

Utah UPSTART Program Evaluation

YEAR 14 PROGRAM RESULTS

Submitted to Utah's Governor's Office of Economic Opportunity

October 31, 2023



EVALUATION AND TRAINING INSTITUTE *A non-profit center for research & evaluation*
100 Corporate Pointe, Suite 387, Culver City, California 90230 || 310-473-8367

ACKNOWLEDGEMENTS

The Evaluation and Training Institute (ETI) thanks Kori Ann Edwards (Managing Director, Strategic Initiatives) from the Governor's Office of Economic Opportunity for the collaboration on this evaluation project. We also thank David Mackay and Stephanie Su (USB E Data and Statistics Team) for their efforts in sharing state KEEP data with us in order to conduct our analysis.

Additionally, we extend our thanks to Claudia Miner, PhD and Haya Shamir, PhD from Waterford for continuing to provide ETI with the necessary UPSTART data used to complete the evaluation each program year.

ABOUT EVALUATION AND TRAINING INSTITUTE

Founded in 1974, the Evaluation & Training Institute (ETI) is a non-profit consulting firm, headquartered in Los Angeles, dedicated to working with schools, post-secondary institutions, public agencies, private foundations, community-based organizations, and professional organizations. We specialize in third-party program evaluations covering many fields, including education, literacy, STEM, social services, health, and prevention. Many of our evaluations have been instrumental in the development of public policy as well as state and federal legislation. Throughout, our focus is on helping clients improve their programs as well as maintain accountability to funders and oversight committees.

Table of Contents

<u>EXECUTIVE SUMMARY</u>	<u>1</u>
<u>INTRODUCTION.....</u>	<u>3</u>
<i>EVALUATION PURPOSE</i>	<i>3</i>
<i>EVALUATION OBJECTIVES</i>	<i>4</i>
<i>UPSTART PROGRAM DESCRIPTION</i>	<i>5</i>
<u>COHORT 14 PRE-KINDERGARTEN STUDY</u>	<u>6</u>
<i>RESEARCH DESIGN</i>	<i>6</i>
<i>RESEARCH QUESTIONS</i>	<i>6</i>
<i>PROCEDURE</i>	<i>7</i>
<i>MEASURES.....</i>	<i>8</i>
<u>PRE-KINDERGARTEN IMPLEMENTATION STUDY RESULTS</u>	<u>11</u>
<i>UPSTART ENROLLMENT</i>	<i>12</i>
<i>UPSTART USAGE</i>	<i>15</i>
<i>UPSTART GRADUATION RATE</i>	<i>18</i>
<u>PRE-KINDERGARTEN ACHIEVEMENT STUDY RESULTS</u>	<u>21</u>
<i>PEEP ENTRY PERFORMANCE LEVELS</i>	<i>21</i>
<i>PEEP EXIT PERFORMANCE LEVELS</i>	<i>22</i>
<u>COHORT 13 KINDERGARTEN STUDY</u>	<u>26</u>
<i>COHORT 13 EVALUATION METHOD SUMMARY</i>	<i>26</i>
<u>KINDERGARTEN ACHIEVEMENT STUDY RESULTS</u>	<u>28</u>
<i>KEEP ENTRY AND EXIT PERFORMANCE LEVELS</i>	<i>28</i>
<u>SUMMARY & RECOMMENDATIONS</u>	<u>31</u>
<i>COHORT 14 PROGRAM IMPLEMENTATION</i>	<i>31</i>
<i>ACHIEVEMENT OUTCOMES</i>	<i>32</i>
<i>STUDY LIMITATIONS</i>	<i>32</i>
<i>RESEARCH AND EVALUATION RECOMMENDATIONS</i>	<i>33</i>
<u>APPENDIX A. PEEP PERFORMANCE LEVELS.....</u>	<u>35</u>

List of Figures

Figure 1. UPSTART Program Enrollment Over Time 12

Figure 2. Percentage of Tier Selection 14

Figure 3. Distribution of Hours of Instruction for C14 Families..... 16

Figure 4. Average Weekly Minutes of Program Use 17

Figure 5. Tier Support and Usage Over Time 17

Figure 6. UPSTART Student Performance Level at PEEP Entry 22

Figure 7. UPSTART Student Performance Level at PEEP Exit..... 23

Figure 8. Kindergarten Readiness among Those Needing Support at Entry for Literacy 24

Figure 9. Kindergarten Readiness among Those Needing Support at Entry for Numeracy..... 25

Figure 10. Program and Non-Program Literacy Performance Levels at Entry and Exit..... 29

Figure 11. Low-Income Program and Non-Program Students Literacy Performance at Entry and Exit 30

Figure 12. Low-Income Graduates, Non-Graduates, and Non-Program Students Literacy Performance at Entry and Exit 31

List of Tables

Table 1. UPSTART Reading and Math Curriculum 5

Table 2. PEEP Exit Profile 10

Table 3. Summary of PEEP Performance Level Descriptors 11

Table 4. Demographic Characteristics of C14 Population 13

Table 5. C14 Hours of UPSTART Instruction 15

Table 6. Average Total Hours and Weekly Minutes by Tier..... 15

Table 7. Waterford Graduation Criteria..... 19

Table 8. Demographic Characteristics of C14 Population 20

Table 9. KEEP Exit Profile..... 27

EXECUTIVE SUMMARY

The Evaluation and Training Institute (ETI) has prepared this report for the Utah Governor's Office of Economic Opportunity (GOEO) to evaluate the UPSTART program in its fourteenth year of implementation. Cohort 14 students were enrolled during the 2022-2023 program year. This year's evaluation includes the results for both program implementation and the achievement outcome analysis, as well as results from the kindergarten program analysis (Cohort 13's kindergarten year).

Waterford enrolled 14,330 preschoolers in the 2022-23 program year, up from the previous year's enrollment of 13,404. This year's cohort of children also marks the fourth and final year of implementing and examining three different tiers of program support for families: Tier 1, UPSTART software only; Tier 2, UPSTART with digital communication for parents to receive text-message program support; Tier 3, UPSTART with full implementation support.

Notable Program Impacts on Student Achievement:

- The Year 14 evaluation found that over 90% of the UPSTART sample achieved proficiency in literacy and numeracy by the end of the program, as measured by Utah's Preschool Entry and Exit Profile (PEEP) assessment.
- An analysis of UPSTART children classified as "needing support" in literacy or numeracy at the beginning of preschool, found that a notable proportion of those children increased their proficiency level to "sufficient prerequisite knowledge" by the end of preschool (i.e., demonstrating the expected early skills for kindergarten entry).
- The majority of Year 13 kindergarten students, who participated in UPSTART during their preschool year (2021-2022), started and ended their kindergarten year at Level 3

proficiency or as defined by the USBE, “demonstrated sufficient knowledge and skills and may require minimal intervention to succeed.”

Notable Program Implementation Details:

- Seventy percent of UPSTART students graduated from the program during the 2022-2023 program year.
- All students (regardless of graduation status) used the UPSTART program for an average of 33 hours during the program year; students who graduated from the program had approximately 42 hours of average use. Logging regular and consistent time with the program was important to meeting the program’s time-centered goals.
- Participation in Tier 3, the most comprehensive level of parental support, yielded higher program usage compared to families opting for Tier 1 (no parental support).
- Children who did not graduate from UPSTART were more likely to have parents with lower levels of education, higher levels of poverty, and were members of an underrepresented racial or minority group compared to children who graduated from the program.

Recommendations

We recommend continuing the UPSTART program based on its association with students’ early skill development. It is important for the program to address the risk factors that contribute to lower graduation rates and a lack of consistency in use among the most vulnerable populations of children. We recommend encouraging consistent *weekly* use to foster productive program engagement so that all children get the most out of their time in the program.

INTRODUCTION

Evaluation Purpose

The Evaluation and Training Institute (ETI), a non-profit research and consulting firm, is leading a multi-year evaluation of the UPSTART program in the state of Utah. This report focuses on findings from the fourteenth program year, which was implemented in 2022-2023. The overarching goal of the evaluation is to help the state and stakeholders determine the student benefits from participating in the program. This report focused on Year 14 preschool students (Cohort 14, or C14) and a cohort of kindergarten students who participated in UPSTART the year prior and completed their kindergarten year in 2023 (C13 students).

ETI studied two core aspects of UPSTART, 1) children's use of the program during the program year (implementation), and 2) an analysis of educational achievement outcomes at the beginning-of-year and also the end-of-year among study participants (outcomes). This year's cohort of children also marked the fourth and final year of implementing and examining three different tiers of program support for families: Tier 1, UPSTART software only; Tier 2, UPSTART with digital communication for parents to receive text-message program support; Tier 3, UPSTART with full implementation support.

Report Roadmap

The sections of this report are organized in the following way: (1) an overview of the objectives, program description and evaluation methodology, (2) program implementation including enrollment, usage, and graduation results across the total program population, (3) a pre-k achievement outcome analysis among Cohort 14 study participants, (4) a kindergarten achievement analysis for Cohort 13 students, and (5) a summary and recommendations from the 2022-2023 program year.

Evaluation Objectives

The overall objectives of the current evaluation are two-fold, (1) study pre-k program implementation to determine the extent to which students adhere to the program's recommended levels of use, and (2) study program achievement outcomes (both at pre-k and kindergarten). The primary goal continues to be children's literacy development since UPSTART's main focus is on early reading; however, an analysis of students' numeracy outcomes is also included. UPSTART provides curriculum for numeracy learning, and the state's preschool and kindergarten assessments were designed to measure both literacy and numeracy. We found it valuable to provide information for both of these domains, as kindergarten readiness encompasses a combination of the two.

The **C14 preschool evaluation** had several main objectives:

- Document the extent to which participants used the computerized curriculum as it was intended.
- Examine the early literacy skills (primarily) and numeracy skills (secondarily) among those participating in the UPSTART program at the beginning and end of the pre-k year.
- Explore the relationship between the support tier selected, curriculum usage and kinder readiness.

The **C13 kindergarten evaluation** had two main objectives:

- Examine the early literacy and numeracy skills before and after kindergarten for a sample of students who participated in the UPSTART program during their preschool year.
- Explore the kindergarten achievement outcome differences between those participating in the program at pre-k, and those who did not.

UPSTART Program Description

Waterford recognizes that nearly half of our country's 4-years-olds face socioeconomic barriers that keep them from early education. *Utah Preparing Students Today for a Rewarding Tomorrow* (UPSTART) is a home-based computer school readiness program developed and provided by Waterford to prepare young children for school entry and future academic success. The program offers in-home, early education access to unserved and often lower income families by providing support, technology, and internet where needed. Waterford enrolled 14,330 preschool children during the 2022-2023 school year, its fourteenth year of operation. Families were provided with an adaptive program of computer-based early literacy and math instruction to prepare them academically for kindergarten.

The UPSTART software used adaptive lessons, digital books, animated songs, and activities to deliver individualized early literacy and numeracy content. Once children met the required literacy usage, they were able to additionally engage in the math/science curriculum. The skills taught by the UPSTART program are outlined in **Table 1**.

Table 1. UPSTART Reading and Math Curriculum

Early Reading Curriculum	Early Math and Science Curriculum
Phonological Awareness	Numbers and Operations
Phonics	Measurement and Data
Comprehension and Vocabulary	Geometry
Language Concepts	Science Concepts

Children were encouraged to use the UPSTART program for a minimum of 1500 minutes across the program year (or 15 minutes a day, 5 days a week). Depending on tier selection,

families were provided with parental resources and technical support from Waterford customer service representatives.

COHORT 14 PRE-KINDERGARTEN STUDY

The following section presents a summary of the research methods used in the Cohort 14 implementation and achievement outcome studies including the design, guiding research questions, our data collection procedures, sample definitions and the measures used to conduct the analysis.

Research Design

ETI conducted a pretest/posttest treatment group only design following the USBE’s recommended approach for the Year 14 evaluation.¹ We collected pretest and posttest assessment data measuring early literacy and numeracy skills from a sample of children enrolled in the UPSTART program. The pretest assessments were conducted prior to the start of the preschool year (2022) and posttest assessments were conducted at the end of the preschool year (2023). The following timeline shows the single group design used to study program outcomes on emerging literacy and numeracy skills.

	Summer 2022		Summer 2023	
Registered UPSTART Students	Pre-Test	UPSTART	Post-Test	Kindergarten

Research Questions

Several research questions guided the current evaluation to determine the extent to which preschool children used the program as it was intended, how UPSTART students scored on measures of kindergarten readiness, and how (if at all) the different levels of parental program

¹ The Year 14 design was agreed upon while the evaluation contract was within the previous department.

support (tier assignment) affected usage and early achievement outcomes. Our research questions for the C14 evaluation were as follows:

Cohort 14 Implementation Study

Research Question 1.1: To what extent did children use the UPSTART program as defined by Waterford (measured in minutes of instruction per week/total minutes for duration of program)?

Research Question 1.2: What proportion of the participants successfully met the requirement for program completion (i.e., “graduated” as defined by Waterford)?

Research Question 1.3: Did different levels of parental program support (tiers) influence program usage?

Cohort 14 Preschool Outcome Study

Research Question 2.1: What proportion of UPSTART students achieved Level 3 (“proficiency”) for kindergarten readiness at the completion of the program year?

Research Question 2.2: What proportion of UPSTART students classified at the start of preschool as “needing support” in literacy or numeracy skill development, were classified as “proficient” in those skills by the end of the program?

Research Question 2.3: Did UPSTART students with different levels of parental program support (tiers) show differences in early achievement outcomes?

Procedure

Implementation Study. To evaluate the implementation of the UPSTART program, data were obtained from records shared by Waterford for all children who had enrolled in the 2022-2023 program year. The records provided data across all usage variables including average weekly use, average number of days per week, average session duration, and overall average total time with the program. Waterford provided program graduation data indicating whether or not the child successfully met the graduation requirements of the program. Data included information about the type of equipment provided to the family by UPSTART as well as the tier-level of support

selected. Parents additionally provided demographic information about the child and household through the program registration process. These datasets were analyzed by ETI to generate the findings related to program implementation, usage, graduation rates, and support tier descriptive statistics.

Preschool Outcome Study. To evaluate the kindergarten readiness of UPSTART students, we studied the test performance for a sample of children enrolled in the program for Year 14. Pretest and posttest data were collected by ETI, including early literacy and numeracy skill assessments. Prior to any child testing, parents of UPSTART children provided informed consent for the testing procedures and also completed a parent survey via an online platform. ETI collected all assessment data through one-on-one virtual test administration with a trained test administrator. The tests were conducted at the beginning and end of the preschool year and all children were assessed with the same state-mandated assessment tool that included several literacy and numeracy scales. The entire assessment procedure was completed in 30-40 minutes on average with identical procedures for each UPSTART child.

Measures

Preschool Study. The outcomes of interest for the UPSTART evaluation were primarily basic literacy skills such as phonological awareness, letter knowledge, and vocabulary. Secondly, we examined numeracy abilities such as numeral recognition and quantity discrimination. We used *Utah's Pre-Kindergarten Entry and Exit Profile (PEEP)* to measure the achievement outcomes among our treatment group. This assessment was recently designed and created by Utah educators, higher education faculty members, and members of the USBE.

Utah's PEEP is intended to inform various stakeholders, such as parents, teachers, and leadership, on the academic and lifelong learning practices essential for entering and exiting preschool students. The assessment is designed to provide the following information:

- Current levels of performance upon entry and exit of preschool.
- Identify students who may need early intervention instruction and promote differentiated instruction for all students.
- Analyze the effectiveness of programs.
- Provide opportunities for data-informed decision-making and cost-benefit analysis of early learning initiatives.
- Identify effective instructional practices or strategies for improving student achievement outcomes in a targeted manner.
- Understand the influence and impact of preschool in both the short- and long-term².

We were most interested in identifying performance levels, determining program effectiveness, and understanding the influence of a preschool intervention for the purposes of this evaluation. PEEP exit subscales measured literacy constructs such as *vocabulary* and *letter knowledge*, while the numeracy constructs included skills such as *counting* and *numeral discrimination*. The PEEP exit consists of two subscales and produces a total score for literacy and a total score for numeracy. All question areas are presented in **Table 2**.

² Information cited from the PEEP administration manual. <https://www.schools.utah.gov/file/56a1e992-800e-4f57-b64f-e9d804f720fe>

Table 2. PEEP Exit Profile

Literacy (Score range 0-108)	Numeracy (Score range 0-36)
Oral language	Oral counting
Alphabet knowledge: uppercase	Numeral identification
Alphabet knowledge: lowercase	Number sense: 1-1 correspondence, cardinality, and quality to numeral
Writing letters	Number sense: numeral to quantity
Phonological awareness	Discrimination: quantity
Alphabetic principle	Discrimination: shape creation
Concept of print: directionality	
Concept of print: letter and word	
<i>PEEP Exit Profile updated August 2021</i>	

The PEEP was also designed to categorize students' performance levels based on their overall literacy and numeracy scores. Students at PEEP entry (prior to starting preschool) are grouped into two levels, defined by USBE as either Level 1 "support needed" or Level 2 "at benchmark". Students at PEEP exit (after completing preschool) are grouped into three levels; Level 1, "does not demonstrate prerequisite knowledge and skills," Level 2, "has minimal prerequisite knowledge and skills" or Level 3, "has sufficient prerequisite knowledge and skills". The performance classifications are considered separate for literacy and numeracy, where a student may be considered (for example) "level 2" in literacy but "level 1" for numeracy. The performance levels are summarized in **Table 3**.

Table 3. Summary of PEEP Performance Level Descriptors

	PEEP Entry	PEEP Exit
Performance Level 1	Demonstrates limited prerequisite knowledge and skills in literacy or numeracy (support needed)	Does not demonstrate prerequisite knowledge and skills in literacy or numeracy
Performance Level 2	Has sufficient prerequisite knowledge and skills in literacy or numeracy (at benchmark)	Has minimal prerequisite knowledge and skills in literacy or numeracy
Performance Level 3	--	Has sufficient prerequisite knowledge and skills in literacy or numeracy

See Appendix A for more detail on PEEP performance level descriptions.

It is important to note that the PEEP is a relatively new measure, developed by the state to use a common assessment for preschool students. The state of Utah administered the PEEP for the first time in 2021, due to a suspension of state-wide testing in the Spring of 2020 as a result of the pandemic. Therefore, the evaluators did not have access to the psychometric properties of the PEEP and cannot report on reliability or validity of the measure. The evaluators used untransformed (raw) PEEP scores to conduct the descriptive analysis.

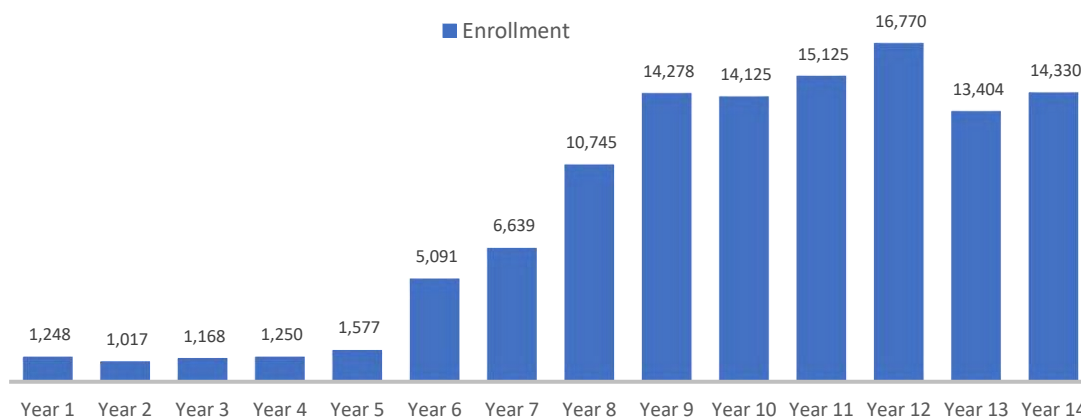
PRE-KINDERGARTEN IMPLEMENTATION STUDY RESULTS

The following presents our findings from Year 14 of UPSTART implementation. Waterford provided a comprehensive dataset to ETI including 2022-2023 UPSTART enrollment, demographic information, provisioned educational technology, UPSTART program usage, and whether or not children completed program requirements as defined by Waterford. The implementation findings are based on all children state-wide who participated in the program.

UPSTART Enrollment

The 2022-2023 program year enrolled 14,330 preschoolers into C14, an uptick from the previous year's cohort (N = 13,404). **Figure 1** illustrates the growth of the program's reach over time from (Year 1, N=1,248) to the most recent Year 14 (N=14,330). UPSTART provided free personal computers to those in need of the technology, resulting in approximately 10% of the C14 children.

Figure 1. UPSTART Program Enrollment Over Time



The demographic makeup of C14 was similar to previous program years, with the vast majority of the cohort classified as Caucasian/White (78%) with 9% from Latino/a origin (**Table 4**). Approximately one third (32%) of the families in Year 14 were under 185% poverty.

Table 4. Demographic Characteristics of C14 Population

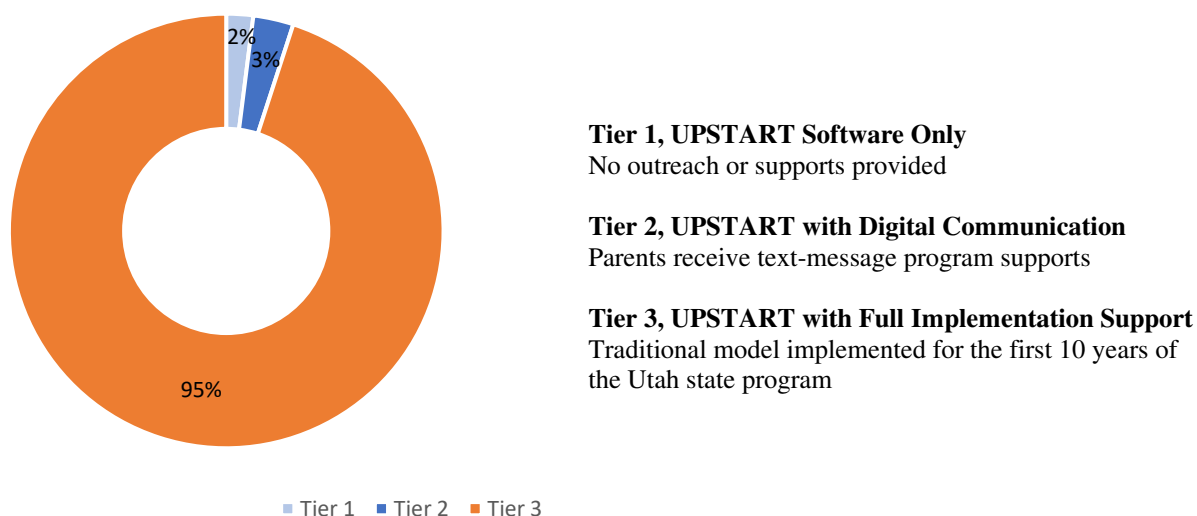
Demographic Categories		C14 UPSTART
Child's Gender		(n=14,330)
	Male	51%
	Female	49%
Child's Ethnicity		(n=13,864)
	Caucasian/White	78%
	Latino/a	9%
	Asian/Pacific Islander	3%
	African American/Black	1%
	Other	9%
Child's Language		(n=14,330)
	English	94%
	Spanish	5%
	Other	1%
Parent Educational Attainment		(n=14,117)
	High school graduate	10%
	Some College	24%
	College Graduate	44%
	Advanced Degree	20%
Parent Marital Status		(n=14,108)
	Married	92%
	Otherwise	8%
Poverty Level		(n=11,759)
	Under 185% Poverty	32%
	Above 185% Poverty	68%
<i>Percentages in the table are based on those providing a response in Waterford's participant records and may not add to 100% due to rounding.</i>		

ETI acknowledges that not all families provided responses to every demographic question, therefore the percentages represent the proportion among those who provided information. This case-wise deletion method is a conventional, traditional, and simple technique used to manage missing data, particularly when reporting univariate statistics (Schafer, J. L., & Graham, J. W. 2002).

For any demographic category with less than the total sample (i.e., 14,330), the sample size of those providing a response was noted within the table.

Parent Support Tiers. For the past several years, parents enrolling their children in the program were asked to select the level of support that worked best for their family. All families were given the option to select one of three support tiers (defined in **Figure 2**). Each year, the vast majority of the participants selected Tier 3 with full implementation support. This current year was no different with Tier 3 appealing to 95% of the registered families (n=13,644) and Tier 1 and 2 garnered approximately 5% (Tier 1, n=335 and Tier 2, n=351).

Figure 2. Percentage of Tier Selection



As illustrated in **Figure 2**, the most popular choice by families was the traditional program support model characteristic of the first decade of the program's implementation. Beginning in Year 15 and moving forward, Tier 3 will be the level of support given to all families participating in UPSTART.

UPSTART Usage

The C14 academic year covered 38 weeks of instruction, beginning the week of September 6, 2022, and ending May 23, 2023. Program usage data were analyzed to understand the extent to which families were using the program and meeting requirements for graduation. The average duration in the program for enrolled families was approximately 33 weeks. Waterford's recommended total usage criteria was 1500 minutes (or 25 hours) across the entirety of the program. The average level of usage for all students enrolled (N=14,330) was approximately 1,980 minutes (or 33 hours) of instruction, down marginally from last year but still well above the recommended usage. The average level of usage specifically among those who met the graduation requirement (i.e., 1500+ total minutes or high achievement on WACS³), was a total of 42 hours. The Year 14 data for instructional hours are summarized in **Table 5**.

Table 5. C14 Hours of UPSTART Instruction

Group	N	Mean	SD	Range
All UPSTART	14,330	32.57	18.91	0.02-456.75
UPSTART Graduates	9,994	41.93	14.07	0.27-456.75

We also examined if usage with the program differed by the parental support tiers. As we have seen in prior years, as the level of support increased, so did the time spent with the program (**Table 6**).

Table 6. Average Total Hours and Weekly Minutes by Tier

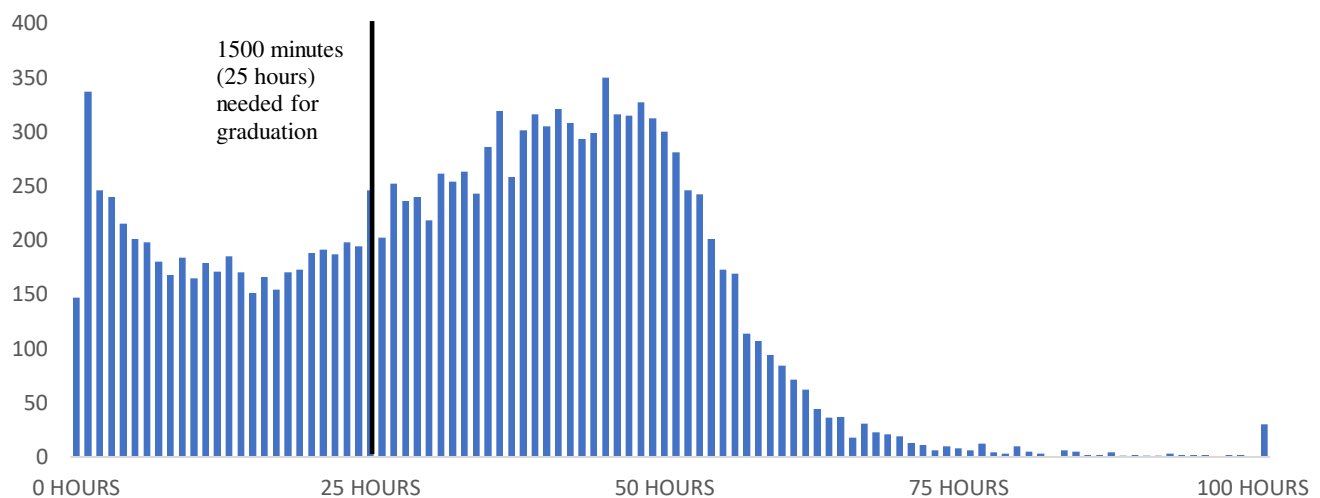
	Avg. Total Hours	Avg. Weekly Minutes
Tier 1	26	48
Tier 2	30	52
Tier 3	33	58

³ WACS Waterford Assessment of Core Skills.

An independent samples t-test was conducted to determine if the average total hours for each tier were statistically different from one another. Results showed that families who received no support (Tier 1) ($M = 26$, $SD = 16.00$), spent less time on a weekly basis and averaged significantly fewer total hours with the UPSTART program compared to those who had full implementation support (Tier 3) ($M = 33$, $SD = 19.02$); $t(13977) = 6.7$, $p = 0.000$. There was a similar relationship seen when comparing families who received some support (Tier 2) ($M = 30$, $SD = 15.53$) and families who had full implementation support (Tier 3) ($M = 33$, $SD = 19.02$); $t(13993) = 2.6$, $p = 0.0089$.

Figure 3 shows the distribution of hours of instruction for the total C14 population ($N=14,330$). As illustrated, the majority of children exceeded the program's minimum requirement of 25 hours, yet roughly 30% of participants logged less than the recommended criteria for program completion.

Figure 3. Distribution of Hours of Instruction for C14 Families.



The bottom quartile of the C14 population completed 18.22 hours of instruction or less, the midpoint of the C14 distribution was 34.47 hours, and the top quartile completed 46.18 hours or

more of instruction. Only 30 students used the program for 100 hours or more. The pattern of weekly

usage across the duration of the program was similar to prior years where dips in program use map to holidays and breaks from school (see **Figure 4**). On average, the 2022-2023 UPSTART children recorded just under the recommended weekly usage amount (75 min) across the program year.

Figure 4. Average Weekly Minutes of Program Use

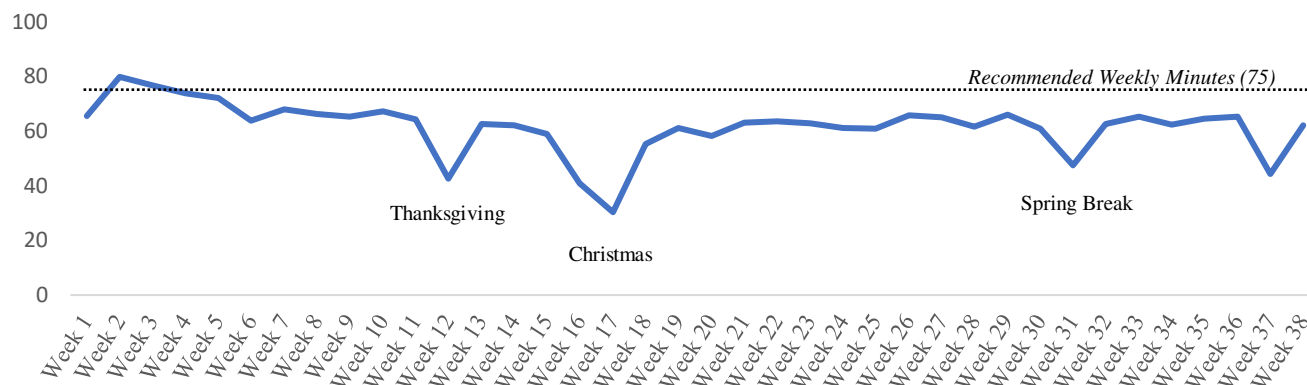
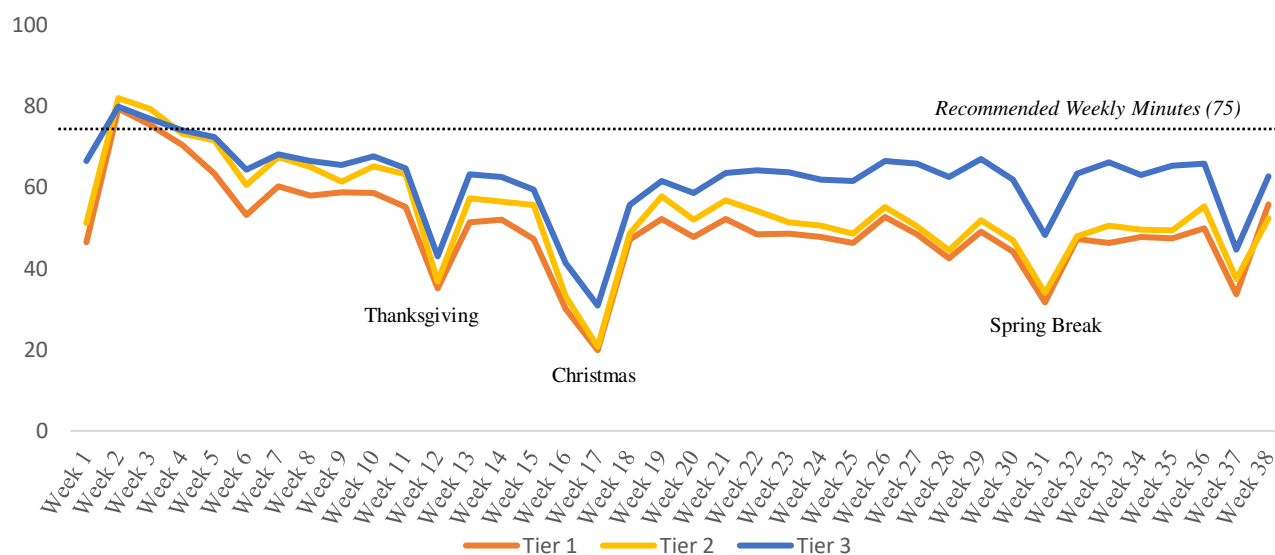


Figure 5 illustrates the same week to week trend for families based on their tier level. Families with more program support (Tier 3, blue line), consistently surpassed the use levels of those in the other tiers. Usage among all tiers was similar in the beginning of the program, yet immediately after

Figure 5. Tier Support and Usage Over Time



Winter break (Week 18), Tier 3 families hovered above both Tier 1 and Tier 2 families by as much as 18 minutes. This same pattern was documented across multiple years of studying the support tiers and provides further evidence that those who received the full implementation support had children who used the program more week to week.

UPSTART Graduation Rate

Prior ETI evaluation studies have shown that successful usage and hence graduation from the program were strongly associated with helping children develop school readiness skills. That is, curriculum usage was significantly and positively related to literacy outcomes as measured by early literacy instruments.

Waterford's graduation criteria changed several years ago to include a Waterford Assessment of Core Skills (WACS) score. The WACS is an adaptive assessment, designed to provide precise data for high and low scoring students by testing students on the following core literacy skills: phonological awareness, vocabulary, listening comprehension, phonics, and reading comprehension⁴. The graduation requirement can now be met in a couple of different ways (a) logging at least 1,500 minutes (25 hours of instruction) with the UPSTART curriculum, and/or (b) WACS achievement level (see **Table 7**).

⁴ Source: Waterford.org

Table 7. Waterford Graduation Criteria

	Low Program Usage (Less than 1500 minutes)	High Program Usage (1500 minutes or more)
High WACS achievement ($\geq$ Kindergarten Beginning AND $\geq$ 1 Level of Growth) OR ($\geq$ Kindergarten Intermediate AND No Pre-test)	4% High achievement Low Usage	39% High Achievement High Usage
Low WACS Achievement ($<$ Kindergarten Beginning OR $<$ 1 Level of Growth) OR ($<$ Kindergarten Intermediate AND No Pre-Test)	3% Low Achievement Low Usage	16% Low achievement High Usage
No WACS	27% No EOY Assessment Data Low Usage	11% No EOY Assessment Data High Usage

Graduation rate and categories provided by Waterford; bold text denotes graduate

As an example, students with high achievement on the WACS (i.e., demonstrated kindergarten readiness skills) qualify as graduates of the program, regardless of their program usage level (low or high). Additionally, students who meet the usage requirement (regardless of achievement level on WACS) are also considered graduates of the program. About 39% of students had high achievement on the WACS assessment and high usage, compared to only 4% of students who had high achievement and low usage. About 30% of students were unable to engage with the program as recommended and also had either low or no WACS scores, hence not meeting the graduation criteria of the UPSTART program.

Waterford classified 9,994 children as program graduates out of the 14,330 enrolled in Year 14. Cohort 14's graduation rate was therefore 70% (i.e., $9,994/14,330 = 0.697$), which is down slightly from the graduation rate from Cohort 13 (73%). **Table 8** displays the demographic characteristics of graduates and non-graduates within Cohort 14.

Table 8. Demographic Characteristics of C14 Population

Demographic Categories		UPSTART Graduates	UPSTART Non- Graduates
Child's Gender		(n=9,994)	(n=4,336)
	Male	51%	50%
	Female	49%	50%
	Unspecified	<1%	<1%
Child's Ethnicity		(n=9,670)	(n=4,194)
	White	80%	72%
	Hispanic	8%	11%
	Asian/Pacific Islander	3%	4%
	Native American	<1%	1%
	African American	1%	1%
	Other	9%	11%
Child's Language		(n=9,994)	(n=4,336)
	English	94%	93%
	Spanish	4%	5%
	Other	1%	2%
Parent Educational Attainment		(n=9,869)	(n=4,249)
	High school graduate	8%	15%
	Some College	22%	29%
	College Graduate	47%	37%
	Advanced Degree	22%	16%
Parent Marital Status		(n=9,862)	(n=4,246)
	Married	94%	86%
	Otherwise	6%	14%
Poverty Level		(n=8,264)	(n=3,495)
	Under 185% Poverty	28%	40%
	Above 185% Poverty	72%	60%
<i>Percentages in the table are based on those providing a response in Waterford's participant records and may not add to 100% due to rounding.</i>			

As we have seen in previous years of study, children who did not meet the graduation requirements were more likely than UPSTART graduates to be a member of an underrepresented

racial or ethnic minority group, have parents with lower levels of education, and higher levels of poverty.

PRE-KINDERGARTEN ACHIEVEMENT STUDY RESULTS

Our analysis on literacy and numeracy outcomes for a sample of Cohort 14 UPSTART students is presented below. PEEP assessment data were analyzed to answer the following research questions:

Research Question 2.1: What proportion of UPSTART students achieved a Level 3 (“proficiency”) for kindergarten readiness at the completion of the preschool program year?

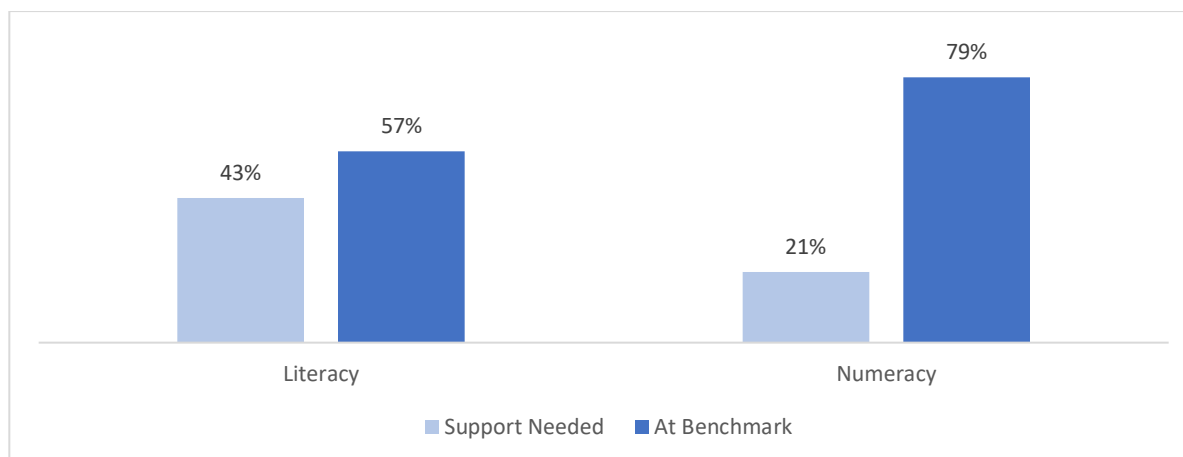
Research Question 2.2: What proportion of UPSTART students classified at the start of preschool as ‘needing support’ in literacy or numeracy skill development, were classified as “proficient” in those skills by the end of the program?

Research Question 2.3: Did UPSTART students with different levels of parental program support (tiers) show differences in early achievement outcomes?

PEEP Entry Performance Levels

We first examined baseline performance among our UPSTART evaluation sample. There were two possible performance levels at PEEP entry, “At Benchmark,” (i.e., demonstrated sufficient prerequisite knowledge and skills) or “Support Needed,” (i.e., demonstrated limited prerequisite knowledge and skills). Among our UPSTART evaluation sample, we found that 57% of students were at benchmark in literacy and 79% were at benchmark in numeracy prior to beginning their preschool year (**Figure 6**; darker blue bars). This illustrates that the majority of the UPSTART students had sufficient prerequisite knowledge (as measured by the PEEP entry) prior to program participation.

Figure 6. UPSTART Student Performance Level at PEEP Entry



Literacy and Numeracy- n=349

Next, we examined the performance levels achieved at the end of the program year.

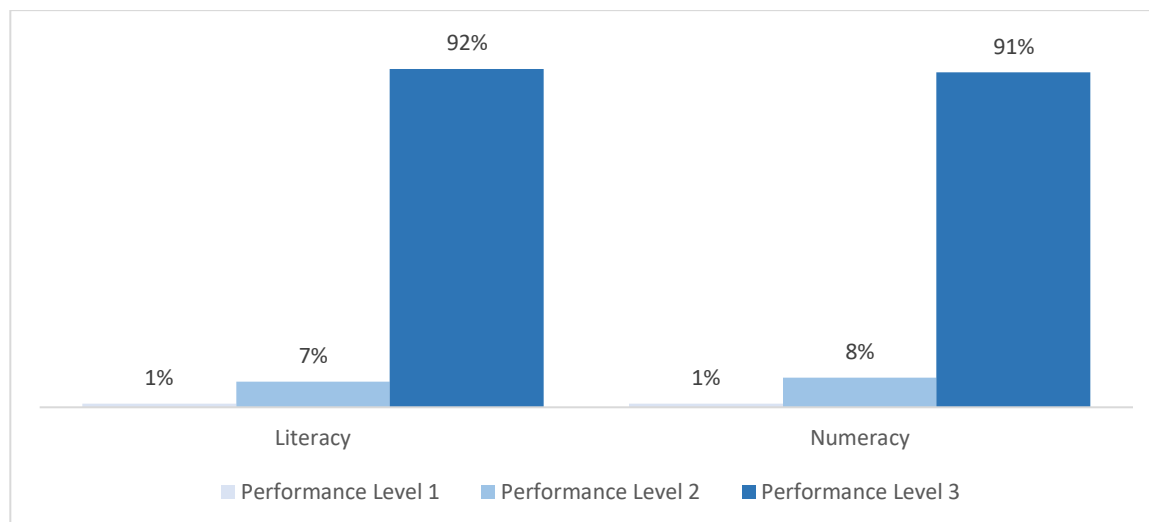
PEEP Exit Performance Levels

What proportion of UPSTART students achieved a Level 3 (“proficiency”) for kindergarten readiness at the completion of the program year?

There were three performance levels a student could achieve on the PEEP Exit, which reflected proficiency in prerequisite kindergarten skills and aligned to the level of intervention a student *may* need in kindergarten. These performance levels are not connected to the PEEP entry classifications and are therefore examined independently from the PEEP entry outcomes. A student at performance Level 3 at exit, is considered “Proficient” in either literacy or numeracy and has demonstrated sufficient knowledge and skills for entering kindergarten. Students at performance Level 2 and Level 1 have not demonstrated the necessary skills for kindergarten and may require additional intervention to succeed.

As shown in **Figure 7**, the vast majority of the UPSTART evaluation sample exited the program “proficient” in both literacy (92%) and numeracy (91%), indicating readiness for kindergarten in skills and prerequisite knowledge. Very few UPSTART students were starting kindergarten needing additional support in either domain (8% or 9%, respectively).

Figure 7. UPSTART Student Performance Level at PEEP Exit



Literacy and Numeracy- n=283

These results suggest that participating in the UPSTART program was related to the children's kindergarten readiness in both literacy and numeracy.

We conducted another set of analyses to see how students most in need of support (i.e., most at-risk of falling behind their peers) at the *beginning* of preschool, progressed in literacy and numeracy skills by the end of the UPSTART program year. The following question guided this inquiry:

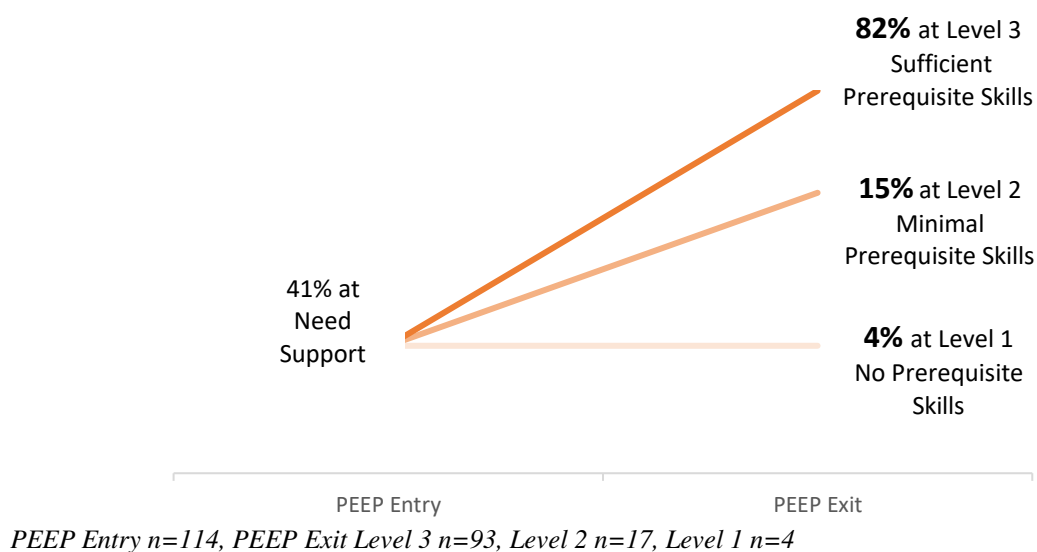
What proportion of UPSTART students classified at the start of preschool as “needing support” in literacy and numeracy skill development, were classified as “proficient” in those skills by the end of the preschool year?

For the analysis, we looked only at the group of students who were classified as needing support at PEEP entry, and then mapped that same group of students to see how they performed at the PEEP exit. We did not connect entry to exit scores in any way, but rather used the USBE's

classification system of performance levels to understand potential academic progress among a single group of students⁵.

Literacy. Among children classified as in need of support for literacy at the start of the program (n=114), 82% of them achieved “sufficient prerequisite knowledge in literacy” or Level 3 at the end of the program (**Figure 8**).

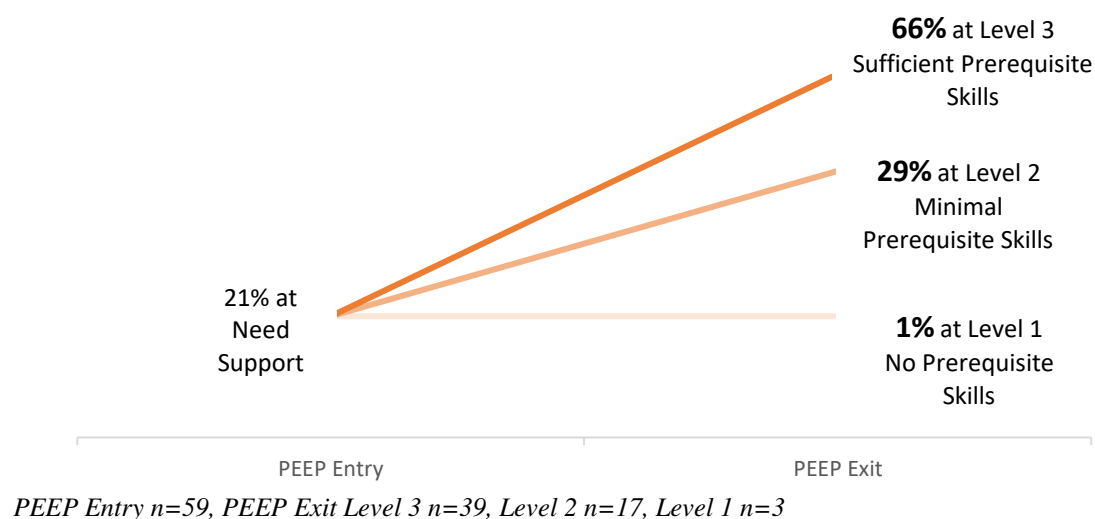
Figure 8. Kindergarten Readiness among Those Needing Support at Entry for Literacy



Numeracy. Similar to literacy, we analyzed how pre-k students most in need of support in numeracy at the *beginning* of the program, progressed in numeracy skills by the end of the program year. We found that 66% achieved Level 3 at the end of preschool (**Figure 9**).

⁵ For additional information on how USBE defines Performance levels and cut points refer to Appendix A.

Figure 9. Kindergarten Readiness among Those Needing Support at Entry for Numeracy



Both sets of findings suggest that the UPSTART program may be particularly beneficial for providing the additional boost in literacy and numeracy skills for the students in need of the most support before entering their kindergarten year.

As a final part of the preschool analysis, we also assessed if student PEEP outcomes were different between those receiving different levels of parental support. We sought to answer the following question:

Did UPSTART students with different levels of parental program support (tiers) show differences in early achievement outcomes?

We examined literacy and numeracy performance levels on the PEEP exit assessment across participants in Tier 3 (full parent support), Tier 2 (digital communication support), or Tier 1 (no parent support). We found that the vast majority of UPSTART students were considered proficient at the end of the program year regardless of parent support tier. That is, there were no significant differences between the selected support tier level and student achievement outcomes on the PEEP.

COHORT 13 KINDERGARTEN STUDY

As part of the Year 14 UPSTART evaluation, we also looked at a sample of past UPSTART students who participated during the 2021-2022 program year, and who finished kindergarten in 2023.⁶ Kindergarten assessment data were available for more than 46,000 students which allowed more flexibility in our analysis including exploring differences between program graduates vs non-graduates as well as a deeper dive into the achievement outcomes of low-income students.

Cohort 13 Evaluation Method Summary

Kindergarten Research Design. ETI conducted a pretest/posttest observational research design utilizing assessment data provided by USBE. We identified a large sample of Year 13 UPSTART students within the state's dataset and then studied the kindergarten literacy achievement outcomes for those C13 students. This process also allowed us to create a state-wide comparison group (excluding the UPSTART children from the dataset). The following timeline shows the evaluation design used to study outcomes in kindergarten.

2021-2022	Fall 2022		Spring 2023
Preschool UPSTART Program Year	KEEP Entry	Kindergarten	KEEP Exit
No UPSTART			

The kindergarten analysis was designed to answer the following research questions:

Research Question 3.1: How did the literacy KEEP performance levels for kindergarteners who participated in the UPSTART program at pre-k, compare to the pattern of performance for Non-UPSTART students?

Research Question 3.2: How did the literacy KEEP performance levels for low-income kindergarteners who participated in the UPSTART program at pre-k, compare to pattern of performance for low-income Non-UPSTART students?

⁶ For this report, we focused on literacy outcomes in the kindergarten analysis

Research Question 3.3: How did the literacy KEEP performance levels for low-income kindergarteners compare across three groups, (1) those who graduated from the UPSTART program at pre-k, (2) those who participated but did not graduate, or (3) those who did not participate at all?

Kindergarten Data Collection. We used secondary assessment data provided by the USBE to evaluate the kindergarten literacy outcomes for the Cohort 13 UPSTART students and their non-program peers. We examined KEEP literacy data for 11,765 UPSTART students and 32,034 non-program students at entry and 11,740 UPSTART students and 32,722 non-program students at exit⁷. Assessments were conducted by kindergarten teachers at the beginning and end of the kindergarten year.

Kindergarten Measures. Our kinder analysis used *Utah's Kindergarten Entry and Exit Profile (KEEP)* to measure the achievement outcomes among all kindergarten students. Like the PEEP, the KEEP consists of two subscales and produces a total score for literacy and a total score for numeracy. All question areas are presented in **Table 9**.

Table 9. KEEP Exit Profile

Literacy (Score range 0-118)	Numeracy (Score range 0-55)
Listening Comprehension	Oral Counting
Phonemic Awareness	Counting Objects
Phonics and Word Recognition	Identify and Compare Numerals
Emergent Reading	Decompose Numbers
Encoding	Decompose Numbers
Writing	Compose and Decompose Teen Numbers
	Addition and Subtraction- Word Problems
<i>KEEP Exit Profile updated October 2018</i>	

⁷ Cohort 13 sample sizes are based on those identified in both the state SIS dataset and in the KEEP dataset for the 2022-2023 school year.

Students who take the KEEP are grouped into three levels; Level 1, “does not demonstrate prerequisite knowledge and skills,” Level 2, “has minimal prerequisite knowledge and skills” or Level 3, “has sufficient prerequisite knowledge and skills.” Similar to PEEP, the performance classifications are considered separate for literacy and numeracy⁸.

KINDERGARTEN ACHIEVEMENT STUDY RESULTS

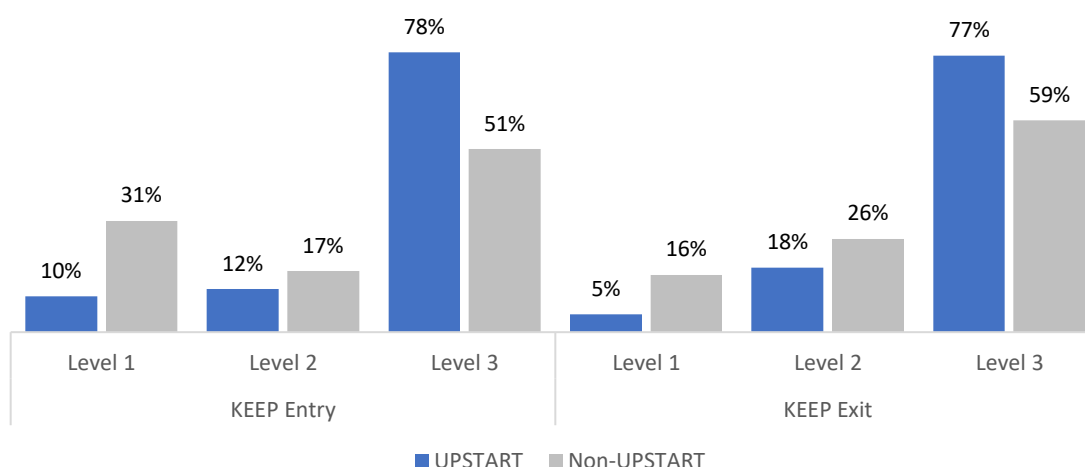
KEEP Entry and Exit Performance Levels

How did the literacy KEEP performance levels for kindergarteners who participated in the UPSTART program at pre-k, compare to the pattern of performance for Non-UPSTART students?

As shown in **Figure 10**, 78% of the Cohort 13 UPSTART students (blue bars) began their kindergarten year with sufficient prerequisite literacy knowledge as measured by the KEEP entry (vs 51% of non-program students). They maintained this level of proficiency through the completion of the kindergarten year (77%) as their non-program peers played catch-up (59%). While a certain level of learning would be expected for all students in kindergarten, the UPSTART children were associated with higher scores at KEEP exit. It should be noted that we did not control for differences in student characteristics and thus this analysis was descriptive in nature.

⁸ See Appendix B for more detail on KEEP performance level descriptions.

Figure 10. Program and Non-Program Literacy Performance Levels at Entry and Exit

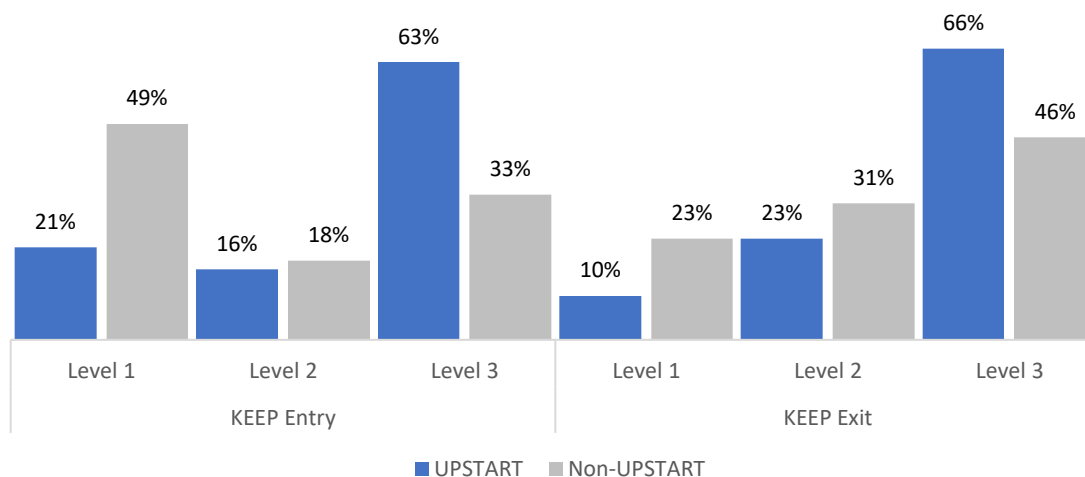


Note: Cohort 13 KEEP Entry sample size =11,765; KEEP Exit sample size =11,740. Non-UPSTART KEEP entry= 32,034; Non-UPSTART KEEP Exit= 32,722

How did the literacy KEEP performance levels for low-income kindergarteners who participated in the UPSTART program at pre-k, compare to low-income students who did not?

We conducted a separate analysis of literacy outcomes of kindergarteners identified as low-income at the beginning and end of the school year. In **Figure 11**, we see that two-thirds (63%) of low-income pre-k UPSTART students started the kindergarten year with sufficient knowledge and skills, twice the proportion of low-income students who did not participate in UPSTART at pre-k (33%). When we examined the end of year literacy outcomes among economically disadvantaged students, 66% of pre-k Upstart students and just 46% of non-program students achieved sufficient prerequisite knowledge and skills in literacy by the end of kindergarten. Again, higher levels of proficiency were expected from beginning to end of year for all students, however for these low-income and likely at-risk students, those participating in UPSTART appear to fare better than their non-program peers.

Figure 11. Low-Income Program and Non-Program Students Literacy Performance at Entry and Exit



Note: UPSTART n= 2,453, Non-UPSTART n=11,356

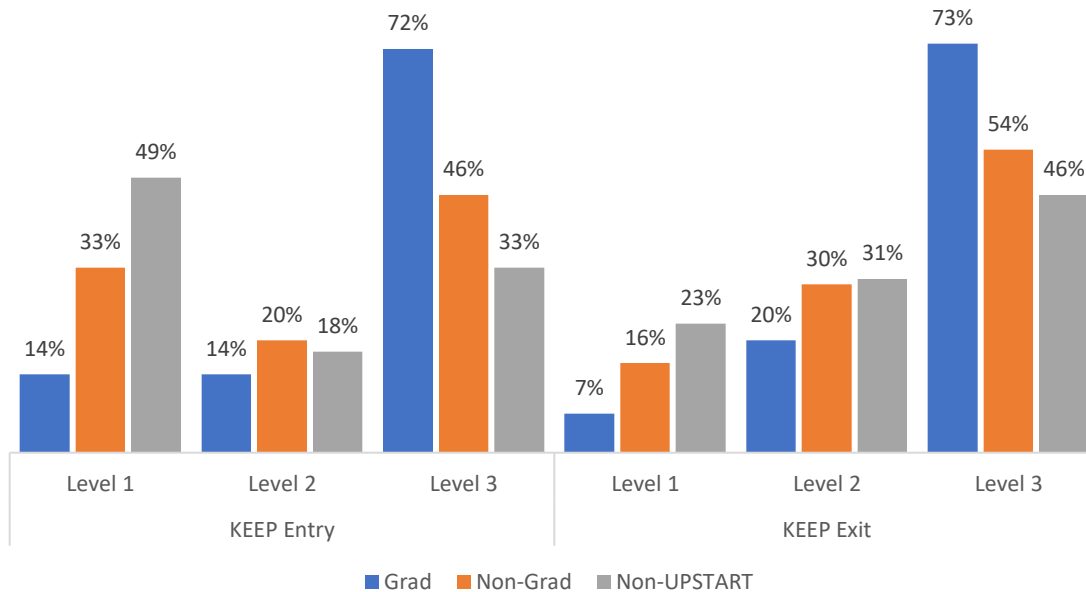
We took the analysis one step further and compared low-income non-program students to two different types of UPSTART students, those who actively participated in the program and met the graduation requirement, and those that used the program less consistently and hence did not meet the graduation criteria. The guiding research question was:

How did the KEEP literacy performance levels for low-income kindergarteners compare among the following three groups, (1) those who graduated from the UPSTART program at pre-k, (2) those who participated but did not graduate, or (3) those who did not participate at all?

Figure 12 depicts our findings of the three groups of low-income students. A greater percentage of low-income students who participated in UPSTART at pre-k (blue bars), even when the graduation criteria was not met (orange bars), achieved Level 3 at KEEP entry and exit compared to students not participating at all (gray bars). As shown, 72% of low-income program grads and 46% of program non-grads started their kindergarten year with sufficient literacy knowledge and skills compared to just a third of non-program students. By the end of kindergarten all students demonstrated gains in skill development (as would be expected), however the pre-k UPSTART

students, and most notably the graduates (73%), show a clear advantage in literacy skills over their non-program peers (46%).

Figure 12. Low-Income Graduates, Non-Graduates, and Non-Program Students Literacy Performance at Entry and Exit



Note: C13 Graduate sample size = 1,569; C13 Non-Graduate Sample size = 884; C13 Non-UPSTART sample size = 11,356.

Both low-income analyses suggest the potential benefit that the pre-k UPSTART program may provide at-risk students into and throughout their kindergarten year.

SUMMARY & RECOMMENDATIONS

Cohort 14 Program Implementation

During the 2022-2023 program year, 70% of all C14 program students met the criteria for program graduation. This rate is slightly down from last year (C13; 73%) and was based on two main criteria, (1) meeting or exceeding a total of 1500 minutes of program usage, or (2) achieving a high score on the WACS. Nearly all of the children who earned their graduation status, did so by exceeding the usage requirements set forth by Waterford, with a small minority of children

graduating strictly based on their achievement score. Families with children who did not graduate from UPSTART tended to have lower levels of parental education, more poverty and were members of underrepresented ethnic groups. Graduation is an important program goal, and families with these risk-related characteristics may benefit from extra program resources to help them achieve it.

Achievement Outcomes

Our pretest/posttest analysis found that at the end of the UPSTART program, most preschool students were considered proficient in literacy and numeracy and thus ready for kindergarten. As we followed the previous year's cohort through kindergarten, we also found descriptive evidence that UPSTART students continued to fare better all the way through the kinder year and observed particularly encouraging results among the state's low-income families.

Study Limitations

We have identified possible limitations to consider when interpreting the evaluation results, namely, the research design and the measures used.

Research Design Limitations. The Utah State Board of Education requested that ETI conduct a program group (only) pretest-posttest research design to study the PEEP proficiency levels of pre-k students in the UPSTART program. This type of observational research design does not support direct causal statements about the impact of the program on students' literacy and numeracy outcomes beyond the descriptive analysis presented here, which is why we report on the "association" or "relationship" between program participation and student achievement outcomes.

Measurement Limitations. We had limited information about the PEEP's psychometric properties, as well as scoring procedures that would go beyond raw summative scores (such as age standardized scoring or other methods). In consultation with the USBE we used raw summative scores for treatment students in our analysis, which are not standardized by age. This limited our

ability to interpret the PEEP mean scores and thus the performance level outcomes for the pre-k study.

Research and Evaluation Recommendations

There are several things that we would recommend for improving future evaluations, based on lessons learned over recent evaluation years:

- Quasi-experimental, matched treatment and control group research design should be used for future preschool program evaluations so that causal statements about program effects on student academic achievement can be made.
- Adequate information is needed regarding the PEEP/KEEP (or any assessment adopted for use by the state), including psychometric properties, and possible score transformations such as standardized by age.
- Subgroup analysis with KEEP data suggest that the UPSTART program may be especially beneficial for the literacy skill development of low-income students, providing a strong footing for continued proficiency in literacy through their kindergarten year. Future research should explore this trend in more detail.
- Future work may consider examining potential longer-term benefits of the UPSTART preschool program into early elementary years. Leveraging the state-wide averages for assessment data creates a viable comparison point given the large sample sizes available.

Program Recommendations

As UPSTART's reach within the state grows, it is important to monitor program implementation so that increased enrollment does not erode graduation or usage rates, two key areas for ensuring strong student literacy and numeracy achievement and future program success.

Specifically, we recommend that the program vendor consider improving strategies for addressing

usage challenges and graduation rates among the most at-risk students. Encourage and promote consistent usage as the building blocks for reaching the graduation requirement. We look forward to studying how new partnerships and outreach in the school districts and communities may yield new ways to support the upcoming cohorts of the UPSTART program.

Appendix A. PEEP Performance Levels

The USBE created PEEP performance levels to help further identify student needs. At the PEEP Entry, literacy and numeracy scores fall into two performance levels. For both subscales, students who score within the performance level 1 range (0-23 points) are classified as “support needed” and demonstrate limited prerequisite knowledge and skills in literacy and numeracy. Students who score within the performance level 2 category are considered to be at benchmark. **Table A1** provides more information on the skills associated with each performance level.

Table A1. PEEP Entry Performance Levels

Subscale	Performance Level	PEEP Score Range	Level of Support	Associated Skills
Literacy	Performance Level 1	0-20 Points	Support Needed	Identifies few to no correct objects in the picture, Demonstrates few to no concepts of print, Identifies few to no letters of the alphabet, Draws few to no lines or shapes correctly, Can produce the first part of few to no compound words, Associates few to no sounds to the letters of the alphabet, Responds to a question in an incomplete sentence or off- topic or does not respond
	Performance Level 2	21-31 points	At Benchmark	Identifies most correct objects in the picture, Demonstrates concepts of print, Identifies some letters of the alphabet, Draws most lines or shapes correctly, Can produce the first part of most compound words, Associates some sounds to the letters of the alphabet, Responds to a question on-topic and speaks in complete sentences
Numeracy	Performance Level 1	0-23 Points	Support needed	Identifies few to no shapes, Counts few to no numbers from 1 to 10 Recognizes few to no numbers from 0 to 5, Identifies few to no groups that have more objects
	Performance Level 2	24-32 points	At Benchmark	Identifies all or most shapes, Counts most numbers from 1 to 10, Recognizes most numbers from 0 to 5, Counts objects using one-to-one correspondence up to 5, Demonstrates emerging cardinality in counting, Sorts objects correctly by given category, Duplicates and/or extends simple patterns, Identifies all or most groups that have more objects

At the PEEP exit, literacy and numeracy scores are divided into three performance levels. For both subscales, students who score within the performance level 1 range at the exit have not

demonstrated prerequisite knowledge and may require significant intervention. Students who are within the performance level 2 range (Literacy - 26-46 points, Numeracy- 18-28 points) have minimal prerequisite knowledge and may require some intervention in order to succeed. Students with the highest literacy and numeracy scores (literacy- 47-108 points and numeracy- 29-36 points) have demonstrated sufficient knowledge and skills and may require minimal intervention to succeed. **Table A2** provides more information on the skills associated with each performance level.

Table A2. PEEP Exit Performance Levels

Subscale	Performance Level	PEEP Score Range	Level of Support	Associated Skills
Literacy	Performance Level 1	0-25 Points	Significant Intervention	Names few to no objects/actions in a picture, identifies no letters of the alphabet, writes few to no letters, isolates few to no first sounds, associates few to no sounds to the letters of the alphabet, demonstrates few to no concepts of print
	Performance Level 2	26-46 points	Some Intervention	Names some objects/actions in a picture and describes it with a phrase(s), identifies few letters of the alphabet, writes some letters, which may include those in their name, isolates some first sounds, associates some sounds to the letters of the alphabet, demonstrates some concepts of print
	Performance Level 3	47-108 Points	Minimal Intervention	Names most objects/actions in a picture and describes it with a sentence(s), identifies some letters of the alphabet, writes his or her name and many letters, isolates many first sounds, associates many sounds to the letters of the alphabet, demonstrates concepts of print
Numeracy	Performance Level 1	0-17 Points	Significant Intervention	Counts few to no numbers from 1 to 20, identifies few to no numbers between 0 and 10, identifies one to no groups that have more objects, draws one to no shapes
	Performance Level 2	18-28 points	Some Intervention	Counts some numbers from 1 to 20, identifies some numbers between 0 and 10 counts some objects using 1-1 correspondence associates numbers to small quantities less than 5, identifies some groups that have more objects draws some shapes
	Performance Level 3	29-36 points	Minimal Intervention	Counts many numbers from 1 to 20, identifies many numbers between 0 and 10, counts objects using 1-1 correspondence, associates numbers to large quantities less than 10, identifies groups that have more objects, draws most shapes

Appendix B. KEEP Performance Levels

Similar to the PEEP, KEEP performance levels were created to help further identify student needs. Though the score range changes from entry to exit, literacy and numeracy scores fall into three performance levels. At entry and exit, students who score within the performance level 1 range have not demonstrated prerequisite knowledge and may require significant intervention. Students who are within the performance level 2 range have minimal prerequisite knowledge and may require some intervention in order to succeed. Students with the highest literacy and numeracy scores have demonstrated sufficient knowledge and skills and may require minimal intervention to succeed. **Tables B1 and B2** provide more information on each level and the different score ranges at entry and exit.

Table B1. Literacy KEEP Performance Levels

Subscale	Performance Level	KEEP ENTRY Score Range	KEEP EXIT Score Range	Level of Support
Literacy	Performance Level 1	0-25 points	0-68 points	Significant Intervention
	Performance Level 2	26-46 points	69-96 points	Some Intervention
	Performance Level 3	47-108 points	97-118 points	Minimal Intervention

Table B2. Numeracy KEEP Performance Levels

Subscale	Performance Level	KEEP ENTRY Score Range	KEEP EXIT Score Range	Level of Support
Numeracy	Performance Level 1	0-17 points	0-32 points	Significant Intervention
	Performance Level 2	18-28 points	33-44 points	Some Intervention
	Performance Level 3	29-36 points	45-55 points	Minimal Intervention



Evaluation and Training Institute
100 Corporate Pointe, Suite 387
Culver City, CA 90230
www.eticonsulting.org

For more information on the
Evaluation and Training Institute, contact ETI:

Jon Hobbs, Ph.D., President
Phone: 310-473 8367
jhobbs@eticonsulting.org