HQSR-E Key Findings Summary

This memorandum was prepared to share key findings from the High-Quality School Readiness Expansion Program (HQSR-E) evaluation with stakeholders ahead of the full evaluation report (to be submitted in October 2017). The Utah State Board of Education (USBE) contracted with the Evaluation and Training Institute (ETI) to evaluate the impacts of the HQSR-E on preschool children's school readiness. The goal of the Year 1 evaluation was to begin developing an understanding of how the program works across three high-quality preschool program implementation models: in public preschool settings, private preschool settings and through an at-home, computer administered software program (known as UPSTART: "Utah Preparing Students Today for a Rewarding Tomorrow"). <u>It is important to note that the Year 1 findings are the beginning of the story (not the end), and that our evaluation was designed to look at the program impacts across three years, using three cohorts of students.</u> Where applicable, children's outcome scores from each program model were compared to scores from children who were not in a high-quality preschool setting ("control group"). These findings were based on data collected from Cohort 1 students, during the 2016-2017 preschool year. Key findings were organized by research question followed by a brief description of the research methods and limitations to the evaluation.

Findings

SUMMARY OF KEY FINDINGS

✓ **Literacy Skill Development: We found different literacy outcome results depending on the HQ preschool program model: students participating in the UPSTART Group had higher rates of growth** on knowledge age equivalency scores, which were used as a benchmark to determine if a child had early literacy skills that contribute to school readiness. In addition, students in the UPSTART Group were the only group with statistically higher literacy test scores than the control group.

Participation in Public and Private Preschool Groups did not make a positive impact on early literacy skills when compared to a control group of children, and, in certain instances, the control group had higher scores on measures of literacy skills than these program children.

Pilot Study:

✓ **Social Emotional Development (SED):** In general, program groups had similar SED growth as the control group. The majority of children from all groups had developed the SED skills appropriate for their age, including relationships with adults, play and peer relationships, and prosocial skills, by the end of their preschool year.

✓ **Math Skill Development:** Children in the control group had higher math test scores on average than children in the Public and Private Preschool Groups, but there were no statistical differences between the UPSTART group and control groups' average test scores.

Literacy

The development of early literacy skills in preschool is a crucial component of school readiness. Children who are behind their peers at entry into Kindergarten might become struggling readers, something which could have a negative impact on their academic success, and poses challenges for educators trying to meet the needs of every student in their class. Our Literacy findings are organized by research questions for clarity and ease of interpretation.

Question 1: To what extent were children ready to learn, based on their literacy skill development, at entry into kindergarten?

To answer this question, we operationally defined students who were ready to learn in Kindergarten as children who had an age equivalency score, or mental age, of at least five years. We chose age equivalency scores of at least five years as our benchmark to reflect the age children enter Kindergarten in Utah. To make it easier for the reader to understand we refer to age equivalency as, "content knowledge or knowledge age," in the remainder of this memo.

Answer:

- Most of the UPSTART Group children (75%) were ready to learn[1] at entry into Kindergarten, while a little more than half of the Public Preschool Group and control group were ready, and 43 percent of Private Preschool Group were ready.

- UPSTART Group children who had a content knowledge age of at least five years old increased by 58% from pre-to-posttest, by 39% for the Public Preschool Group, 23% for Private Preschool Group children, and 40% for control group children.

Table 1. Percent of Children with Content Knowledge Age Scores greater than or equal to an Average 5-Year-Old

UPSTART		Public		Private		Control	
Pre 17%	Post 75%	Pre 17%	Post 56%	Pre 20%	Post 43%	Pre 15%	Post 55%

*Note: Scores were based on the Brigance Literacy Composite

Question 2: How did children's level of learning change from pre-to-post compared to age specific norm group achievement levels?

Children's content knowledge scores can also be used to determine to what extent a child is ready to learn as they matriculate into Kindergarten. One of our literacy measures offers norm group reference data, from which we can calculate if a child was performing at, above or below their normative counterparts at the age of entry into Kindergarten. If a child scores below the average content knowledge for their biological age, they might be at risk of struggling in Kindergarten. The risk goes up for students who performed far below their age equivalent norm group comparisons. For example, if a child is performing at 90% of his/her normed achievement level, he/she is closer to the content knowledge target than a child performing only at 60 or 70%.

---

[1] We defined "ready to learn in kindergarten" as having a content knowledge age greater than or equal to the average five-year-old, which is the age children enter Kindergarten in Utah.

Answers:
- A majority of children across groups were on target, or within 90% of the target to learn, based on their biological age and knowledge age scores: at least 50% of children had content knowledge scores greater than the average for their age group at post-test, and another 17-23% across all groups were within 90 percent of the average.

- The percentage of children with scores higher than their biological age increased from pre-to-post-test by 35% for the UPSTART Group, 27% for the Public Preschool Group, 23% for the control group, and 19% for the Private Preschool Group.

- Children who scored 70% below the average for their age were the most at risk: the UPSTART Group and the control group had fewer than 10% of children in this category at post-test, while the Public and Private Preschool Groups had 17 and 28 percent of children in this range, respectively.

Table 2. Ratio of Content Knowledge Age Range based on Biological Age (BA)

	UPSTART		Public		Private		Control	
	Pre	**Post**	**Pre**	**Post**	**Pre**	**Post**	**Pre**	**Post**
At or Above Biological Age (BA)	30 32%	62 67%	22 31%	42 58%	14 31%	23 50%	34 37%	55 60%
.90 below BA	13 14%	21 23%	10 14%	13 18%	4 9%	8 17%	13 14%	16 17%
.80 below BA	19 20%	4 4%	5 7%	5 7%	3 7%	2 4%	10 11%	13 14%
.70 or below BA	31 33%	6 7%	35 49%	12 17%	24 53%	13 28%	35 38%	8 9%

*Note: Ratio of Biological Age (BA) to Content Knowledge Age=Knowledge Age (KA)/Biological Age (BA)

Question 3: Was enrollment in the HQSR-E program associated with higher scores on measures of early literacy skills when compared to children not enrolled in the program?

Answers:
- Out of the two early literacy measures (Composite Early Literacy Scale and Oral Comprehension Scale), the UPSTART Group was the only program group to produce a statistically significant positive effect.

- The UPSTART Group produced a strong impact on the development of early literacy skills compared to control children (ES: .38) on the Composite Early Literacy Scale, a global measure of alphabet knowledge, phonics, vocabulary, and language concepts.

- The control group children did better than the Private Preschool Group on the PELI oral comprehension composite measure (ES: .57).

Table 3. Post-test Analysis of Literacy Composite Effect Sizes, OLS Regression Model

Construct	UPSTART	Public	Private
Brigance Composite Early Literacy Scale	**.38***	-.14	-.26
PELI Composite Oral Comprehension Scale	-.03	-.25	**-.57****

*Note: $^*p<.05$, $^{**}p<.01$, $^{***}p<.001$; any effect size above .26 is higher than the average effect size seen in similar education evaluations.

Question 4: Did high-quality preschool children have better early literacy skills at entry to kindergarten on specific literacy domains compared to control group children?

Answers:

- There were few statistically significant differences between program and control group children on individual Brigance literacy subscales, with the exception of UPSTART, for which there were medium to strong effects in areas measuring decoding skills (ES=.41 and .51) and letter knowledge (ES=.33).

- Control group children showed statistically stronger outcome scores compared to the Public Preschool Group in visual discrimination (language concepts), and compared to the Private Preschool Group in areas measuring visual and auditory discrimination, alphabet knowledge, and oral comprehension. Effect sizes ranged from medium to large effects.

Table 4. Effect Size Estimates by Literacy Construct and Group

Literacy Construct	Subscale	UPSTART	Public	Private	Instrument
Language Concepts	Visual Discrimination	.24	**-.50****	**-.39***	Brigance
	Auditory Discrimination	.19	-.23	**-.67****	Brigance
Vocabulary & Syntax	Expressive Vocab	.19	-.09	-.33	Brigance
Decoding	Post-primer Vocab	**.51****	-.20	-.06	Brigance
	Sight Words	**.41****	-.09	-.12	Brigance
Letter Knowledge	Letter Sounds	.15	-.15	-.10	Brigance
	Recites Alphabet	-.01	-.11	**-.46***	Brigance
	Letter Knowledge	**.33***	.05	-.02	Brigance
Comprehension	Oral Comprehension	-.03	-.25	**-.57***	PELI

*Note: $^*p<.05$, $^{**}p<.01$, $^{***}p<.001$; any effect size above .26 is higher than the average effect size seen in similar education evaluations.

Pilot Study Findings

A pilot study may be used to test specific components of an evaluation with a smaller sample of participants prior to integrating these components into a larger group or integrating these components into a final evaluation design. We conducted a pilot study of children's social-emotional and math skills development to determine if the instruments used to measure these skills were sensitive to detecting differential effects between the program and control groups, and to determine if integrating these measures were logistically feasible. For example, we wanted to determine if adding these scales would increase the time of the assessment to the point that mental fatigue might impact the findings.

Socio-Emotional Development

Children's social and emotional development are important skillsets for determining if children are emotionally ready to learn and interact with their peers in a school setting. According to the California Department of Education, *"Social-emotional development includes the child's experience, expression, and management of emotions and the ability to establish positive and rewarding relationships with others" (*Early Education and Support Division, 2017). We administered a survey to parents that measured children's social-emotional development in four areas, including areas measuring their interpersonal skills (relationships with adults and peers) and self-regulatory skills (prosocial skills and confidence and motivation).

Question 5: What effects did the program have on the social-emotional development (SED) of program children, when compared to a group of children who did not use the program?

Answers:

- There was little variation between the results of program and control groups on children's SED. In general, parents among all four groups rated the SED skills of their children similarly, with the highest percentage of children having developed age appropriate behaviors related to relationships with adults (96-97%), followed by play and peer relationships (83-88%), prosocial skills (82-86%), and motivation and self-confidence (55-60%).

- In addition, most children across groups developed the socio-emotional development skills appropriate to their age, except for the motivation and self-confidence scale.

- Parents rated their children's motivation and self-confidence higher at pre-test compared to post-test, indicating parents' views might have changed after children's preschool exposure.

Table 5. SED Pre-and-Post Analyses: Percent of Total Skills Developed

Construct	UPSTART		Public		Private		Control	
	Pre	Post	Pre	Post	Pre	Post	Pre	Post
Relationships with Adults	97%	97%	97%	96%	96%	96%	97%	97%
Play and Peer Relationships	83%	88%	79%	85%	79%	83%	80%	85%
Prosocial Skills	79%	86%	86%	86%	81%	82%	78%	83%
Motivation and Self-Confidence	83%	57%	87%	60%	83%	55%	84%	59%

*Note: Percentages equal the average of children's total raw score/total possible score

Question 6: To what extent did children gain SED skills?

Answers:

- Of the children who did not achieve the maximum possible total score at pre-test (e.g. had the opportunity to grow from pre-to-post), a majority experienced positive growth on individual SED items, regardless of the SED measure or group.

- On average, 81-85% of the UPSTART, Public, Private, and Control group children gained additional skills from pre-to-post-test across all SED scales.

Table 6. Children who Gained SED Skills from Pre-to-Post Test

Construct	UPSTART	Public	Private	Control
Relationships with Adults	14 88%	11 85%	12 100%	15 88%
Play and Peer Relationships	53 80%	39 71%	27 82%	49 80%
Interpersonal composite	56 84%	43 77%	27 82%	51 82%
Prosocial Skills	55 83%	43 80%	26 74%	47 77%
Motivation and Self-Confidence	56 84%	38 83%	25 76%	45 83%
Self-regulatory composite	66 92%	53 90%	29 74%	60 88%

*Note: Percentages were based on the number of children who gained skills on each SED scale at post-test, divided by the number of children who were able to gain additional skills after pre-test.

Math

Educators begin teaching foundational math skills in Kindergarten and children must often reach certain benchmarks prior to learning more advanced skills in later grades. Preschool can provide an opportunity for children to get a head start in building these foundational skills, and by helping children understand numeracy, "*provide early childhood educators and elementary school teachers with the tools they need to nurture mathematical thinking*" (Bisanz, J., 2011).

Question 6: Did high-quality preschool children have better early math skills at entry to kindergarten compared to control group children?

Answers:
- Control Group children had statistically stronger development in numeracy skills compared to the Public and Private Preschool Groups on the math composite and on individual subscales.

- There were no statistically significant differences between the UPSTART Group and the control group children in measures of math[2].

---

[2] As a computer based program, not all children received the same curriculum, and children may not have participated in the math component of UPSTART.

Table 7. Effect Size Estimates by Math Construct

Subscale	UPSTART	Public	Private	Instrument
Counts by Rote	0.05	-.47**	-.47*	Brigance
Reads Numerals	0.03	-.38*	-.37	Brigance
Missing Numerals	-0.14	-.33*	-.40*	Brigance
Composite	-0.02	-.47**	-.52**	Brigance

*Note: *$p<.05$, **$p<.01$, ***$p<.001$; any effect size above .26 is higher than the average effect size seen in similar education evaluations.

Evaluation Methods & Limitations

Research Design. We used a repeated measures quasi-experimental research design to measure the school readiness skills of program[3] students (the "treatment group") and non-program students (the "control group") before the program started (the "pretest") and after the program ended (and before students enrolled into Kindergarten; the "posttest"). Our main school readiness measures included several aspects of early literacy, but we also pilot-tested measures of early math learning and social and emotional development.

Measures. We measured early reading, math and social and emotional development[4] using the Brigance Inventory for Early Child Development (IED II and III; Brigance 2010, 2013, respectively) and the Preschool Early Literacy Indicator: Comprehension section (PELI; Dynamic Measurement Group, 2016).

Sample. ETI recruited and tested students in four experimental groups: High-Quality Public Preschool (Public), High-Quality Private Preschool (Private), High-Quality Computer/Home-Based Preschool (UPSTART), and the control students (Control). After removing cases with incomplete or missing data, and cases in which the program or control condition changed (e.g. control children who enrolled in HQ preschool after pre-test), the final analysis sample is listed in the following table:

Experimental Group	Number Tested	Number in Analysis
Public	75	72
Private	58	45
UPSTART	93	92
Control	134	92

---

[3] We defined program students as economically disadvantaged students who attended a high-quality preschool.
[4] Children's social and emotional development were measured using a parent survey.

Analysis. Data were analyzed using STATA v. 14 (StatCorp, 2015), and SPSS v. 23 (SPSS, 2013) to conduct an ordinary least squares (OLS) regression and descriptive analyses of student outcomes.

Limitations. All research projects have limitations, and our study is no exception. Two types of limitations will be reviewed, theoretical and design/methodological, and each should be considered when interpreting our findings.

Theoretical
In theory, we are evaluating three types of preschool program implementation models, and comparing each to what is known as a counterfactual condition (i.e., the absence of the HQSR-E program; our control group); however, in the "real world" we have one narrowly defined program model (UPSTART) and two program models that include a wide variety of differences across preschool curricula, attendance, locations, staffing, resources, mission statements, and so on. By comparing the models, we are abstracting the many components of individual programs into categories (in Public and Private, for example). By abstracting programs the advantage is to take a high-level look at student outcomes, but we cannot tailor our outcome metrics to each individual program.

These theoretical issues also apply to control students (and the counterfactual condition): we do not know what types of preschool activities the control students engaged in during their pre-kindergarten year. We cannot imagine that parents would do nothing with their children during such an important time in their development, but we only know that these children did not attend an HQSR-E preschool. It is possible, however, that these students had equally rich environments that stimulated their academic and social development. It would be too difficult to control for all preschool childhood experiences in our control group, so we are left with little knowledge of what unobserved- but potentially important- activities these control group children engaged in during their preschool year.

Research Design and Methods
Our research design was based on using repeated measures (pretest and posttest) and controlling for pre-existing differences in important covariates (such as baseline levels of achievement, and socio-economic and demographic characteristics) using linear regression as our analysis technique. We ran into two limitations within the context of our design, having small sample sizes (specifically for the Private Provider program group) and ceiling effects for our pilot SED measures.

Small Sample Sizes. With all inferential statistics, having an adequate sample size is important, and we ran a power analysis to determine what sample size (N) we would need to be able to detect small effect sizes. The Private Preschool Group was just below our required N threshold, which was a minimum of 55 cases per group, and this made it difficult to detect small effects in regression models for Private Preschool Group outcomes.

The low evaluation sample size was due to things beyond our control, such as having fewer high-quality private preschools participate in Cohort 1 than expected (only 3 sites), which constrained the number of students we could draw from to begin with. In addition, since it was the first year of the program and evaluation, there was a considerable lag in enrollment and transmitting enrollee's contact information to us. Attrition was relatively high for the Private Preschool Group, and we are adopting new strategies to work with providers to try our best to reduce attrition rates for this group.

Measurement Issues. Measuring developmental and academic skills of preschool students is challenging, and for our Cohort 1 pilot test of SED measures, we ran into "ceiling effects," which makes the measure less useful. Ceiling effects occur when a child (or parent, in the case of parent rating scales) scores at the top of a scale at pretest, and they do not have much room for growth in their posttest scores. This was prevalent in the parent SED rating scales, and the majority of parents reported that their child had met most (if not all) developmental skills listed in our measure at pretest. There are many reasons why this might have occurred, but two possibilities are: parents may have over-estimated their child's level of SED (bias), or their reports were accurate and the children in our samples had healthy and robust socio-emotional development. We suspect that most parents are being accurate, and that our SED test did not have enough response options per item to discriminate well. In other words, the scale was maxed out at pre-test due to normal development and children with healthy SED. As a pilot study, this is an important finding, and we have adjusted the scale response options accordingly for our Cohort 2 testing going forward.

References

StataCorp. 2015. Stata Statistical Software: Release 14. College Station, TX: StataCorp LP.

IBM Corp. Released 2013. IBM SPSS Statistics for Mac, Version 22.0. Armonk, NY: IBM Corp

Bravo Aguayo, K., Abbott, M., and Kaminski, R. (2016). *Preschool Early Literacy Indicators: Time for Bed*. Eugene, OR: Dynamic Measurement Group.

Bader, L. A. (2009). *Bader Reading and Language Inventory* (6th ed.). New York, NY: Pearson.

Bisanz, J. (2011, May). Numeracy. In Encyclopedia of Early Childhood Development. Retrieved September 14, 2017, from http://www.childencyclopedia.com/numeracy/introduction

Brigance, A. (2010). Inventory of Early Development II - Standardized (IED-II). North Billerica, MA: Curriculum Associates.

Brigance, A. (2013). Inventory of Early Development III - Standardized (IED-III). North Billerica, MA: Curriculum Associates.

Early Education and Support Division. (2017). Social-Emotional Development Domain: California Infant/Toddler Learning & Development Foundations. In California Department of Education. Retrieved from http://www.cde.ca.gov/sp/cd/re/itf09socemodev.asp