginning of kindergarten, as measured by the DIBELs Next Beginning Kindergarten Composite. The final model controlled for the significant influence of ELL status, special education status, and gender on reading proficiency outcomes at the beginning of kindergarten.
- UPSTART participants maintained their achievement gains through the middle of kindergarten as measured by the DIBELs Next Middle Kindergarten Composite. The final model controlled for the significant influence of ELL status, special education status, low income status and race on reading proficiency outcomes in middle kindergarten.

Based on the first 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 first year of operation during the 2009-10 school year, the project's implementation contractor – the Waterford Institute – enrolled 1,248 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 first year of implementation was designed specifically to assess the program's impact on developing the children's reading proficiency once they enrolled in kindergarten. Other objectives included documenting the extent to which participants used the computerized curriculum; establishing the relationship between curriculum usage and reading proficiency outcomes; and documenting the program's completion or "graduation" rate.

A majority (56%) of the 1,248 preschool children that enrolled in the first year of UPSTART were from low income families, according to data provided by the Waterford Institute. Slightly more boys (52%) were enrolled than girls (48%). In terms of ethnicity, the vast majority (81%) of the enrollment was Caucasian, 13% were Hispanic, 2% were of Asian descent, 1% were Black, and 1% were Native American. Ethnicity was unknown for 1% of the first year enrollment.

Background

The Utah State Department of Education's (USOE) quasi-experimental framework for the evaluation involved a variation of the Posttest-Only Design with Nonequivalent Groups. The USOE's scope of work suggested the need for repeated posttests in kindergarten and grade one. This design is sketched below where NR means "not randomly assigned," X stands for the implementation of the UPSTART treatment and O1, O2, and O3 stand for repeated posttests at the beginning and end of kindergarten and at the beginning of grade one.

NR	X	O1	O2	O3

NR		O1	O2	O3

In this research design, the UPSTART children received the Waterford Early Learning Program through the use of interactive personal computers at home in the year before kindergarten and are then compared with a group of children that did not participate in the UPSTART preschool program. The two main problems with this design are:

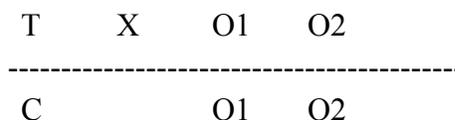
- **Selection bias:** We don't know if any observed posttest differences between the treatment and control group are due to pre-existing differences.
- **Mortality:** Even if the groups had been the same at the beginning, we don't know if any observed posttest differences are due to differential dropout rates of families from the study groups that changed the composition of the groups over time.

Thus, design limitations include not knowing whether selection effects or differential dropout rates account for any between group differences that might be observed after the preschool program. We proposed two things that could be done to improve this design: (1) use a pretest in fall 2009 with both treatment and control groups, or (2) use multiple control groups that bracket the expected effects. The USOE declined to permit the use of pretests and delays in obtaining data from the Waterford Institute and the Utah public schools derailed our plans to select a "criterion" (or All Star) control group.

Lacking the desired degree of control through experimental design, our fallback option in the first year evaluation was to exercise statistical control methods through the use of linear multiple regression analysis and analysis of covariance.

Evaluation Design

The kindergarten evaluation design that was implemented is a posttest-only repeated measures design with nonequivalent groups and is diagramed below, where T stands for children who received the UPSTART preschool program, and C stands for the general population Control group of students that 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 at the beginning of kindergarten, and O2 indicates measurements taken in the middle of kindergarten.



Research Questions

We hypothesized that if UPSTART has no effect on improving reading readiness, then the kindergarten children who participated in UPSTART – the treatment group – would be expected to perform at the same level as the general population control group on kindergarten measures of reading proficiency at the beginning of kindergarten. If UPSTART does have an effect on improving reading readiness, then the treatment group should perform significantly better than

the control group when first measured at the beginning of kindergarten. If UPSTART sustains its gains over time, then the treatment group would be expected to continue to perform better than the control group when retested in the middle of kindergarten.

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

RQ1: *Does UPSTART improve reading readiness?*

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

$$T > C @ O1$$

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

$$T = C @ O1$$

RQ2: *Does UPSTART sustain improvements in reading readiness?*

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

$$T > C @ O1 \text{ and } O2$$

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

$$T = C @ O2$$

In the kindergarten analysis, the outcomes of interest are measures of early literacy skills relevant to emerging readers such as early phonemic awareness, letter recognition, awareness of concepts of print and oral language comprehension.

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 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 curriculum¹ 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 Utah public schools in which the UPSTART participants enrolled for kindergarten in the fall of 2010 mostly used the latest (7th) edition of *The Dynamic Indicators of Basic Early Literacy Skills* (DIBELS), called the DIBELS Next, to measure student reading proficiency. We were able to obtain DIBELS Next reading proficiency data from seven Utah public school districts for 137 kindergarten students who had participated in UPSTART. The cooperating school districts also provided us with DIBELS Next data for over 9,000 kindergarten students who had not participated in UPSTART.

At the kindergarten level, the DIBELS Next (DN) primarily measures phonemic awareness and early phonics skills. We used DN Composite Scores to measure student reading proficiency. The DN Composite for beginning kindergarten is the total score obtained by summing the scores of the First Sound Fluency (FSF) subtest and the Letter Naming Fluency (LNF) subtest. The DN Composite for middle kindergarten is the total score obtained from summing the scores of the FSF subtest, LNF subtest, Phoneme Segmentation Fluency (PSF) subtest, and the Nonsense Word Fluency Subtest. The latter measures a child's competency with the alphabetic principle and with early phonics skills.

Test Data Collected

Ten Utah school districts supplied test data for 258 children enrolled in public kindergartens in 2010 who had participated in UPSTART the previous school year. One school district submitted Developmental Reading Assessment (DRA) data for 62 students. Two school districts submitted DIBELS test data for 59 students. Seven school districts supplied DIBELS Next (DN) test data for 137 students.

¹ 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.

Only the DN data were sufficient for the desired statistical analysis. Consequently, the kindergarten analysis was carried out with DN data using a sample of 137 kindergarten students who had participated in the UPSTART preschool program the previous year and a control group of 247 kindergarten students who had not participated in UPSTART. Control students were randomly selected from the pool of DN test data at each of the seven participating school districts (Box Elder, Granite, Iron, Ogden, Salt Lake, Tooele, and Weber) using a 2:1 ratio of control to treatment group students in order to optimize sample size for the analysis.

See Appendix A for a summary of the statistical power analysis that was conducted to inform the sampling plan. See also Appendix B for further details on the samples.

Kindergarten Analysis

A kindergarten student data file was developed based on data collected from the Waterford Institute and from ten Utah public school districts that agreed to supply data for the first year evaluation. A summary of the variables examined in the DIBELS Next analysis is shown in Table 1. The DIBELS Next composite scores for the beginning and middle of kindergarten were chosen as the dependent variables for the analysis since they provide the best overall estimate of the student’s reading proficiency, according to the DIBELS Next Technical Manual.

*Table 1
Variables in the Kindergarten Analysis of Reading Readiness*

Predictor Variables	Data Values
Group	Treatment vs. Control
District	School District
Gender	Male vs. Female
Ethnicity	Asian Black/African American Caucasian/White Hispanic American Indian/Alaska Native Hawaiian/Pacific Islander Multi-Ethnic Other Unknown
Low Income (Free or Reduced Price Lunch)	Yes/No
English Language Learner (ELL)	Yes/No
Primary Language Spoken	English Spanish Other
Attended Any Preschool	Yes/No
Title 1 Compensatory Education	Yes/No
Special Education	Yes/No
Birth Date	Mm/dd/yyyy format
UPSTART “Graduate”	Yes/No
Usage	UPSTART Minutes Logged
UPSTART Provided Equipment	9 categories of equipment
Entry Date	Date enrolled in UPSTART
UPSTART Usage by week	Minutes per week
DIBELS Next Outcome Variables	Data Values
Composite Score Beginning K	0 to High
Composite Score Middle K	0 to High

Exploratory Analyses

The treatment group and control group data were examined separately using descriptive statistics and the SPSS Explore procedure to describe the two groups on the variables assessed and to check the assumptions underlying use of the general linear model (e.g., normally distributed data and homogeneity of variance). This helped determine how the data were distributed on the variables assessed, identified the extent to which there was missing data, and the extent to which the assumptions of the general linear model were met. Correlations were also run to assess the degree of relationship between the independent and dependent variables and to examine the independence of the predictor variables. Preliminary regression analyses also provided diagnostic data for checking assumptions, particularly homogeneity of variance with respect to how well the models functioned in predicting outcomes over different levels of the predictors.

Because of substantial missing data, the following variables could not be used in the subsequent analyses: primary language spoken, preschool attendance, and Title I status. The preliminary analyses also identified problems with the assumption of homogeneity of variance, particularly with the control student data. This led to the use of transformed response scores for some analyses.

Variable Transformations

Based on the distribution of the ethnicity variable, it was necessary to re-code it as a dummy variable measured as White vs. other (coded 1 for White, 0 otherwise). Gender was also re-coded as a dummy to create a continuous scaled variable called Male (1 if male and 0 if female).

As noted above, non-constant variance was diagnosed in preliminary regression runs and the response variables were subsequently transformed in an attempt to resolve this problem. Variance stabilizing transformations using square root and log transformations were carried out for the two DIBLES Next composite variables to correct for violations of the assumption of homogeneous residuals (i.e., the difference between predicted and observed reading proficiency scores).

OLS Regression Analysis

The initial kindergarten analysis used an ordinary least squares (OLS) multiple regression approach to estimate the effect of UPSTART participation on reading proficiency at the beginning and middle of kindergarten. Reading proficiency was measured by DIBLES Next composite scores for the beginning and middle of kindergarten. A 2-block hierarchical regression design was employed in which a set of five covariates was entered first to control for differences between the treatment and control group, followed by the treatment-control group comparison. The covariate control set included dummy variables for gender (Male vs. Female), ethnicity (White vs. Other), low income (Yes vs. No), English Language Learner (Yes vs. No),

and whether the child received special education services in kindergarten (SPED vs. Otherwise). The treatment and control group samples are described in Appendix B by school district and compared (unadjusted) on the five covariate measures.

Separate OLS regressions were run in estimating reading proficiency effects for the beginning and middle of kindergarten. The OLS regression procedure was successful in estimating the impact of UPSTART as measured by the DN Composite for the middle of kindergarten. However, difficulties were encountered in estimating impacts for beginning kindergarten students as discussed immediately below.

Diagnosing Heteroskedasticity

In the beginning kindergarten analysis, scatterplots of the residuals in both the initial OLS analyses and in the subsequent analyses using transformed composite scores (square root and log transformations) showed that the assumption of homogeneity of variance had been violated. The diagnosis of heteroskedasticity brought into question the accuracy of the effects estimates for beginning kindergarten using the OLS regression procedure. The solution was to determine what was causing the error variances to differ over levels of the covariates and to shift to a weighted least squares analysis. Exploratory analyses revealed that there was significantly greater variability on the DN Composite for the low income and special education control group students relative to non-low income and regular education control students or treatment group students within the beginning kindergarten sample.

Weighted Least Squares Analysis

The principle source of heteroskedasticity turned out to be the low income variable as opposed to the special education variable. The low income dummy was then used as a weight variable in a weighted least squares (WLS) analysis to correct for heterogeneous variances in the beginning kindergarten analysis. Covariates for special education status, ELL status, ethnicity (White vs. Other) and gender (percent male) were entered as control variables in the WLS analysis in estimating reading proficiency differences between treatment and control group children at the beginning of kindergarten. The initial WLS analysis showed that the ethnicity covariate was not statistically significant in the beginning kindergarten analysis, so it was dropped from the final analysis. The final results reported for the effects of UPSTART participation on reading proficiency at the beginning of kindergarten are based on the WLS analyses and include covariates for ELL, SPED, and gender, with low income status as the weight variable.

Analysis of Implementation Time

An analysis of covariance was used to determine the relationship between the amount of instruction received by UPSTART participants and reading proficiency outcomes. The usage variable (UPSTART Minutes Logged) was co-linear (redundant) with the treatment-control

group dummy variable since its value was zero for the control group students (who never used the UPSTART curriculum) and a positive value for the UPSTART treatment students (who did). For that reason, the usage variable could not be incorporated into the multiple regression analysis with the Group variable. The alternative was to run an analysis of covariance subset for the UPSTART treatment group using an ordinal version (ordered categories) of the usage variable to see what the impact of time in the program was on reading proficiency. This was accomplished by creating a new variable called Usage Group in which usage was factored into four levels corresponding to quartiles of usage for the DN test samples. The analysis of covariance that was run on Usage Group incorporated the five covariates used in the regression analysis as control measures.

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 questions 1 and 2 about the extent to which UPSTART may have facilitated the development of reading proficiency in kindergarten 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 first 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 reading proficiency outcomes.

UPSTART Enrollment

The Waterford Institute provided documentation for a first-year UPSTART enrollment of 1,248 children. The Waterford Institute enrolled most (almost 80%) of the first-year UPSTART students in April and May of 2009. Waterford continued to enroll families over the summer and into the fall of 2009, concluding enrollment in January 2010.

UPSTART Equipment Provided

The kind of education technology supports provided to children enrolled in UPSTART is shown in Table 2 for all 1,248 students enrolled in the first year and for the DIBLES Next kindergarten sample. The majority of the first 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 majority of students in the kindergarten test sample (approximately 64%) also received a computer drive with the curriculum loaded on it.

Next most often, UPSTART loaned personal computers to almost 11% of the enrolled students and gave them free access to the Internet while they used the equipment. A slightly higher percentage of the kindergarten test sample (about 17%) received a free computer loan with free Internet access in the first year of the program. Another 7% of the program participants were given access to a home computer for free while they participated in the program. Similarly, 8% of the kindergarten test sample was given access to a home computer for free while they participated in the program. The remaining 10 to 12% of the enrolled children received various combinations of computer technology to enable them to access the UPSTART curriculum.

Table 2
Percent of Students Provided Equipment by UPSTART

Equipment Provided	All UPSTART	Kindergarten Sample
Drive	70.5	63.5
Computer & Internet	10.7	16.8
Computer	6.8	8.0
Computer & Cellular	5.9	6.6
Computer & Wireless	2.4	2.2
Internet & Drive	1.9	0.7
Computer with Wireless & Internet	0.8	1.5
Cellular & Drive	0.6	0.0
Other	0.4	0.7
Sample Size	N = 1,248	N = 137

UPSTART Graduates

The Waterford Institute defined an UPSTART "graduate" as a participant who:

- Was actively engaged in the program for at least nine months; and
- Had at least 1,000 minutes of usage (16.67 hours) while in the program.

Children who enrolled in the spring of 2009 and only used the program for four months (e.g., through the summer of 2009 only) were not considered “graduates” by the Waterford Institute.

Of the 1,248 children documented as enrolled in UPSTART in the first year of the program, the Waterford Institute classified 741 as graduates of the program. This converts to a *graduation rate of 59%* ($741/1248 = .59$, or 59%).

While the Waterford Institute’s definition of an UPSTART graduate could be considered somewhat arbitrary, it does focus the program design on the theoretical desirability of enrolling preschool children for a full year and not just the summer. It also sets a desired level of minimum exposure to the curriculum, which is 16.7 hours or approximately 13.3 weeks at 75 minutes per week. The logic of the Waterford Institute’s definition of a graduate, however, is somewhat elusive.

The arbitrariness of the definition of what constitutes an UPSTART graduate is reinforced by its lack of an empirical grounding in being related to a valued literacy outcome like reading

proficiency. We found that UPSTART graduation status was not significantly correlated with reading proficiency at either the beginning ($r=.13$) or middle ($r=.10$) of kindergarten based on the test samples in the kindergarten analysis. However, we did find that increasing exposure to the curriculum is a significant predictor of reading proficiency.

In order for UPSTART “graduate” status to be more meaningful, an empirical cut point needs to be established that has some reliable bearing on being predictive of reading proficiency. That will require further research using a criterion-referenced test of reading proficiency. However, we do examine the relationship between level of instruction and reading proficiency in the final section on implementation below.

UPSTART Usage

UPSTART instruction was initiated in mid-April 2009 and continued through the first week in July 2010 for a total possible of 65 weeks of instruction in the “first year” of the program. The program design called for 15 minutes of instruction per day, 5 days a week for a total of 75 minutes of instruction per week. A student enrolled for 65 weeks in the first year would be expected to have engaged in at least 4,875 minutes of instruction, or 81.25 hours of instruction. The average level of usage for all students enrolled in the first year was 53 hours of instruction, which converts to 42 weeks of UPSTART implementation on the average. The students in the DIBELS Next test samples used the UPSTART curriculum for 68 hours of instruction on the average, or approximately 54 weeks of instruction over the course of the first program year.

Students considered to be UPSTART graduates by the Waterford Institute used the UPSTART curriculum for 76 hours of instruction on the average, or approximately 61 weeks over the course of the first program year. Based on these calculations, an UPSTART “graduate” would have to have been engaged for approximately 94% of the total possible instructional time available during the first year of the program (i.e., $61/65 = .94$, or 94%).

The hours of instruction observed for all students documented to be enrolled in the first year of UPSTART are summarized in Table 3 compared to “graduates” and the students in the kindergarten analysis sample. On the average, UPSTART students received 53 hours of instruction whereas those classified by the Waterford Institutes as graduates of the program received almost 76 hours of instruction on the average. The kindergarten analysis sample was in between at almost 68 hours of instruction.

Table 3
Hours of UPSTART Instruction

Group	N	Mean	SD	Range
All UPSTART	1,248	53.16	40.92	<1 – 313.57
UPSTART “Graduates”	741	75.95	37.31	16.86 - 313.57

Kindergarten Sample	137	67.73	44.04	<1 – 237.92
---------------------	-----	-------	-------	-------------

Inspection of the histograms in Figures 1-3 showing the distributions of hours of instruction for the three groups suggests that the kindergarten analysis sample is more representative of the first year UPSTART program population than the “graduates” group.

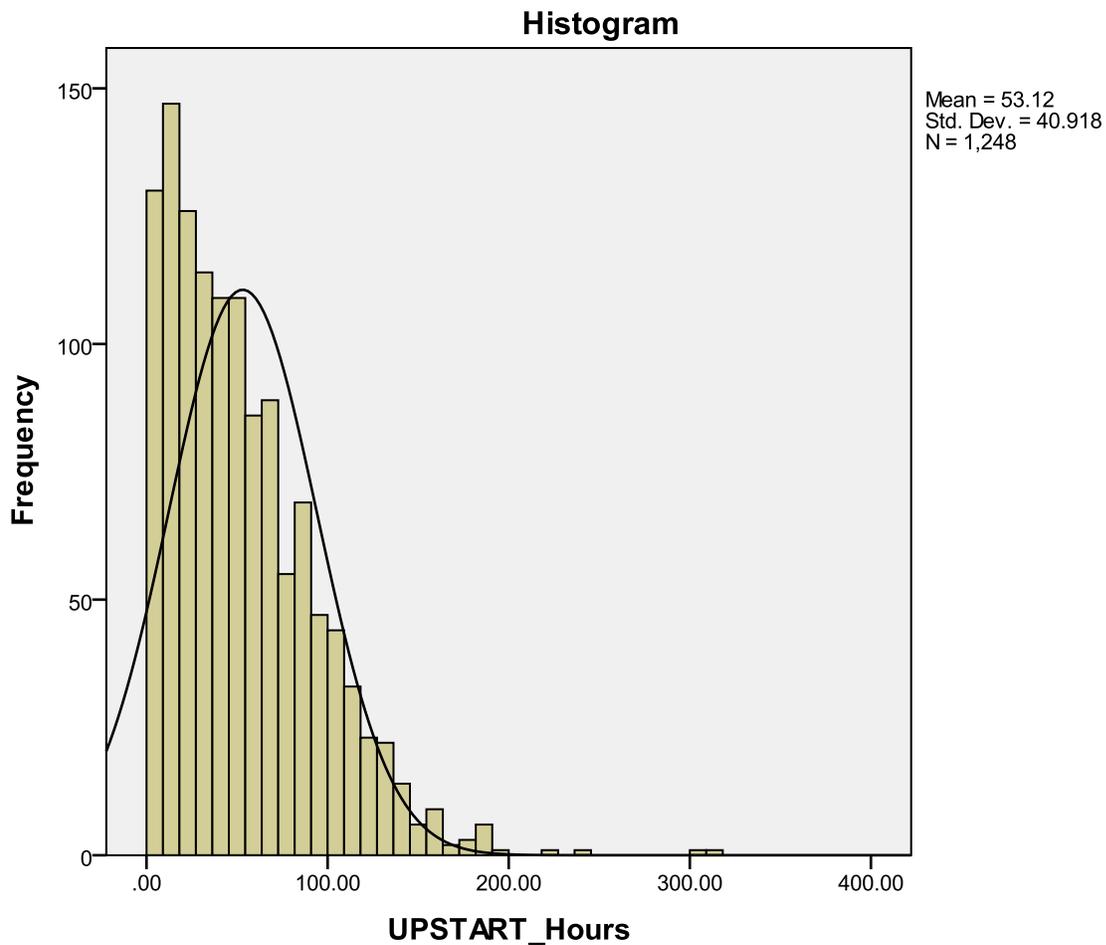


Figure 1. Hours of Instruction for All Students Enrolled in UPSTART in Year 1

Usage of the UPSTART curriculum for all students enrolled in the first year of the program (see Figure 1) is right-skewed with a mean of approximately 53 hours of instruction and a standard deviation of 41 hours. Because of the right-skewed nature of the distribution, the median – 45 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 one hour of instruction. At the other end of the distribution, approximately 1% of the enrollment completed 180 or more hours of instruction.

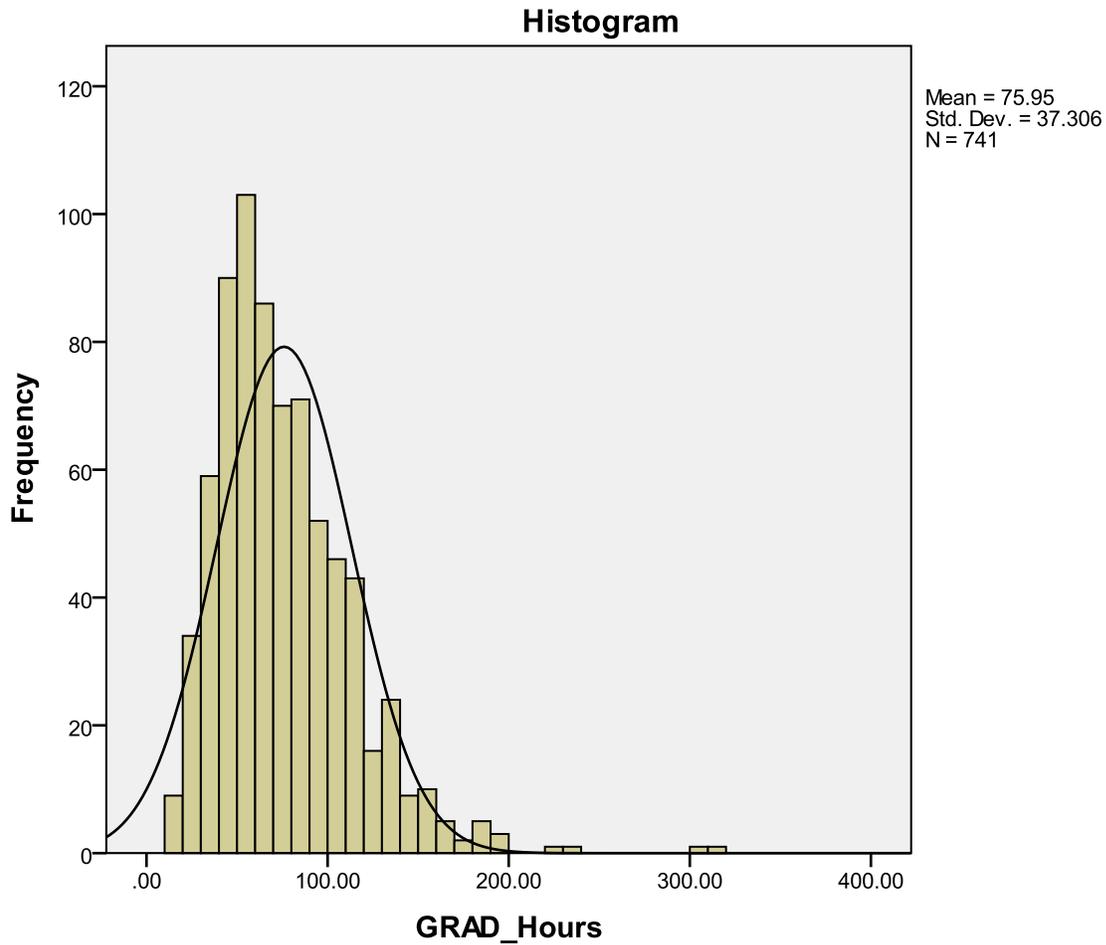


Figure 2. Hours of Instruction for UPSTART Graduates in Year 1

Lop off the bottom quintile from Figure 1 and you have Figure 2. The usage distribution for the “graduates” subgroup starts with 16.9 hours of instruction and runs to 313.6 hours of instruction, as with the full program group. The graduates’ distribution is right-skewed as well, which makes its median value of approximately 69 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 1.407 compared to 1.237).

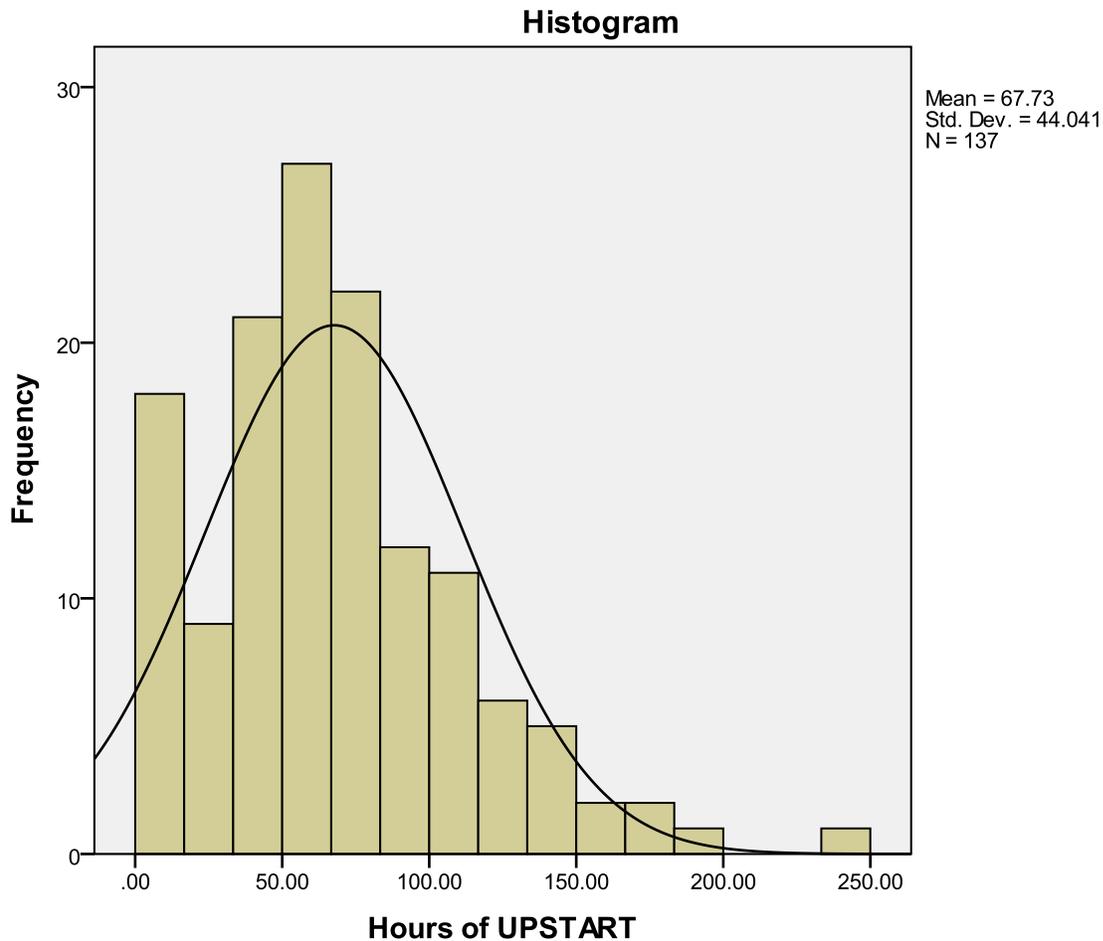


Figure 3. Hours of Instruction for Year 1 UPSTART Kindergarten Analysis Sample

UPSTART usage for the kindergarten analysis sample (see Figure 3) is right-skewed (skew statistic = .845) with a mean of approximately 68 hours of instruction and a standard deviation of 44 hours. The analysis sample’s median is 60 hours of instruction. At the low end, less than 4% of the analysis sample completed less than one hour of instruction. At the high end of the distribution, approximately 2% of the sample completed 180 or more hours of instruction.

How UPSTART Usage Relates to Reading Proficiency

As previously noted, we did establish in the kindergarten analysis that UPSTART curriculum usage is positively and significantly correlated with reading proficiency. This moderately strong relationship was observed at both the beginning ($r=.30$) and middle ($r=.29$) of kindergarten. We also found that increases in reading proficiency were also related to increasing levels of UPSTART curriculum usage. Table 4 shows that UPSTART usage is significantly related to

reading proficiency, statistically controlling for the effects of ELL and special education status (income level, race, and gender are not significant predictors in this model).²

Table 4
Tests of Between Subjects Effects: DN Composite Score - Beginning K-

Source	Type III Sum of Squares	df	Mean Square	F	Sig.	Partial Eta Squared
Corrected Model	21768.801 ^a	5	4353.760	8.629	.000	.257
Intercept	274324.319	1	274324.319	543.678	.000	.813
ELL	12687.081	1	12687.081	25.144	.000	.167
SPED	3185.005	1	3185.005	6.312	.013	.048
UsageGroup	8436.421	3	2812.140	5.573	.001	.118
Error	63071.366	125	504.571			
Total	358277.000	131				
Corrected Total	84840.168	130				

Adjusted R Squared = .23

In Table 5 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 kindergarten analysis sample and displays the difference in DN Composite 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 (under 40 hours in the kindergarten sample) score approximately 23 points lower on the DIBLES Next at the beginning of kindergarten than participants in the fourth quartile of usage (90 or more hours in the kindergarten sample). The parameter estimate for the second quartile of usage indicates that those who used the UPSTART curriculum between 40 and 60 hours scored almost 17 points lower on the average than fourth quartile UPSTART users (90 or more hours). Finally, the parameter estimate for the third quartile of usage indicates that those who used the UPSTART curriculum between 60 and 90 hours scored approximately 13 points lower on the average than fourth quartile UPSTART users (90 or more hours). These estimates clearly indicate a linear trend such that reading proficiency in kindergarten increases with increasing levels of UPSTART curriculum usage in preschool.

² The Partial Eta Square statistic for Usage Group in Table 4 indicates that increasing exposure to the UPSTART curriculum accounts for about 12% of the reading proficiency measured by the DN composite at the beginning of kindergarten. This indicates a moderately strong effect of UPSTART at the beginning of kindergarten.

Table 5
Parameter Estimates: DN Composite Score - Beginning K-

Parameter	B	Std. Error	t	Sig.	Partial Eta Squared
Intercept	64.053	4.244	15.094	.000	.646
ELL	-27.772	5.538	-5.014	.000	.167
SPED	-26.229	10.440	-2.512	.013	.048
[UsageGroup=1.00]	-22.852	5.808	-3.935	.000	.110
[UsageGroup=2.00]	-16.564	5.559	-2.980	.003	.066
[UsageGroup=3.00]	-12.681	5.596	-2.266	.025	.039
[UsageGroup=4.00]	0 ^b

a. Parameter set to zero because it is redundant

This positive linear trend in reading proficiency associated with usage quartile is evident in Figure 4 below

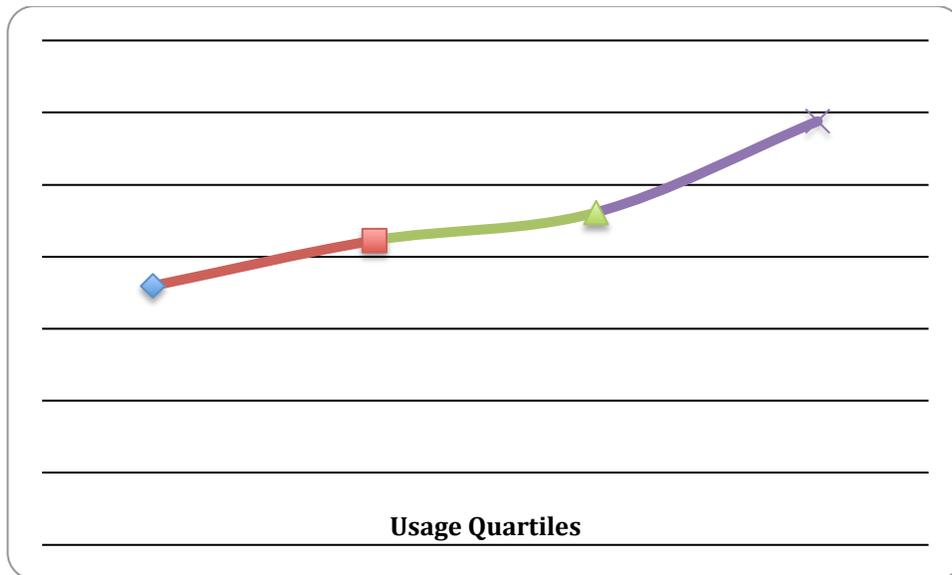


Figure 4. Mean DN Composite B Scores by UPSTART Usage Quartile

The results are much the same for the UPSTART children in middle kindergarten – see Tables 6 and 7 -- and show that increases in reading proficiency correspond with increasing usage of the UPSTART curriculum. In middle kindergarten, UPSTART usage accounts for about 19% of the children’s reading proficiency as measured by the DN Composite at that level.

Table 6
Tests of Between Subjects Effects: DN Composite Score - Middle K-

Source	Type III Sum of Squares	df	Mean Square	F	Sig.	Partial Eta Squared
Corrected Model	95390.609 ^a	5	19078.122	7.346	.000	.223
Intercept	3081911.705	1	3081911.705	1186.759	.000	.903
ELL	57905.486	1	57905.486	22.298	.000	.148
SPED	15451.424	1	15451.424	5.950	.016	.044
UsageGroup	31826.081	3	10608.694	4.085	.008	.087
Error	332404.943	128	2596.914			
Total	3759082.000	134				
Corrected Total	427795.552	133				

Adjusted R Squared = .193

In Table 7, the parameter estimates for Usage Group indicate that participants in the first quartile of usage (under 40 hours) score almost 40 points lower on the DIBLES Next in the middle of kindergarten compared to participants in the fourth quartile of usage (90 or more hours). The parameter estimate for the second quartile of usage indicates that those who used the UPSTART curriculum between 40 and 60 hours scored almost 38 points lower on the average than fourth quartile UPSTART users (90 or more hours). The parameter estimate for the third quartile of usage (60-90 hours) indicates that these participants scored about 24 points lower on the average than fourth quartile UPSTART users.

Table 7
Parameter Estimates: Composite Score - Middle K- DIBELS Next

Parameter	B	Std. Error	t	Sig.	Partial Eta Squared
Intercept	194.622	9.454	20.585	.000	.768
ELL	-59.227	12.543	-4.722	.000	.148
SPED	-53.328	21.862	-2.439	.016	.044
[UsageGroup=1.00]	-39.811	12.929	-3.079	.003	.069
[UsageGroup=2.00]	-37.523	12.501	-3.002	.003	.066
[UsageGroup=3.00]	-24.112	12.583	-1.916	.058	.028
[UsageGroup=4.00]	0 ^b

Again, these estimates suggest a linear trend, with reading proficiency increasing with increasing levels of UPSTART curriculum usage. The data for this trend are presented in Figure 5 below. The shallow slope in Figure 5 is consistent with the with partial eta squared statistic for Usage Group in Table 6 which suggests a smaller effect for UPSTART at the middle kindergarten level.

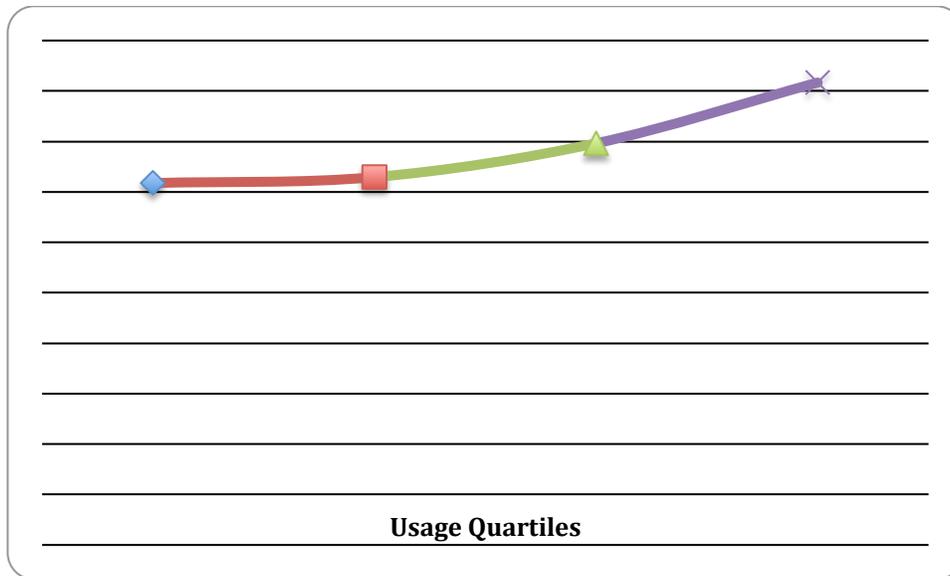


Figure 4. Mean DN Composite M Scores by UPSTART Usage Quartile

UPSTART Outcomes

A weighted least squares regression approach was used to estimate reading proficiency at the beginning of kindergarten as measured by the DIBLES Next (DN) beginning kindergarten composite. Low income status was used as a weight variable to correct for non-constant variance in estimating the regression coefficients. ELL status, special education status and gender were used as control variables in the final model to estimate reading proficiency for children who had participated in UPSTART the previous school year compared to children from the general kindergarten population who had not participated in UPSTART.

Does UPSTART improve reading readiness?

The WLS ANOVA summary table shows that the four-factor regression model is statistically significant. The results of interest are shown in Table 9.

Table 8
WLS ANOVA Summary Table for DN Composite – Beginning of Kindergarten

	Sum of Squares	Df	Mean Square	F	Sig.
Regression	19573.703	4	4893.426	11.422	.000
Residual	54837.500	128	428.418		
Total	74411.203	132			

Table 9 shows that ELL status, special education status and being male all tend to depress beginning kindergarten reading proficiency (DN Composite) scores by approximately 8 to 23 points on the average. Special education students score 23 points lower on the DIBLES Next composite at the beginning of kindergarten compared to regular education students on the average. ELL students score 15 points lower than English proficient students on the average at the beginning of kindergarten. And boys score almost 8 points lower than girls on the DN Composite at the beginning of kindergarten. Statistically controlling for these three factors and weighting the data on the basis of income status, we find that children who had participated in UPSTART during preschool scored almost 18 points higher in reading proficiency as measured by the DN Composite compared to beginning kindergarten children who did not participate in UPSTART prior to enrolling in public school. *Based on these results, the evidence is that UPSTART appears to significantly improve reading readiness for beginning kindergarten students.*

Table 9
WLS Regression Coefficients for DN Composite – Beginning of Kindergarten

	Unstandardized Coefficients		Standardized Coefficients		t	Sig.
	B	Std. Error	Beta	Std. Error		
(Constant)	30.122	3.074			9.797	.000
ELL	-15.278	3.909	-.304	.078	-3.909	.000
SPED	-23.303	6.668	-.271	.078	-3.495	.001
MALE	-7.664	3.625	-.162	.077	-2.114	.036
GROUP	17.772	3.946	.345	.077	4.503	.000

Adjusted R Square = .24

Considering the amount of variance accounted for by the beginning kindergarten model (an adjusted R Square of .24) and judging from the Beta statistic (the standardized regression coefficient value of .35), the UPSTART effect size would be considered to represent a *moderately strong impact*.³ Thus, the available evidence from the first year of the program is that UPSTART does appear to have a moderately strong impact on helping young children develop beginning reading skills to a significantly greater extent than they would have without participating in the program.

Does UPSTART sustain improvements in reading readiness?

The DN Composite for middle kindergarten covers more reading skills than the beginning reading composite. At the middle kindergarten level, phonemic awareness continues to be measured on the DIBLES Next through the First Sound Fluency and Letter Naming Fluency subtests. However, the Phoneme Segmentation Fluency subtest is added along with measures of

³ See Chapter 9 in Cohen, J. (1988). *Statistical Power Analysis for the Behavioral Sciences*. Hillsdale, NJ: Erlbaum.

the alphabetic principle using the Nonsense Word Fluency Subtest and whole word reading is scored on this subtest as well.

The OLS regression model summary for middle kindergarten is displayed in Table 10 and shows that the four measured covariates (Model 1) account for 15% of the variance in DIBLES Next reading proficiency in the middle of kindergarten. The model summary also shows that adding the GROUP variable (Model 2) makes a statistically significant contribution in accounting for variation in DIBLES Next test scores, meaning that UPSTART participation continued to make a difference in kindergarten reading proficiency through the middle of kindergarten. The overall model accounts for 17 percent of middle kindergarten reading proficiency as measured by the DIBLES Next and indicates that UPSTART students were able to sustain their gains in reading proficiency through the middle of kindergarten.

Table 10
OLS Regression Model Summary
DN Composite – Middle of Kindergarten

Model	R	R Square	Adj R Square	Std. Error of the Estimate	Change Statistics				
					R Square Change	F Change	df1	df2	Sig. F Change
1	.402 ^a	.162	.153	54.544	.162	18.126	4	376	.000
2	.430 ^b	.185	.174	53.866	.023	10.524	1	375	.001

Table 11 shows ELL status, low income status, and special education status all tend to depress middle kindergarten reading proficiency by approximately 14 to 46 points. Ethnicity also has an effect on reading proficiency in the middle kindergarten model. Specifically, Caucasian students tend to have higher reading proficiency scores – almost 17 points higher on the average -- than non-White students by the middle of kindergarten.

Statistically controlling for these four factors, we find that middle kindergarten children who had participated in the UPSTART preschool program scored approximately 19 points higher in reading proficiency as measured by the DN Composite for middle of kindergarten compared to middle kindergarten children who did not participate in UPSTART prior to enrolling in public school. Based on this evidence, we can say that the *UPSTART participants appear to have sustained their gains in reading proficiency through the middle of kindergarten.*

Table 11
 OLS Regression Coefficients for DN Composite – Middle of Kindergarten

	Unstandardized Coefficients		Standardized Coefficients		t	Sig.
	B	Std. Error	Beta	Partial r		
(Constant)	136.378	6.991			19.508	.000
ELL	-27.881	8.438	-.184	-.168	-3.304	.001
LOW INCOME	-13.724	6.105	-.113	-.115	-2.248	.025
SPED	-46.108	11.741	-.186	-.199	-3.927	.000
WHITE	16.592	6.953	.133	.122	2.386	.018
GROUP	19.089	5.884	.154	.165	3.244	.001

Adjusted R Square = .17

Considering the amount of variance accounted for by the middle kindergarten model (an adjusted R Square of .17) and judging from the Beta statistic for the Group variable (the standardized regression coefficient value of .15), UPSTART’s *sustained effect size* would continue to be considered a *moderately strong impact*. Thus, the available evidence indicates that the UPSTART effect has been sustained through the middle of kindergarten.

In our follow-up, we will see if the UPSTART effect continues to be sustained when we check the children’s reading proficiency skills again in first grade.

Summary and Conclusions

This final section of the first year 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 reading proficiency as measured in kindergarten.

Test Data Collected and Analyzed

Test data were obtained from 10 Utah public school districts for 258 children who had participated in UPSTART during its first year of operation during the 2009-2010 school year. The test data came from three assessments: the DIBELS Next, the DIBELS, and the DRA. Control group data were provided by the school districts for all nonparticipating students tested. DRA test data were supplied for 62 participants from one school district; DIBELS data were supplied by two districts for 62 participants; and DIBELS Next (DN) data were supplied by seven districts for 137 participants.

We needed a sample size of at least 90 UPSTART participants in order to use multiple regression analysis to test a model with six factors, which was our intent. The model to be estimated involved five covariates (gender, race, ELL status, low income status, and special education status) plus the independent variable of interest: the UPSTART treatment group compared to nonparticipant controls. Test data were sufficient only for an analysis using the DIBELS Next data supplied by seven school districts for 137 former UPSTART students.

The response variables selected for the kindergarten analysis included the DN Composites for the beginning and middle of kindergarten. Thus, the final analysis of UPSTART's impact in kindergarten was based on DN Composite scores for 137 children enrolled in seven Utah school districts.

The Analysis

A weighted least squares (WLS) regression approach was used to estimate reading proficiency at the beginning of kindergarten as measured by the DIBELS Next (DN) beginning kindergarten composite. Low income status was used as a weight variable to correct for non-constant variance in estimating the regression coefficients. ELL status, special education status, and gender were used as control variables in the final WLS model to estimate reading proficiency for children who had participated in UPSTART the previous school year compared to children from the general kindergarten population who had not participated in UPSTART.

An ordinary least squares (OLS) regression approach was used to estimate reading proficiency in middle kindergarten as measured by the DIBLES Next (DN) middle kindergarten composite. The final OLS model employed a hierarchical block design in which a control set of four covariates (dummy variables for race, ELL status, low income status, and special education status) were entered first followed by the set of interest (treatment vs. control).

The dependent variables modeled were:

- The DN beginning kindergarten composite, which included the First Sound Fluency subtest and the Letter Naming Fluency subtest; and
- The DN middle kindergarten composite, which in addition to higher levels of the first two subtests also included the Phoneme Segmentation and Nonsense Word Fluency subtests.

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 kindergarten analysis sample. The final ANCOVA models statistically controlled for ELL status and special education status in estimating the effect of UPSTART usage on reading proficiency at the beginning and middle of kindergarten using the DN composites. 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 first-year UPSTART enrollment of 1,248 children. A majority (56%) of the 1,248 preschool children that enrolled in the first year of UPSTART were from low income families, according to data provided by the Waterford Institute. Slightly more boys (52%) were enrolled than girls (48%). In terms of ethnicity, the vast majority (81%) of the enrollment was Caucasian, 13% were Hispanic, 2% were of Asian descent, 1% were Black, and 1% were Native American. Ethnicity was unknown for 1% of the first year enrollment.

Most of the first year participants (70%) received a computer drive with the UPSTART curriculum loaded on it. Approximately 11% of the first year participants received a computer loan and free Internet access to help them access the UPSTART curriculum. Another 7% of the first year participants were loaned a personal computer to use at home while participating in UPSTART. The remaining 10 to 12% of the first 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 kindergarten test sample had a mean of 68 hours of UPSTART curriculum usage over the first year of the project. This compares with an average of 76 hours of instruction for program “graduates” and an average of 53 hours of instruction for all students enrolled in UPSTART in the first year.
- Length of participation in the UPSTART curriculum was significantly and positively correlated with reading proficiency at the beginning ($r=.30$) and middle ($r=.29$) of kindergarten.
- Reading proficiency improved with increasing levels of UPSTART curriculum usage for both beginning and middle kindergarten children.
- The UPSTART graduation rate in the first year of the program was 59%. UPSTART graduation status was not significantly correlated with reading proficiency at either the beginning ($r=.13$) or middle ($r=.10$) of kindergarten.

Findings: UPSTART Impact on Reading Proficiency in Kindergarten

- UPSTART participation had a moderately strong impact on improving the reading proficiency of UPSTART participants compared to nonparticipants at the beginning of kindergarten, as measured by the DIBELS Next beginning kindergarten composite. This analysis controlled for the significant influence ELL status, special education status, and gender on reading proficiency outcomes at the beginning of kindergarten.
- UPSTART participants maintained their achievement gains through the middle of kindergarten as measured by the DIBELS Next middle kindergarten composite. This analysis controlled for the significant influence of ELL status, special education status, low income status and race on reading proficiency outcomes in middle kindergarten.

Appendix A

Statistical Power Analysis

Two power analyses were conducted in planning the sample sizes needed for the first year kindergarten analysis of UPSTART. The power analyses differed in terms of assumptions about the minimum effect size sizes to detect for UPSTART in terms of the R-Squared increment that would be detected over and above that of the control variable set. Both analyses assumed a hierarchical regression analysis in which a set of 5 covariates would be entered in a control block that would yield an R-squared of .20, followed by the UPSTART treatment vs. control group variable of interest. Alpha was set at .05 and power was set at .80.

In Model 1, the minimum UPSTART effect was specified as an R-squared increment of .02. This was the absolute smallest effect that could be meaningfully detected. The power analysis determined that a sample size of 310 cases per group would be required to detect such a small treatment effect.

In Model 2, the minimum UPSTART effect was specified as an R-squared increment of .07. This was also a small effect but was considered more meaningful as a realistic estimate of the UPSTART effect. The power analysis determined that a sample size of 90 cases per group would be required to detect a treatment effect of this size.

In planning the samples, it was determined that the analysis required a treatment group of at least 90 cases. However, it was recognized that 300 cases per group would be more optimal. The 137 treatment group cases met the minimum sample size criterion. Since we had a pool of over 9,000 control cases, we decided to optimize the control group sample by using a 2:1 ratio of control cases to treatment group cases and consequently selected a random sample of 274 controls stratified by school district.

Appendix B Samples

*Table B.1
Treatment-Control Group Sample Sizes by School District*

School Districts	Sample Sizes	
	Treatment Group	Control Group
Box Elder	16	32
Granite	58	116
Iron	10	20
Ogden	6	12
Salt Lake	27	54
Tooele	10	20
Weber	10	20
Total	137	274

*Table B.2
Treatment-Control Sample Demographics*

Demographic	Group	N	Mean	SEM	t	Significance
% Male	Treatment	137	59	4	2.18	**
	Control	274	48	3		
% White	Treatment	137	77	4	3.20	**
	Control	274	62	3		
% ELL	Treatment	137	15	3	-1.27	NS
	Control	274	20	2		
% Low Income	Treatment	137	30	4	-2.87	**
	Control	274	44	3		
% SPED	Treatment	137	5	2	-0.44	NS
	Control	274	6	2		

**p<.01

The reader should note that the two critical covariates in most of the first year impact analyses are the ELL and SPED variables and that the treatment and control groups are both statistically equivalent on these two covariates.

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	3
INTRODUCTION	6
BACKGROUND	6
EVALUATION DESIGN	7
Research Questions	8
Outcome Measures	9
Data Collection	10
PRESCHOOL DATA ANALYSIS	11
Attrition Analysis	11
Analysis Strategy: Research Question 1	12
Analysis Strategy: Research Question 2	12
Analysis of Implementation Time	12
UPSTART IMPLEMENTATION	13
UPSTART Enrollment	13
UPSTART Equipment Provided	13
UPSTART Graduates	14
UPSTART Usage	14
How UPSTART Usage Relates to Literacy Outcomes	18
UPSTART OUTCOMES	22
Do UPSTART students have better literacy skills at kindergarten than control students?	22
Do UPSTART students show stronger literacy growth rates from preschool to kindergarten than control students?	25
SUMMARY AND CONCLUSIONS	29
Test Data Collected and Analyzed	29
The Analysis	29
Findings: UPSTART Implementation	30
Findings: UPSTART Impact on Literacy Skills at the beginning of Kindergarten	31

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.

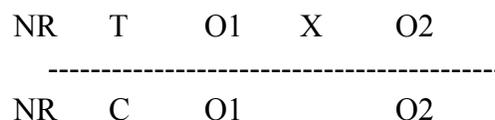
- **Mortality:** We could not rule out whether the observed posttest differences were due to differential dropout rates of families from the study groups that might have changed the composition of the groups over time.

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.



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 its size and direction (i.e., which group is favored over the other by a particular inequality). The pretest also allows us to examine the nature and degree of attrition in the study and whether it differentially affects one group more than the other.

Research Questions

We hypothesized that if UPSTART has no effect on improving early literacy skills, then the preschool children who participated in UPSTART – the treatment group – would be expected to perform at the same level as the comparison group on posttest measures of early literacy development at the beginning of kindergarten. If UPSTART does have an effect on improving early literacy, then the treatment group should perform significantly better than the comparison group on the posttest at the beginning of kindergarten. For purposes of triangulation, we also wanted to take a slightly different look at the data by examining 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: *Do UPSTART students have better literacy skills at kindergarten than comparison group students?*

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

$$T > C @ O2 \text{ (controlling for differences at O1)}$$

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

$$T = C @ O2$$

RQ2: *Do 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:

$$T > C @ O2 - O1 \text{ (growth)}$$

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

$$T = C @ O2 - O1 \text{ (growth)}$$

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 curriculum² 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 comparisons). 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.

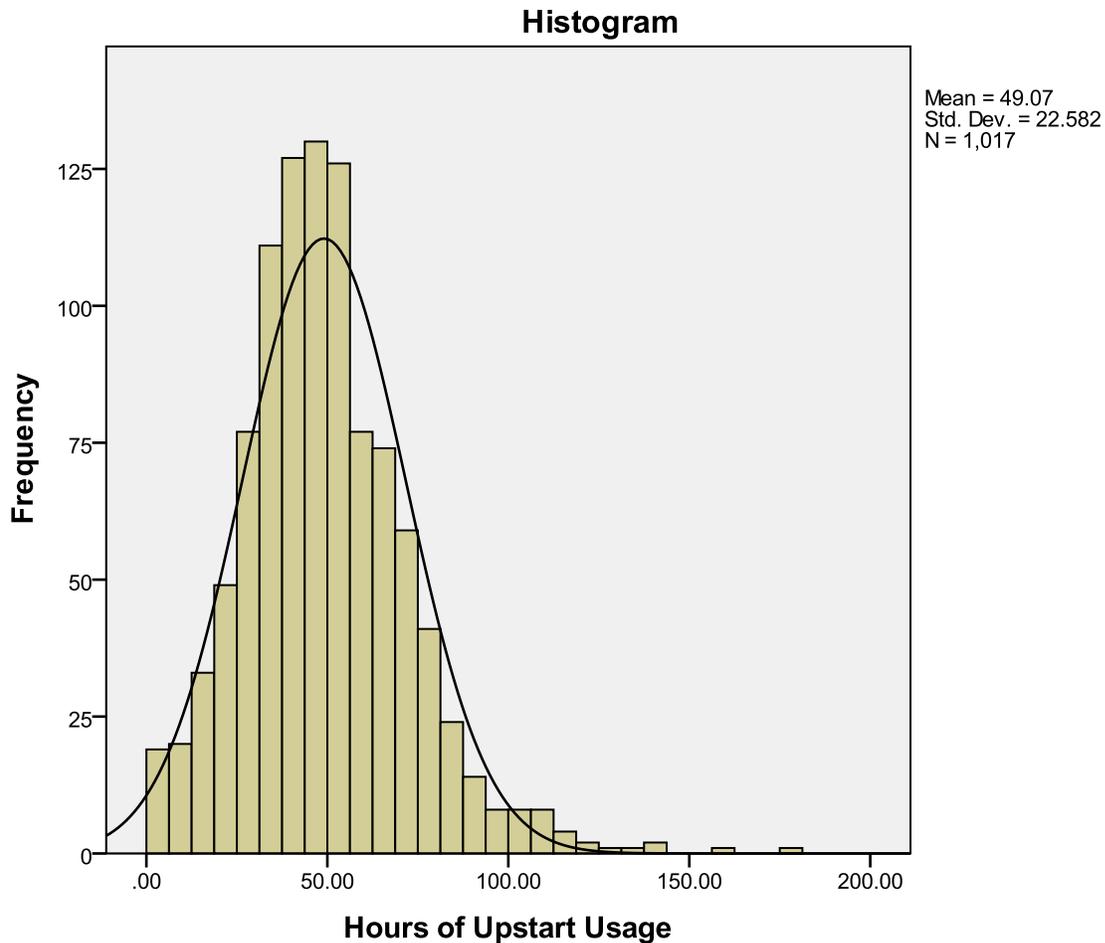


Figure 1. Hours of Instruction for All Students Enrolled in UPSTART in Year 2

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.

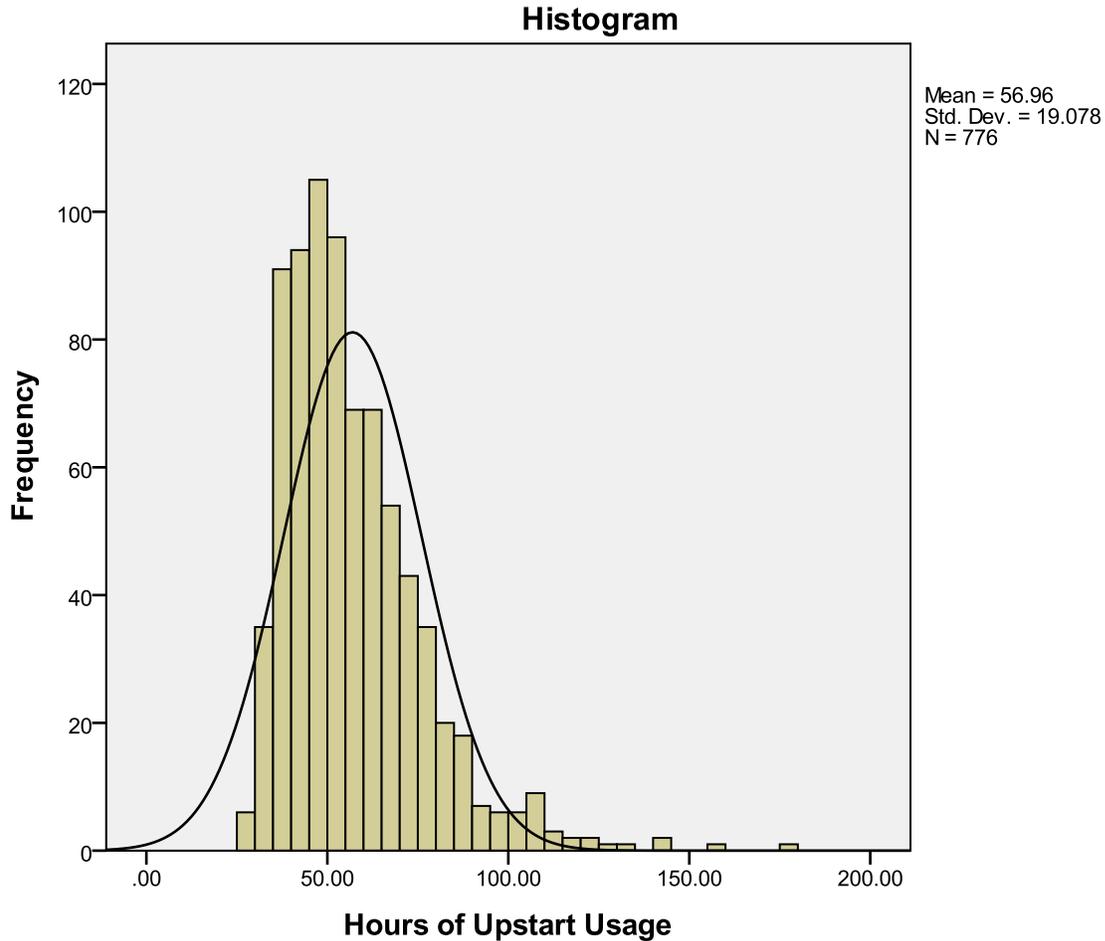


Figure 2. Hours of Instruction for UPSTART Graduates in Year 2

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).

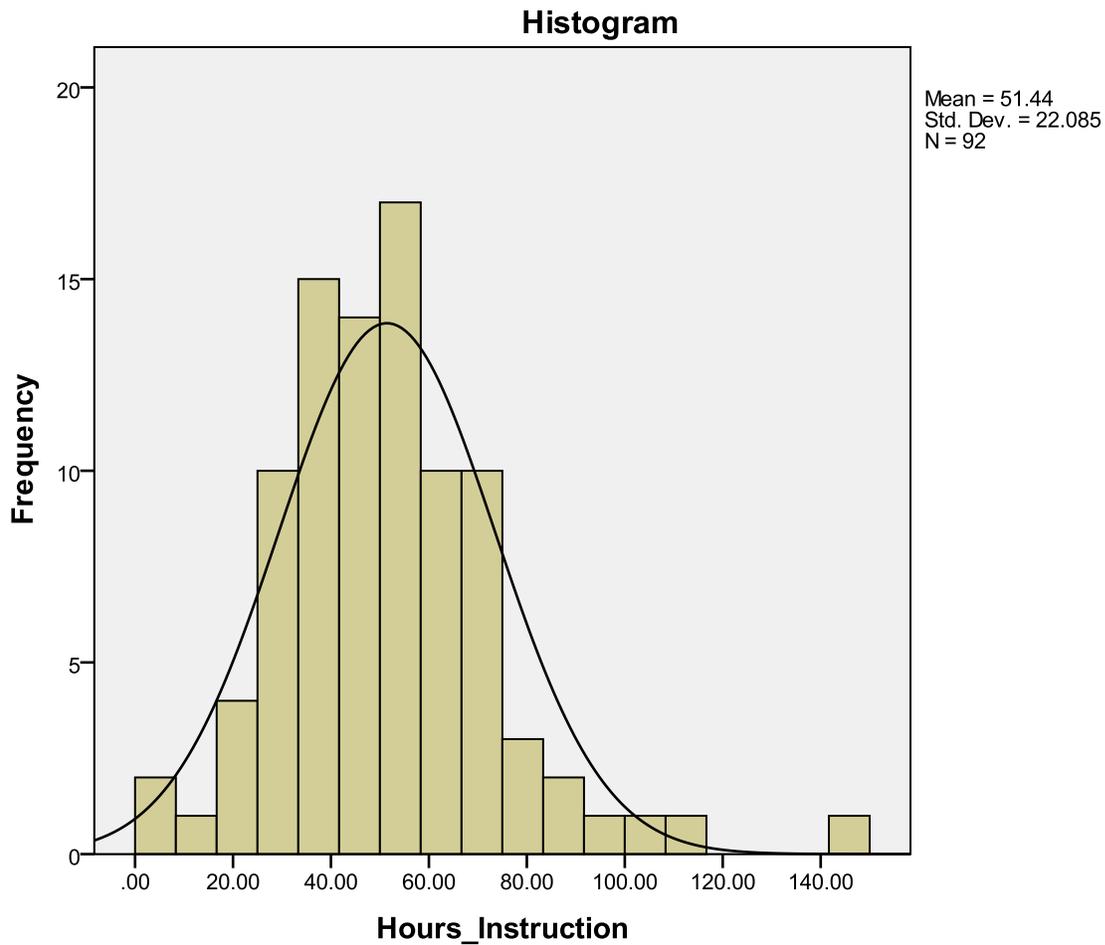


Figure 3. Hours of Instruction for Year 2 UPSTART Preschool Analysis Sample

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 literacy⁵ 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

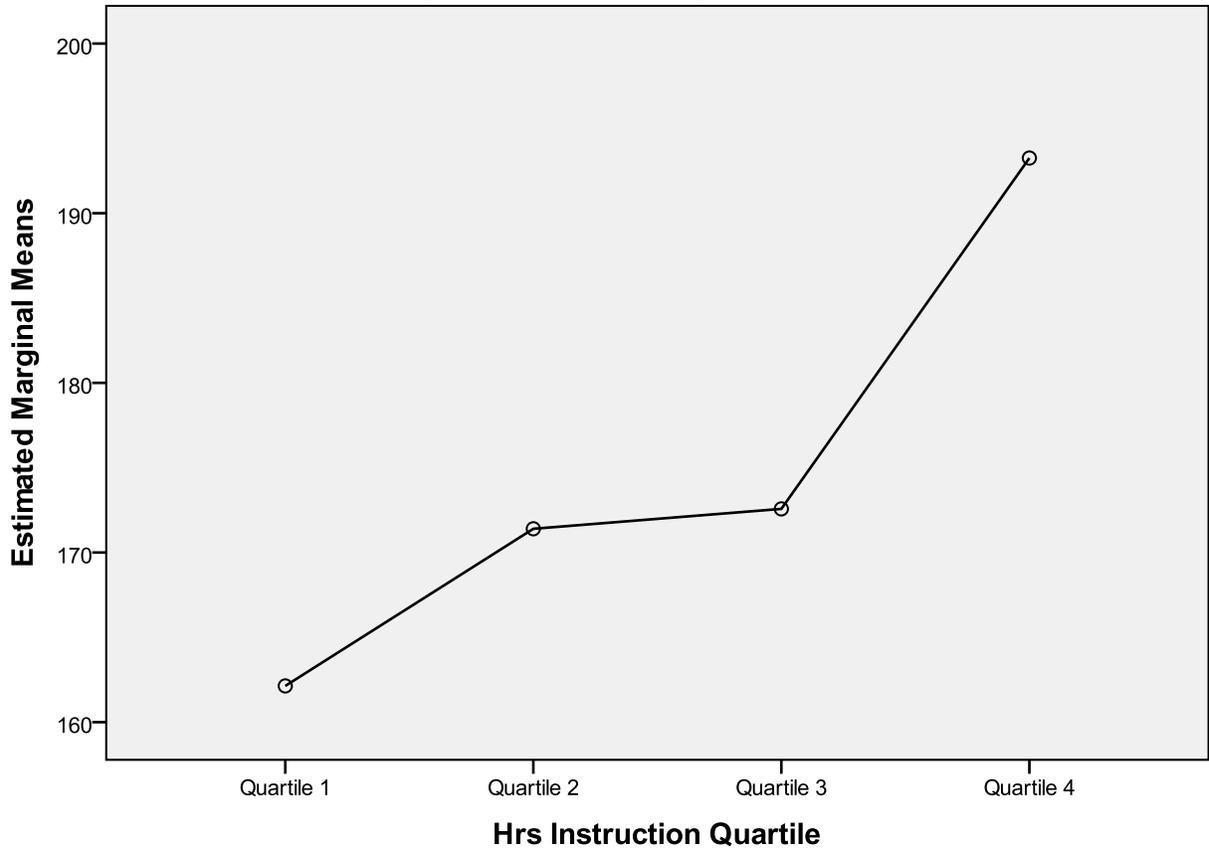
Estimates

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.

Estimated Marginal Means of Brigance Total Posttest



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

Figure 4. Mean Brigance Total Posttest 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		Sig.	Partial Correlation
	B	Std. Error	Beta	t		
(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 F⁸. 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		Sig.	Partial Correlation
	B	Std. Error	Beta	t		
(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* subtest, 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.

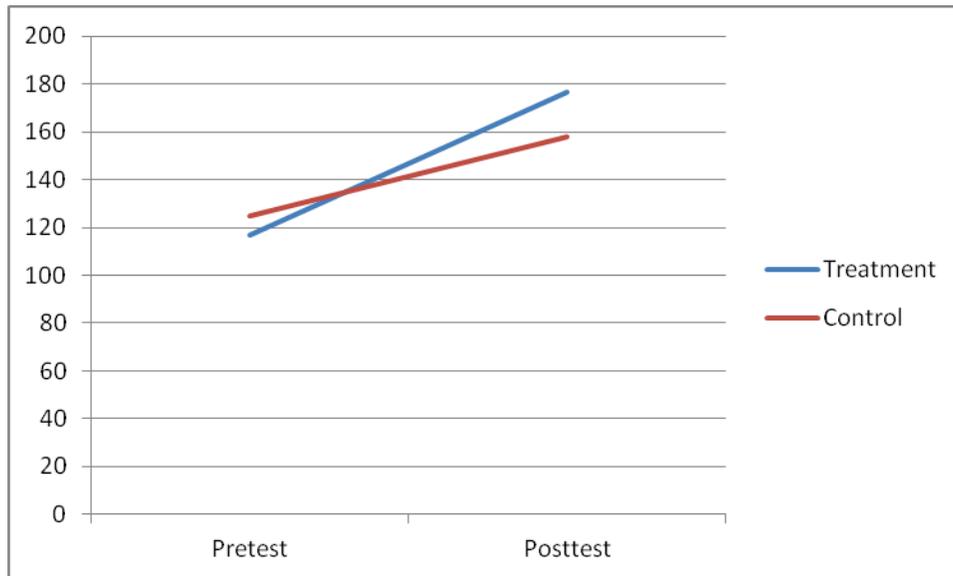


Figure 5. Treatment and Control Group Growth on the Brigance from Pretest to Posttest

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.

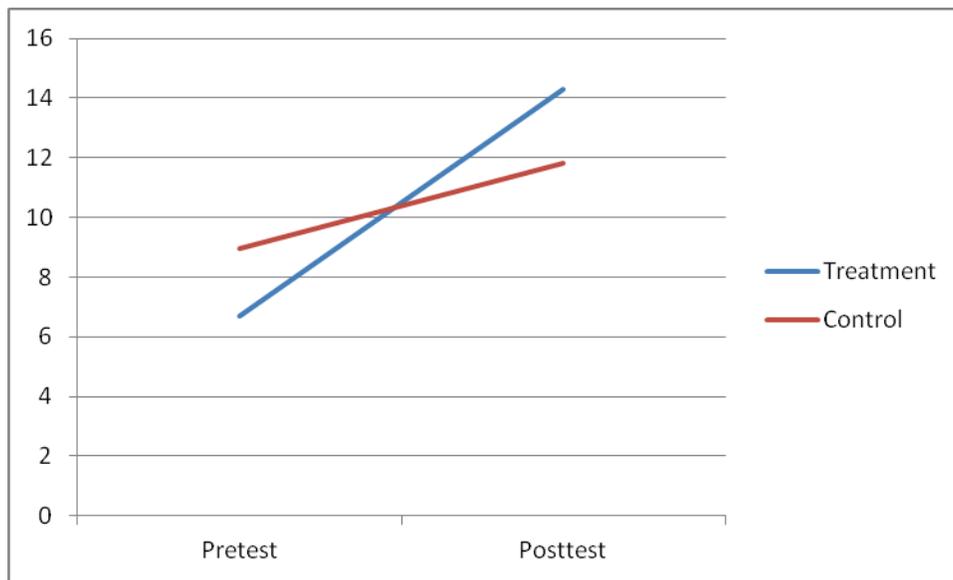


Figure 6. Treatment and Control Group Growth on the Bader from Pretest to Posttest

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 ₃ 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

<i>Covariate</i>	<i>Group</i>	<i>N</i>	<i>Mean</i>	<i>SEM</i>	<i>t</i>	<i>Significance</i>
% Male	Treatment	77	48	5.80	0.53	NS
	Control	81	44	5.60		
% Caucasian	Treatment	76	80	4.65	0.34	NS
	Control	81	78	4.65		
% Hispanic	Treatment	76	09	3.38	-1.26	NS
	Control	81	16	4.10		
% Primary language is English	Treatment	77	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	76	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	76	89	3.59	-0.43	NS
	Control	81	91	3.14		
% Parent is Caucasian	Treatment	76	87	3.95	0.05	NS
	Control	81	86	3.83		
% Parent is Hispanic	Treatment	76	07	2.90	-0.98	NS
	Control	81	11	3.51		
% Parent's primary language is English	Treatment	76	93	2.90	0.98	NS
	Control	81	90	3.51		
Parent Educational Attainment (recoded) ¹⁰	Treatment	76	3.36	0.10	0.72	NS
	Control	81	3.26	0.10		
% Parent is married	Treatment	76	93	2.90	2.49	**
	Control	80	80	4.50		
Parent employment status	Treatment	76	1.61	0.10	-0.41	NS
	Control	81	1.67	0.09		
Spouse employment status	Treatment	72	2.80	0.07	0.59	NS
	Control	70	2.74	0.08		
Household size	Treatment	76	5.04	0.10	2.27	NS
	Control	81	4.68	0.12		
Household income category	Treatment	76	3.71	0.12	0.11	NS
	Control	80	3.69	0.13		

**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

<i>Bader PreTest</i>	<i>Group</i>	<i>N</i>	<i>Mean</i>	<i>SEM</i>	<i>t</i>	<i>Significance</i>
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

<i>Bader Posttest</i>	<i>Group</i>	<i>N</i>	<i>Mean</i>	<i>SEM</i>	<i>t</i>	<i>Significance</i>
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

<i>Covariate</i>	<i>Group</i>	<i>N</i>	<i>Mean</i>	<i>SEM</i>	<i>t</i>	<i>Significance</i>
% 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