Studying Pre-kindergarten Upstart in Rural Great Plains States: Impact, Implementation, Cost-effectiveness and Scale-up findings

L. Jon Hobbs

Evaluation and Training Institute

Allison Coordt

Evaluation and Training Institute

October 20, 2023

Authors Note

This study is supported by a grant from the United States Department of Education, Education Innovation and Research (EIR) Grant # U411A180001. The content is solely the responsibility of the authors and does not necessarily represent the views or policy of the U.S. Department of Education.

Correspondence concerning this paper should be addressed to L. Jon Hobbs, Evaluation and Training Institute, 100 Corporate Pointe, Suite 387, Culver City, CA 90230. Contact: jhobbs@eticonsulting.org.

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.

Abstract

The Upstart Great Plains TASK Force: Taking All to Success in Kindergarten was a five-year Education, Innovation and Research (EIR) expansion grant (U411A180001) funded through the US Department of Education Office of Elementary and Secondary Education. Four research studies were conducted during the grant period of 2018 to 2023: a Student Impact Study, a Program Implementation Study, a Program Scale-Up Study and a Program Cost Effectiveness Study. The Student Impact Study was conducted to assess the effectiveness of the Upstart school readiness program on pre-kindergarteners' emerging literacy and social-emotional skills using an RCT research design. Results revealed significantly higher group mean scores on measures of early literacy following the pre-kindergarten program year for the treatment condition compared to the control condition. The Program Implementation Study was conducted to determine if key program activities were achieved during the grant as planned. The Scale-Up Study assessed the overall plan to expand and develop funding, partnerships and resources required to continue Upstart's mission and serve a broader population after the grant was over. The Program Cost Effectiveness Study was conducted to calculate average costs-per-student for standardized effect sizes across multiple early literacy domains.

Forward

This document contains the final reporting for the EIR funded grant, *The Upstart Great Plains TASK Force: Taking All to Success in Kindergarten*. Several studies were included in the five year lifecycle of this expansion grant beginning in 2018 and ending in 2023. The studies are represented here as standalone chapters, namely a Student Impact Study, Program Implementation Study, Scale-Up Evaluation Study and Cost Effectiveness Study. Below is a brief introduction to the main chapters of this report.

Chapter 1: The Impact Study was conducted to assess the effectiveness of the Upstart school readiness program on pre-kindergarterns' emerging literacy and social-emotional skills. The study used an RCT design. Participating pre-kindergarten families were randomly assigned to the treatment condition (computer-based school readiness program that focused on reading) or a comparison condition (computer-based school readiness program that focused on Math/Science). Measures of early literacy skills and social-emotional learning (SEL) were compared between the treatment and comparison groups. Data were collected for three cohorts of students enrolled in the program during different years:

COHORT	PRE-TEST	POST-TEST	DELAYED POST-TEST
1	Summer 2019	Summer 2020	EOY Kindergarten 2021
2	Summer 2020	Summer 2021	EOY Kindergarten 2022
3	Summer 2021	Summer 2022	

The intervention was implemented during the time between the pre-test and post-test. The complete report, including research questions, detailed methodology, and literacy outcomes, can be found in Chapter 1 of this report.

Chapter 2: The focus of the Program Implementation Study was to determine if the key activities of the program, such as provisioned technology, appropriate use of the curriculum, and

parent support were delivered to the participating families as planned. Studying program implementation was important to provide information about the key program components and the extent to which the implementation was executed with fidelity. In accordance with the federal guidelines, fidelity was defined as "adherence to the planned delivery of the key program activities that were hypothesized to lead to the targeted outcomes of the program," (Wolf et al., 2023). Chapter 2 presents the findings from the Upstart program implementation spanning three program years (2019-2020), (2020-2021) and (2021-2022), including information about the cohorts and data sources used for assessing the success of the program.

Chapter 3: The Scale-Up Study assessed the overall plan to expand and develop funding, partnerships and resources required to continue Upstart's mission and serve a broader population after the grant is over. *The Upstart Great Plains TASK Force* project had several big picture scale up goals beyond strictly serving students. Establishing a partnership network of support was critical for Waterford to exceed yearly recruitment targets and serve more students throughout the five partner states. Chapter 3 presents the findings from a scale up analysis, assessing the grantee's ability to create a foundation within the EIR states from which the Upstart program could continue beyond the federal grant funded years.

Chapter 4: A Cost Effectiveness Study was conducted to determine the interventions's cost-per-student and the relationship between the net costs and the effect on student literacy achievement, as measured across multiple literacy outcome variables. This analysis was designed to help identify strategies for minimizing cost while maximizing students' early literacy gains. This chapter contains a brief description of the types of activities included in program costs, a summary of cost-effectiveness by learning domain, and details about the calculations related to the cost-effectiveness estimates.

Table of Contents

List of Tables

List of Figures

CHAPTER ONE

EIR Great Plains Task Force Grant

Upstart Program Effects on Pre-Kindergarteners' School Readiness Impact Study

Abstract

Upstart is a computer-based learning program used to develop the school readiness skills of young children across the U.S. In five Great Plain states, researchers used a randomized control trial design to examine the impact of the program on advancing children's early literacy and social-emotional skills. Pre-kindergarteners in the treatment group were randomly assigned to the Upstart Reading program, while control group students were assigned to the Upstart Math and Science program. Children in the treatment group also received curriculum on social-emotional learning (SEL). Standardized early literacy and math assessments were administered prior to program commencement, at the completion of the program and then again one year later. Social-emotional learning (SEL) was also measured. Results revealed significantly higher group mean scores on measures of Reading Comprehension, Letter and Word Recognition, and Phonological Processing following the pre-kindergarten program year between those randomly assigned to Upstart Reading compared to those assigned to Upstart Math, after controlling for pre-test scores. Results also showed significantly higher math outcomes among those randomly assigned to Upstart Math and Science compared to those in Upstart Reading. There were no differences between treatment and control groups on assessments measuring social emotional learning, or on any of the outcome measures at the end of kindergarten.

Impact of Upstart Participation on Pre-Kindergarteners' School Readiness

The importance of formally preparing children for kindergarten is reflected now in 44 states that fund formal pre-kindergarten programs (Friedman-Krauss et al., 2022). The expansion of pre-k programs is encouraging for those focused on early childhood education as a critical steppingstone to school readiness. The call for greater pre-k access begins at the education policy level, yet we know that not all states will proactively develop statewide legislation prioritizing the funding of early education programs. Some of the country's more rural states, for example, do not fund state-wide pre-k programs; instead, they prioritize education investments in K-12 over public preschool programs (Miller & Votruba-Drzal, 2013; Showalter et al., 2017, Friedman-Krauss, et al., 2022). A growing body of work on school readiness shows that students who arrive at kindergarten prepared and ready to learn in a classroom, have an easier transition and better chance for later academic success (Jozsa, Amukune, Zentai, & Barrett, 2022; Burchinal, Magnuson, Powell, & Hong, 2015).

The expectations for becoming 'ready to learn' are complicated by disparate access to early learning programs and by the continuous remodeling of what is considered a critical skillset for school readiness. The main building blocks of kindergarten readiness typically contain early literary skills, early math skills and age-appropriate social emotional abilities. Exposure to and purposeful practice with these primary skills, pave the way for a smoother transition into the K-12 system. Access to quality early education programs for families with pre-k students is therefore central to this skill acquisition. These programs often aim to address the educational needs of underserved students, close the school readiness gap, and instill supports to shore up competencies for successful early achievement in the K-12 system. Researchers and practitioners

alike emphasize the critical role that high-quality and accessible early childhood education programs can play in preparing children for educational success.

Experts call the early childhood period between pre-kindergarten and third grade "the tipping point" (Atchison & Diffey, 2018)—the most opportune time to eradicate early achievement gaps. Over time, these gaps not only widen but also harden, resisting intervention and possibly stifling children's potential. If children struggle to read at grade level at the end of third grade, they are four times more likely to drop out of high school; add poverty to this achievement gap and dropout rates multiply by 13 times (Atchison & Diffey, 2018). Though more research is needed, evidence for high quality preschool programs has shown great promise in reversing this trend and helping young, underserved children succeed (Phillips, et al., 2017).

A Rural Reality for Early Education

Despite the growing knowledge and intentional efforts behind school readiness programs, the lack of opportunity and accessibility challenges are all too real for many families with 3- and 4-year-olds. According to census data, preschool enrollment was just over 50% nationally in 2021, yet systematically lower for children living in one of the EIR states (U.S. Census Bureau, 2022). This access gap is especially devastating for rural children who are among the most difficult to serve and one of the most underserved populations. Preschool capacity in rural areas, often lags when compared to preschool capacity in larger metropolitan areas (Allard, 2019).

It is no surprise that children in rural areas begin kindergarten with lower levels of school readiness than their more urban peers. In fact, on measures of reading and math, "rural children lag about 2–3 points (or .20 of a standard deviation) behind children living in small urban and suburban areas" (Miller & Votruba-Drzal, 2013). Barriers to preschool access vary by geographic region, but in rural areas these obstacles often include transportation, qualified

teacher workforce, and facilities in distant and remote locations. As a result, high-quality in-person preschools in these states can be prohibitively expensive for statewide implementation and often have an admission waitlist (Charlson, 2017). Both lack of open preschool spots and preschool location are often cited as a barrier by rural parents (National Center for Education Statistics, 2023). Furthermore, existing programs are both geographically and financially impractical for many rural families, especially when participating at their own expense.

Predominant rural conditions—lower property values, smaller tax base, lower levels of parental education attainment, larger families, limited access to educational programs, and lower expectations for academic achievement—influence state, school, and family investments in educational services and supports. Consequently, rural schools receive just 17% of state education funding, despite educating a quarter of the country's children (Showalter, et al., 2017). Families residing in rural communities demand an approach tailored to increasing alternatives specific to their rural needs and one that builds capacity to support young learners for the long-term, all while keeping available state revenue and cost issues in mind. A viable option for our most rural students who cannot get to preschool, may require bringing pre-k programs to the students.

Bridging Access with Educational Technology

Educational technology offers an opportunity for underserved children to receive early education instruction regardless of geo-location or existing pre-kindergarten challenges. We know that internet access in homes and educational settings has continued to increase nationally from 81% to 97% from 2010 to 2018 (Hussar et al., 2020), and we are seeing the trend spread into our most rural communities. Further, the passing of the Bipartisan Infrastructure Law (2021) promises to deliver $65 billion in funding to build out broadband infrastructure in

unserved and underserved areas. Despite a lag behind their urban and suburban peers, a recent study found that approximately 72% of rural residents surveyed in early 2021 reported that they had access to a broadband internet connection at home as well as dramatic increases in access to technological devices, like computers, tablets, and smartphones (Vogel, E., 2021). As access to highspeed internet and technology grows, so does the abundance of educational software programs targeted to young learners, childcare centers, parents, and schools to support the development of children's early literacy and math skills (Wood et al., 2012).

The pandemic brought educational technology to the forefront of instruction across all grades, making it a mainstay during school closures and other restrictions to in-person learning. Rural communities benefitted from online solutions to educational content and school readiness preparation because they eliminated the most common concerns that families face when considering a pre-k program, like distance, teacher workforce, enrollment space, travel time, and expense. In addition to the logistical benefits, computer-based instruction serves as a helpful tool in pre-kindergarten curriculum because it often incorporates immediate feedback, visual graphics, and adaptative lesson plans (Macaruso, Hook, & McCabe, 2006; Barron et al., 2011), optimally striking a healthy balance between playful learning and structured lessons and providing individualized programming to meet each child's needs (Rogowsky et al., 2017). A meta-analysis of computer-assisted instruction with beginning readers showed a positive effect on literacy (Blok et al., 2002), and other work demonstrated that a computer-based reading program had significant positive effects on children's early literacy skills, phonological awareness, listening comprehension, and vocabulary knowledge (Abrami et al., 2015).

There are numerous educational software programs to help children prepare for school, and the evidence is rapidly growing for their effectiveness (Daugherty et al., 2014), but research

is still needed on the impact that educational technology interventions have on different subgroups of students, including pre-kindergarteners growing up in rural communities. It is possible that through partnerships with state governments, non-profit agencies, community centers and local internet providers, more rural families can leverage the necessary resources and access to educational technology opportunities for their preschool students.

The emphasis on developing educational technology programs designed to target the needs of rural students, and especially those with gaps in school readiness skills, underscores the need for research on program effectiveness.

The Common Denominators in School Readiness

Studies show that success at school entry is enhanced by programs using high-quality, age-appropriate curricula that target academic areas such as early literacy, math, and social emotional growth (Phillips et al., 2017, Józsa et al., 2022). Recent longitudinal studies have found that pre-kindergarteners who develop early math and reading skills tend to have a smoother academic transition contributing to the early foundation for later more complex skill development (Fyfe et al., 2019; Józsa et al., 2022). A summary of the essential competencies is presented below:

> *Early Literacy*. The process of learning to read is complex and comprised of a series of fundamental building blocks from preschool through kindergarten, (1) the alphabetic principle, (2) phonological processing, and (3) emergent reading comprehension. The pathway to reading begins with the development of these early preliteracy skills.
>
> *Early Math*. Critical early math skills have been the center of increased attention as a critical component of school readiness, like numeral recognition, counting, quantity (more, fewer), shapes, spatial relations, measurement, and patterns -- all important predictors of later academic achievement (Hardy & Hemmeter, 2019; Jordan et al., 2010; Mejias, Muller, & Schiltz, 2019).

Social-Emotional Skills. In addition to the academic building blocks to school readiness, multiple studies have demonstrated how social-emotional development is also a key component (Raver & Knitzer, 2002; Darling-Churchill & Lippman, 2016). Students with strong social and emotional skills are more likely to engage with teachers and peers and experience a positive transition from preschool to kindergarten (Raver & Knitzer, 2002 & Sheridan et al, 2010).

About Waterford's Upstart Program

Waterford is a 501(c)(3) national education nonprofit with a mission to providing high-quality educational resources and early childhood education programming at no-cost for children, families, and communities. Upstart is an in-home computer-based school readiness software program, making the program more accessible without sacrificing high quality and individualized attention to skill development. The Upstart program was launched in 2009 and has since served over 300,000 urban, rural, refugee, and high-need children across 27 states (for more information, see: https://www.waterford.org). Importantly, in certain states, Upstart has become integrated into state legislature to address the growing inequality and inaccessibility of quality education programs for low-income children and families. With Upstart's emerging role in state policies, understanding the effectiveness of the program is crucial for furthering its development, addressing persistent inequalities, and preparing children for early success.

Upstart, the program under study, provided pre-kindergarten children with reading, math/science, and social emotional learning curricula with the primary focus on promoting mastery of early literacy skills. Upstart was unique in its ability to remotely provide individualized lessons and learning content to each participating child. The software utilized a lesson sequencer that adapted to each child's skill level, and continuously adapted its activities based on student performance to best meet the child's current skillset while also supporting their

continuing development. For example, if the child was struggling with a concept or skill, the sequencer would run remedial activities to reteach and practice the skills again; similarly, the sequencer advanced to another objective in the lesson content if students were demonstrating a mastery of concepts. Lesson content was delivered online through the software and utilized a variety of multimedia tools such as digital books, songs, and other online activities to best engage the child.

Upstart Reading. The Upstart Reading program used research-based best practices for early literacy instruction, based on early guidelines outlined by the National Reading Panel (2008). These guidelines emphasized phonemic awareness, phonics, fluency, comprehension, and vocabulary. The Upstart Reading curriculum provided the foundation for skilled reading by emphasizing precursor skills related to decoding and comprehension, two processes that are the hallmark of reading fluency (Hoover & Gough, 1990). After enrolling in the Upstart Reading program, the child was tested and placed in a level, ranging from one to three, based on their performance. Beginning the curriculum at Level One – Pre-Reading, the pre-kindergarten child started the program as a nonreader and was first introduced to the skills of a reader. To foster thorough literacy skill development, Waterford recommended a child used the Upstart software and engaged with the Reading program for at least 15 minutes a day, 5 days a week. **Table 1** showcases the reading domains and skills taught by Upstart Reading at the first level of the curriculum: phonics, comprehension/vocabulary, language concepts, and phonological awareness.

Table 1. Upstart Program Reading Domains and Skills

Upstart Reading Domains	Level 1 Pre-Reading Skill
Phonics Systematically builds from not reading to confident reading at 90 words a minute	• Recognize *A* through Z, and *a* through *z* • Learn 10 letter sounds and 20 sight words to read 10 leveled readers • Spell child's name
Comprehension/Vocabulary Develops vocabulary and critical thinking skills through rich reading experiences	• Read along and understand nursery rhymes • Read along and understand alliterative books • Learn 255 target vocabulary words
Language Concepts Introduces concepts of written language (from letters and pictures to basic grammar)	• Understand print (left-to-right, letters, pictures, words, text) • Develop oral language skills (colors, shapes, numbers, sizes, etc.)
Phonological Awareness Develops awareness of individual sounds in words	• Break words into individual sounds (cat to (/*k*/ /*a*/ /*t*/) • Blend individual sounds into words (/*k*/ /*a*/ /*t*/ to cat) • Change a sound in a word to make a new word (cat to bat)

Upstart Math and Science. Literacy development was the primary focus of the grant; however, the Upstart Math and Science program was used by half of the EIR families, who were randomly assigned to the control group. The Upstart math curriculum[1] was developed to use conceptual math and basic cognitive skills. Conceptual math helps students explore the "meaning of operations, calculators, mental computation, estimation, and thinking strategies." At the same time, basic skills like fact retrieval and drill and practice commit problem-solving strategies to memory ("Waterford Early Learning," 2011). These two philosophies work in tandem to introduce "learned mathematical concepts into real-world situations." ("Waterford Early Learning," 2011). **Table 2** showcases the math domains and skills taught by Upstart math at the

[1] The current study examined literacy, math, and social emotional learning outcomes only, thus the science curriculum was not described in detail in this report.

first level[2] of the curriculum: number & operations, operations & algebraic thinking, measurement & data, and geometry.

Table 2. Upstart Math Program Domains and Skills

Upstart Math Domains	Level 1 Math Skill
Numbers and Operations Teaches number recognition, place value, counting, and arithmetic computation	• Recognize, order, and write numbers 0 through 20 • Order, count, and sequence numbers to 100 by ones and tens • Use strategies to compare group size (more than, less than, or equal to)
Operations and Algebraic Thinking Teaches arithmetic computation	• Use objects, drawing, etc., to represent addition and subtraction • Add and subtract within 10, including solving word problems • Fluently add and subtract within 5
Measurement and Data Develops a foundational understanding of measurement, time, and money. Prepares students to analyze data.	• Compare, classify, and describe measurable attributes of objects • Use digital and analog clocks to tell time to the hour • Identify coins and their value
Geometry Teaches properties of shapes, positioning, and the identification of parts of regions or groups.	• Identify basic shapes regardless of their orientation and environment • Create composite shapes • Learn about shape positioning • Understand similarities and differences in 2- and 3-dimensional shapes

Upstart Social-Emotional Learning. The Upstart program support model and curriculum were designed to enhance young children's development, reinforcing kindergarten teachers' assertion for social- emotional readiness before entering the formal classroom. Social and emotional learning (SEL) was built directly into the curriculum, with additional resources and guidance available to parents via an optional parent engagement curriculum. Guided by and aligned with Waterford's Mindset Skills, Upstart focused on social skill content areas such as

[2] Level One is the beginning point of the curriculum, where the pre-kindergarten child is introduced to skills designed to teach the child mathematics. Levels range from one to three, and the child is tested at the beginning of the program and placed in a level based on performance.

Self Awareness, Social Awareness, Responsible Citizenship and Executive Functioning. These content areas were delivered through video activities, books, and other digital media to model behavior. The key components of the SEL curriculum are displayed in **Table 3**.

Table 3. Upstart Program SEL Domains and Skills

Upstart SEL Domains	SEL Skill
Self-Awareness Demonstrates knowledge of emotions, identity, and self-management	• Identifies personal characteristics, preferences, thoughts, and feelings • Maintains and manages boundaries • Asks adults for help when appropriate
Social Awareness Develops increased relationship skills, ability for empathy, recognize diversity and inclusion	• Demonstrates ability to interact/friendship skills • Recognizes and considers others' feelings • Understands similarities and differences in others
Responsible Citizenship Demonstrates problem-solving and decision making, community building with peers, awareness of digital citizenship	• Solves age-appropriate problems • Practices being a good neighbor and classmate • Demonstrates age-appropriate online behaviors
Executive Function Introduces concepts of memory, mental flexibility, and organization	• Increases in memory abilities • Ability to smoothly transition between tasks • Shows early organization skills

The School Readiness Program Model & Research Questions

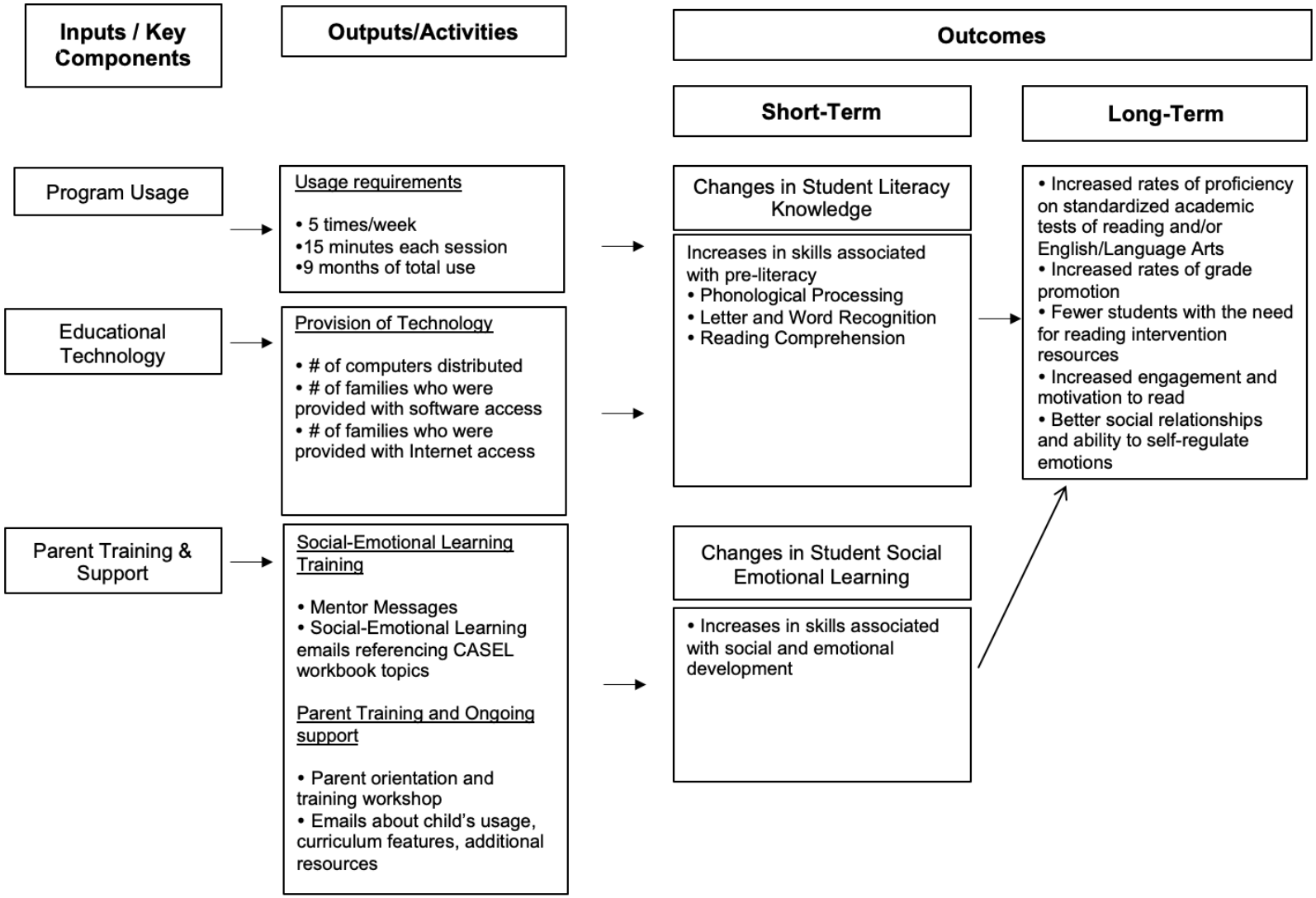
We studied the effectiveness of Waterford's Upstart Reading and Social-Emotional Learning curricula in developing the school readiness skills of rural pre-kindergarten children prior to their entry into kindergarten. The study was based on a logic model (see **Figure 1**), which described the student outcomes that were the specified anticipated changes in participants' behavior, knowledge, skills, and/or attitudes as a result of program exposure (Frechtling, 2017). We distinguished between short and long-term outcomes in the logic model, with short-term

outcomes expected to occur at the end of the Upstart school readiness program and long-term outcomes anticipated in early primary grades.

Short-Term Outcomes. We expected that as a result of participating in the Upstart Reading & SEL program, pre-kindergarteners would have increased skills associated with early literacy and social and emotional development. Literacy skills included letter knowledge (e.g., alphabet sequence, letter names, letter sounds), phonological and phonemic awareness, visual discrimination, and auditory discrimination. Social and emotional skill development was measured across multiple indicators and summed into a composite measure of social and emotional skills. The outcome evaluation assessed the impact of the Upstart project on these short-term outcomes.

*Long-Term Outcomes***.** We expected that children who participated in Upstart Reading would see cumulative benefits that build upon earlier skills once they enter the school environment, such as higher rates of grade promotion and proficiency on standardized tests of reading and/or language arts, along with fewer referrals for reading intervention. Additionally, we predicted that the positive impacts on short-term academic outcomes would influence long-term affective domains, such as a higher engagement with reading and an increased motivation to read.

Figure 1. EIR GPTF Logic Model



Confirmatory Research Questions

The primary research questions included the following confirmatory contrasts between treatment and control students:

1. Did pre-kindergarten children randomly assigned to receive the Upstart Reading & SEL program (treatment condition) during their pre-kindergarten year have higher scores than their counterparts assigned to the Upstart Math and Science program (control condition) on the following measures of school readiness, focused on emerging literacy and social-emotional development:
    a. Phonological processing,
    b. Letter and word recognition,
    c. Reading comprehension, and
    d. Social emotional development.

Hypothesis: If Upstart Reading & SEL influences early literacy and social-emotional skills, then children in Upstart Reading & SEL should perform significantly better than the control group on measures of early literacy and social-emotional learning.

Exploratory Research Questions

1. Did children assigned to receive the Upstart Reading program during their pre-kindergarten year have higher scores at the end of the kindergarten year compared to a control group of children assigned to the Upstart Math and Science program during their pre-kindergarten year on the following measures:
    a. Phonological processing,
    b. Letter and word recognition,

c. Reading comprehension

2. Did children assigned to receive the Upstart Math and Science program during their pre-kindergarten year have higher scores than their counterparts assigned to the Upstart Reading program on the following measures of early math:

    a. Math Concepts and Applications[3]

3. Did children assigned to receive the Upstart Math and Science program during their pre-kindergarten year have higher scores at the end of the kindergarten year compared to a group of children assigned to the Upstart reading program during their pre-kindergarten year on the following measures:

    a. Math Concepts and Applications

Methods

Research Design

We investigated the impact of the Upstart Reading program on improving pre-kindergarten reading outcomes in the year leading up to kindergarten. An intent to treat randomized control trial (RCT) design was used with children from three annual cohorts of program students, each randomly assigned to one of two conditions: an Upstart Reading & SEL program (treatment group) or an Upstart Math and Science program (control group; diagrammed in **Table 4**). Children and families in the treatment group received curriculum designed to support social emotional learning (SEL), a feature that was not implemented with the control group.

[3] The current exploratory research questions targeted math outcomes only as emerging math skills have been identified as important in kindergarten readiness.

Table 4. Evaluation Design

Summer 2019, 2020, 2021			Summer 2020, 2021, 2022		Spring 2021, 2022
Pre-Test	Random Assignment	UPSTART Reading and SEL **Treatment**	Post-Test	Kindergarten	Delayed Post-Test*
		UPSTART Math & Science **Control**			
**Only Cohorts 1 and 2 participated in the third observation*					

Pre-tests were conducted prior to participants' random assignment into conditions to establish baseline levels of the outcome variables, and then followed by two post-test observations, one immediately following the end of the pre-k program and a second at the end of the kindergarten year. A power analysis was conducted to determine the minimum sample size needed for the study to detect program impacts. Using statistical power of 80%, a significance level of .05, and an expected effect size of .18, researchers determined a sample size of 1000 would be sufficient given the parameters. Random assignment was done after families completed baseline assessments.

Procedure

After families enrolled in the Upstart program in the spring of 2019, 2020 or 2021, family contact information was provided to the research team for study recruitment purposes. Researchers contacted families by phone and email and invited them to participate in a school readiness study assessing the effectiveness of Upstart. Qualifying families[4] were offered incentives for taking part in the research, receiving one e-gift card at pre-test, a second one at

[4] Research participant criteria included English speaking and those not receiving special education services.

post-test and a third for participating in the delayed post-test.

In the weeks leading up to each data collection period, the research team recruited and scheduled families for one-on-one assessments with trained test administrators (TA). Each assessment was individually administered, lasted 30-40 minutes on average, and used identical procedures across treatment and control groups as well as observation time periods. Most data were collected using a virtual/online platform, with the exception of Cohort 1 pre-test data, which were collected in-person during summer 2019. All subsequent assessments conducted from 2020 through 2022 were conducted using remote testing procedures developed to address the health and safety regulations (see **Appendix D** for more details). Researchers conducted observations of each TA using a structured protocol to ensure the assessments were conducted with fidelity. A data quality procedure was additionally conducted at three points–beginning, middle, and end–where researchers compared the mean scores of all TAs utilizing bivariate statistical analyses. These observations and data quality procedures were conducted to ensure all staff were implementing the testing protocol consistently and accurately.

Study Participants

The current research study randomly sampled 1,348 families from Idaho, Wyoming, Montana, North Dakota, and South Dakota to evaluate the Upstart program between 2019 and 2022. Among those families, 652 were randomly assigned to the reading program and 696 were randomly assigned to the math/science program. At the completion of the program (i.e., post-test), 1,053 students were tested. To test the exploratory hypothesis regarding enduring treatment effects one year later, (i.e., delayed post-test), two of the three cohorts were tested again at the end of kindergarten (n=645). Families in the study largely mirrored the demographic

make-up of the population of families participating in the Upstart program. **Table 5** shows the balance of demographic characteristics among those randomly assigned to the reading condition or the math/science condition and were also pre-tested and post-tested.

Table 5. Demographic Characteristics of Research Study Sample

Demographic Categories		**Reading (n=506)**	**Math/Science (n=547)**	**Total (n=1,053)**
Child's Gender	Male	253 50%	263 48%	516 49%
	Female	253 50%	284 52%	537 51%
Child's Race/Ethnicity	Caucasian/White	437 86%	459 84%	896 85%
	Latino/a	14 3%	14 3%	28 3%
	Asian/Pacific Islander	6 1%	10 2%	16 2%
	Native American	4 1%	13 2%	17 2%
	African American/Black	7 1%	6 1%	13 1%
	Other	38 8%	44 8%	82 8%
Child's Language	English	494 98%	534 98%	1028 98%
	Other	12 2%	13 2%	25 2%
Parent Educational Attainment	High school graduate	40 8%	39 7%	79 8%
	Some College	160 32%	171 31%	331 31%
	College Graduate	234 46%	244 45%	478 45%
	Advanced Degree	9 2%	19 4%	28 3%
Parent Marital Status	Married	463 92%	493 90%	956 91%
	Otherwise	54 8%	42 10%	96 9%

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. Sample sizes based on available demographic data.

Measures

Two main instruments were used in the assessment of early school readiness skills and social emotional development, the Kaufman Test of Educational Achievement, Third Edition (KTEA-3, 2014) and the Social Skills Improvement System Rating Scales (Gresham & Elliot, 2008), respectively. Both instruments were individually administered at pre-test, post-test and delayed post-test to all treatment and control children.

We assessed children's emerging reading skills as measured by the KTEA-3, across the following domains:

Domain	KTEA Subscale
Phonics and Related Alphabetics	**Letter and Word Recognition** (LWR)- lower and uppercase letter identification, naming letters and corresponding phonetic sounds, and reading basic pre-primer words.
	Phonological Processing (PP)- listen to and manipulate whole words, word parts, and sounds, identify rhyming and non-rhyming words, and hold phonological information in working memory and retrieve that information
Reading Comprehension	**Reading Comprehension** (RC)- basic concepts such as matching pictures with words, reading and following commands, and reading sentences and answering questions

Children's *Letter & Word Recognition* was assessed with a 100-item scale, ranging from 0-100. *Reading Comprehension* was measured using a 68-item scale ranging from 0-68. *Phonological Processing* was measured using a 50-item scale that is the composite score of five subscales: Blending (score range: 0-10), Rhyming (0-8), Sound Matching (0-6), Deleting Sounds (0-11), and Segmenting (0-15). Once raw scores were calculated for each subscale, they were converted to Growth Score Value (GSV) scores. The KTEA-3 has shown good psychometric properties among pre-k students with the reliability ranging from 0.84 for reading comprehension, 0.94 for phonological processing, and 0.97 for letter and word recognition.

SSIS-RS. Parents completed the Social Skills Improvement System Rating Scales (Gresham & Elliot, 2008) as part of a parent survey, targeting their child's behavior during the two months prior to the assessment. The SSIS-RS instrument included seven subscales for measuring different constructs within social-emotional development and a composite score for measuring children's global social skills. The SSIS-RS consisted of 46 items with a score range of 0-138 for the instrument (each item having a possible score of 0-3). Four SSIS-RS subscales, *Communication, Assertion, Engagement,* and *Self-control* were assessed with seven-items and a scale ranging from 0 to 21, respectively. *Cooperation, Responsibility,* and *Empathy* were measured with six-items each, and a scale ranging from 0-18. The raw scores from the seven SSIS-RS subscales were converted to corresponding benchmark levels (Below Average, Average, Above Average). The raw score of the SSIS-RS composite was converted to a standard score and corresponding benchmark level. Conversions for both subscale scores and the SSIS-RS composite were calculated using a set of age-based norms. The SSIS-RS has been documented as reliable and valid among a pre-kindergarten population (Frey et al, 2011). Overall, coefficient alphas for the SSIS-RS Social Skills Scale for the Parent Form are .96, with alpha values ranging from .76 to .88 for each of the subscales.

Data Analysis

Our statistical approach included three phases: a group baseline equivalence analysis, an attrition analysis to determine if sample sizes were within the acceptable attrition boundaries for an RCT set by What Works Clearinghouse Procedures and Standards Handbook (version 5.0; 2022), and ordinary least squares regression (OLS) to test our hypotheses (Cohen, Cohen, West, & Aiken, 2003).

Baseline Equivalence. Baseline equivalency testing was conducted before formal hypothesis testing. Using independent samples *t*-tests, we established equivalence between the reading and math/science groups for all pre-test achievement scores. Children's baseline characteristics were examined to confirm that treatment and control groups were equivalent at the start of pre-kindergarten on factors that may influence school readiness skills, such as pre-test achievement scores and demographic factors. Baseline equivalence results can be found in **Appendix B**.

Attrition Analysis. Ideally, all children who were pre-tested would have been post-tested. However, as in most studies that rely on repeated measures, that ideal is rarely attained. Overall and differential attrition was observed for each research condition and compared to the acceptable levels of bias set by What Works Clearinghouse standards (5.0; 2022). Results for the attrition analysis can be found in **Appendix C**.

OLS Regression. Ordinary least squares (OLS) multiple regression models were conducted to examine the impact of the treatment on reading achievement Growth Scale Value scores (GSVs). The KTEA-3 provides GSVs as a measure of change over time, or growth, to describe the absolute level of performance of a treatment condition as opposed to the relative performance provided by the age-adjusted standardized scores. Comparing GSVs across pre-test and post-test provided information about the magnitude and direction of growth over time among Upstart participants. Regression covariates in the OLS models, included treatment status, baseline achievement scores at pre-test, and a state blocking variable.

Hedges' g standardized effect sizes were calculated to illustrate the magnitude of the difference between the two conditions and compare the size of the impacts across measures. The

effect size estimates presented show the magnitude of the average performance differences in standard deviation units between the treatment group and the control group on outcome measures with statistically significant regression results. Effect sizes were calculated based on the adjusted mean differences between the treatment and control groups divided by the unadjusted pooled standard deviation. Standardized effect sizes are helpful for comparing the magnitude of differences in outcomes based on different measures, but several considerations should be made when establishing a benchmark and subsequently interpreting the magnitude of standardized effect sizes (for a review, see Hill, et al, 2008; Lipsey, et al, 2012 and Kraft, 2020). Determining an average effect size for similar types of research (RCT), studying similar interventions (early literacy programs) with similar populations (preschool) can all be helpful for setting a "benchmark" effect size for comparison (Lipsey, at al, 2012; Kraft, 2020).

We set our benchmark effect size at .20, which represents the average effect size found in a summary of meta-analyses of relevant and similar educational studies (Kraft, 2020; M. Kraft, personal communication, October, 2023). The current study considered intervention effect size estimates greater than the .20 benchmark, as above average and therefore deemed substantive in our interpretation of the findings.

Regression Model Components. We defined the following variables for each pre-kindergarten child in multiple linear regressions to estimate the impact of Upstart reading on our outcome variables of interest: Y_{ij} is the score for child *i* in state-by-cohort block *j* on post-test measures of Reading Comprehension, Phonological Processing and Letter and Word Recognition; Treatment ($\beta_{1UPSTARTReading_{\mathbf{ij}}}$) is an indicator for whether the child received the intervention; $\beta_{2Pretest_{\mathbf{ij}}}$ is the child's score on pre-test measures (pre-test covariate); $\beta_{3State_{\mathbf{sj}}}$

are blocking indicators of the state in which the child resides; and $\beta_{4Cohort_{\mathbf{cj}}}$ are blocking indicators of student cohort. One possible linear regression model that uses these variables is the following:

$$Y_{ij} = \beta_0 + \beta_1(UPSTARTReading_{ij}) + \beta_2(\text{Pretest}_{\text{ij}}) + \sum_{s=1}^{S-1} \beta_{3.s}(State_{sj}) + \sum_{c=1}^{C-1} \beta_4(Cohort_{cj}) + \varepsilon_{ij}$$

The βs in Eq. 1 are regression coefficients that describe the relationship between each variable and the pre-kindergarten's post-test score:

- β_0 is the intercept;
- β_1 is the expected increase in the post-test score for pre-kindergarteners who participated in the Upstart Reading intervention relative to students who participated in the math/science intervention;
- β_2 is the effect of pre-test data; and,
- β_3 is the effect of the state in which the child resides,
- β_4 is the effect of student cohort.

Results

The primary goal of the study was to examine Upstart's impact on children's school readiness skills, namely early literacy, and social emotional learning. Exploratory analyses expanded the focus to early math skill development, as well as the program's ability to show enduring impacts 1 year after participation. The results of the pre-k year intervention for early literacy and social

emotional learning are presented first followed by the exploratory inquires of math outcomes and kindergarten year follow-up results.

Confirmatory Results

Pre-K Year Reading & SEL Outcomes. The predicted mean scores and Hedges' g effect sizes for Reading Comprehension, Letter and Word Recognition, Phonological Processing, and SSIS RS outcomes are presented in **Table 6**. Full regression results can be found in **Appendix A**.

Table 6. Literacy & SEL Regression Analysis Results

	Control Group (Math/Science)			Treatment Group (Reading)						
Measure	Sample Size	Adj. Mean	St. Dev.	Sample Size	Adj. Mean	St. Dev.	Tr – Ctrl Diff	St. Error	Standard Diff[5]	P-value
Reading Comprehension	546	384.23	18.43	507	389.34	18.43	5.11	1.14	0.28	0.000
Phonological Processing	546	465.32	11.10	507	470.52	11.10	5.19	0.69	0.47	0.000
Letter and Word Recognition	546	333.32	25.73	507	348.14	25.73	14.82	1.59	0.58	0.000
SSIS Rating Scales (RS)	546	106.66	0.36	507	106.66	0.38	0.00	0.53	0.000	0.990

[5] Hedges *G* Effect Size: Treatment minus control divided by pooled standard deviation.

Reading Comprehension. Participation in the treatment condition (Upstart Reading) had a statistically significant impact on children's reading comprehension skills ($p = 0.00$). The Upstart Reading treatment group had an average predicted mean score of 389.3 points, compared to the control group's score of 384.2 points, outperforming their control group peers by 5.11 points on a measure of Reading Comprehension while controlling for pre-test scores. A Hedges' g effect size of 0.28 was calculated to contextualize the magnitude of the treatment and control

differences for the Reading Comprehension outcome. Based on the set threshold of 0.20, the Upstart program had a substantive impact on early reading comprehension skills (Table 6).

Phonological Processing. Results displayed in Table 6 indicate that after controlling for pre-test score, random assignment to Upstart Reading had a statistically significant impact on children's Phonological Processing ability ($p = 0.00$). The reading treatment group had an average predicted mean score of 470.52 points, 5.19 points higher than the control group's score of 465.32 points. A Hedges' g effect size of 0.47, exceeded the set threshold (0.20), indicating a meaningful program impact on Phonological Processing skills.

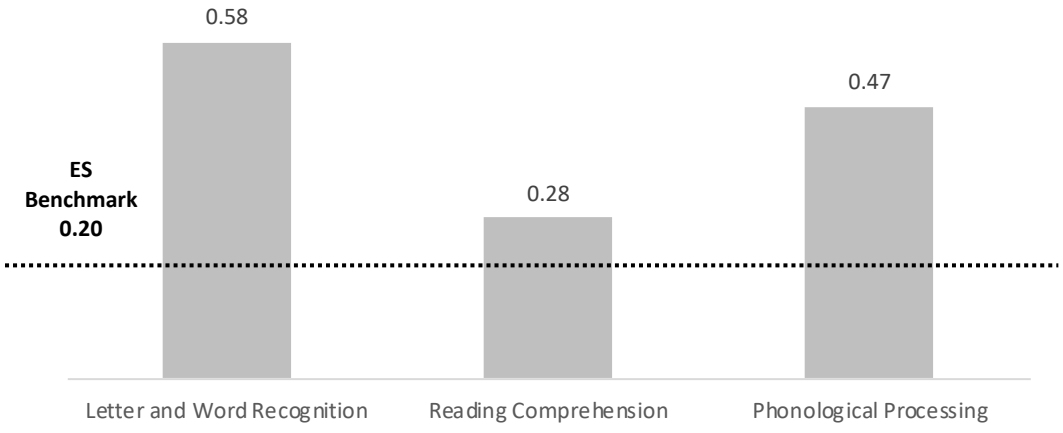
Letter and Word Recognition. Upstart Reading had a statistically significant impact on children's Letter and Word Recognition (LWR) skills ($p = 0.000$) after controlling for pre-test scores. The Upstart Reading treatment group had an average predicted mean score of 348.14 points, 14.82 points higher than the control group's score of 333.32 points. A Hedges' g effect size of 0.58, exceeded the set threshold (0.20), indicating a substantive program impact on Letter and Word Recognition skills (Table 6).

Pre-K Year SEL Outcomes. It was hypothesized that children participating in the treatment group (and therefore exposed to additional social emotional learning curriculum) would show higher SEL outcomes at the end of the program compared to those in the control group. No significant differences were found between treatment and control children on measures of social emotional learning as reported by parents (Table 6).

Effect Size Summary. Based on the effect size benchmark established from similar educational interventions, the Upstart Reading program produced notable impacts on enhancing pre-kindergarten children's reading competencies, including Letter and Word Recognition (g=

0.58), Reading Comprehension (g=0.28), and Phonological Processing (g=0.47) and exceeded the threshold for effective impact across all three literacy outcomes (see Figure 2; no effect sizes are reported for SEL outcomes because no significant mean differences were found).

Figure 2. Effect Size Estimates based on Adjusted Mean Standard Scores by Measure



N=1053; Reading n=507, Math/Science n=546

Exploratory Results

***Kindergarten Year Reading Outcomes*.** As stated previously, the research design included three observations, (1) before the start of the program, (2) at the end of the program, and (3) at the end of the kindergarten year. The first two annual cohorts participated in the third and final observation, referred to as kindergarten year outcomes. As part of our exploratory analysis, we examined if treatment students who used the reading program during their pre-k year continued to show higher scores on measures of early literacy at the end of kindergarten, compared to students who used the math and science program. We did not, however, observe significant differences between reading students and math/science students for any of the reading outcome measures a year following the completion of the intervention.

Pre-K Year Math Outcomes. The primary focus of the current research was literacy development, however with 50% of the participating students enrolled in a math and science intervention, there was an opportunity to additionally look at how the Upstart program impacted early math skill development. Therefore, we explored the hypothesis that students who used the Upstart Math and Science program would score higher on measures of early math than students in the reading condition. We assessed math outcomes at two points in time, (1) at the end of the Upstart program and (2) again at the end of the kindergarten year. Our findings indicate that the math and science program had a significant impact on children's early math skills at the completion of the intervention, as measured by Math Concepts and Applications (see **Table 7**).

Table 7. MCA Regression Analysis Results

	Control Group (Math/Science)			Treatment Group (Reading)						
Measure	Sample Size	Adj. Mean	St. Dev.	Sample Size	Adj. Mean	St. Dev.	Tr – Ctrl Diff	St. Error	Standard Diff*	P-value
Math Concepts and Applications	387	423.71	14.52	387	418.74	15.10	4.97	1.07	0.34	0.000

**Note: Hedges G Effect Size: Treatment minus control divided by pooled standard deviation.*

The Upstart Math and Science group had an average predicted mean score of 423.71, 4.97 points higher than the reading group's mean of 418.74. A Hedges' g effect size of 0.34, exceeded the set threshold (0.20), indicating a substantive program impact on Math Concepts and Application skills. We did not, however, find significant differences in math scores at the end of the kindergarten year between those participating in the pre-k math and science program and those in the pre-k reading program.

Discussion, Research Limitations and Future Research Recommendations

More and more studies are showing that students who arrive at kindergarten prepared and ready to learn, transition more smoothly and may boost their chances for later academic success (Jozsa, Amukune, Zentai, & Barrett, 2022; Burchinal, Magnuson, Powell, & Hong, 2015), yet children in rural areas typically begin kindergarten with lower levels of school readiness than their urban peers. We hypothesized that children participating in the Upstart Reading program during their pre-kindergarten year would perform significantly better than a control group of students who did not receive a reading intervention (children participating in the math/science software) on measures of early literacy and measures of social-emotional development.

Children participating in the pre-kindergarten reading program demonstrated higher scores on all early reading outcome measures, compared to the control group. Treatment effect sizes (Hedges'g) exceeded our threshold to indicate substantive and meaningful changes to students' school readiness. Similarly, as part of our exploratory analysis, we examined if students who used the Upstart Math and Science program scored higher on measures of early math compared to students in the reading program. Our results revealed that math/science students significantly outperformed their reading program peers on measures of early math at the end of the pre-k year. As with reading, the treatment effect size for the math/science intervention also exceeded the threshold for substantive impacts. This finding was exploratory in nature, and not a confirmatory contrast or focus of the study, but no less significant given half the research sample was randomly assigned to the math/science intervention. Two of our confirmatory hypotheses were not supported by the research results: 1) there were no significant differences between the treatment and control groups on measures of social emotional learning outcomes at

the end of the preschool year; and 2) there were no differences at the end of kindergarten (i.e. one year out from the preschool-year intervention) in either reading or math outcomes.

The findings from this study shed light on an important opportunity in early childhood education. First, we know that rural students face obstacles to early learning that their urban peers may not experience. These barriers, which often include transportation, a qualified teacher workforce, and access to facilities in distant and remote locations make preschool expensive for state governments and families. Therefore, the lack of resources is not easily remedied through adding more programs or more staff to existing programs- which is often a solution used in more densely populated urban regions. The Upstart program demonstrated that its innovative technology and program model could be implemented with success in rural regions, and this study added to the growing body of evidence about its positive impact on school readiness.

***Educational Technology as a Solution to Rural Barriers*.** Educational technology as a resource to combat inaccessibility and increase early competencies is a powerful offering, regardless of a program's ability to satisfy all prerequisite skills. During this project, the Upstart program tackled geographic barriers by bringing the solution directly to rural families. The children reaped the benefits from home without negotiating transportation, parental work schedules or time needed to travel long distances. It provided seamless access to curriculum for those who may otherwise not have had the ability to participate in a school readiness program. For rural children, and likely any child without adequate access to educational resources, educational technology becomes the bridge to pre-kindergarten content and instruction. This alternative is particularly crucial in communities where early education may not be a state

priority, where in-person pre-k programs are cost prohibitive, and geographic challenges make participation nearly impossible.

Research Limitations

All research includes limitations related to its design. In this study there were limitations related to sampling and the historical effects related to a global health pandemic that occurred in the middle of our study.

Sample. It should be noted that the current study included a demographically homogenous sample of children residing in rural states in the Great Plains region. The families enrolled in the Upstart program were predominantly white, limiting our ability to generalize our findings more broadly to all types of pre-k children from ethnically and racially diverse backgrounds.

Pandemic factors. This study occurred amid the Covid-19 pandemic, which led to concerns that the pandemic negatively affected kindergarten readiness due to lack of preschool access and staffing shortages reducing the number of children served. Data were collected at times when families were likely feeling the impact of pandemic-related disruptions to daily life. Some families may have delayed their child's start in kindergarten, due to learning challenges brought on by covid restrictions or increased time spent using at-home educational programs during the preschool year. While we acknowledge that these factors exist, treatment and control students were affected equally, both groups being randomly assigned and tested within the same cohorts and timeframes. The effects of the pandemic should have been equally shouldered across the two groups, however, we cannot be sure what, if any, impact the pandemic might have had on our ability to detect treatment effects.

Future Research

There are several things we would recommend for future research, based on the lessons learned during this study, including diversifying the study sample and focusing on math learning and developing new ways to measure long-term program impacts.

Diversify the sample. To broaden our understanding of how Upstart works for students of all backgrounds, we recommend that future research include a larger sample of students from different racial/ethnic and socio-economic groups and from states in different geographic regions in the country. This will help generalize the findings across diverse populations of pre-k students and extend the evidence for the types of students who may benefit most.

Build on early math. This study had an exploratory component examining the impact of early math interventions on kindergarten readiness, and our findings suggest a significant benefit. Future research should target the ways in which pre-k programs can improve the emerging math skills of young students as there is accumulating evidence that early math competencies support learning across other domains such as literacy, making it a critical school readiness skill (Nelson & McMaster, 2019; Litkowski et al., 2020; Terry, 2021). There is also a need to understand how early math skills and a STEM foundation can be optimally fostered, particularly with a national emphasis on STEM education and the importance of diversifying those interested in STEM domains.

***Longer-term impacts*.** Positive impacts to early literacy and math were not observed one-year post-program in kindergarten, however, our research was not designed to study what factors may have mediated the long-term impacts. For example, the following research questions could enhance our understanding of potential long-term impacts: Were students in the control group

catching-up to treatment students through classroom instruction during their kindergarten year? Did having treatment students with strong early literacy skills raise all students' learning in a kindergarten class? How did the development of early reading skills impact the development of early math skills and vice versa? What influence did the parental support model have on children's academic and social skill development? We know that children's experience in kindergarten itself has a strong influence on learning, but we need more creative research to study what mediates long-term impacts.

Given challenges related to access, opportunity, and resources, an effective computer-based kindergarten readiness program is particularly suited for underserved students who have historically struggled with limited early education resources. In the case of the Upstart program, the in-home delivery method with provisioned technology, inherently removed barriers to access due to geographic or socio-economic limitations. This is an important time to study the impacts of early education programs, with funding and attention focused on implementing expanded learning efforts across the country. We encourage early education researchers to continue to provide evidence for programs that support academic school readiness for rising kindergarteners. The opportunity for underserved children to get prepared in the year leading up to kinder and subsequently influencing their early academic trajectory could become an essential step toward leveling the playing field.

References

Abrami, P., & Borokhovski, E., & Lysenko, L. (2015). The effects of ABRACADABRA on reading outcomes: A meta-analysis of applied field research data. Journal of Interactive Learning Research, 26, 337- 367.

Allard, S. W. (2019). *Spatial patterns of work, poverty, & safety net provision in the U.S. US2050 Initiative* Research Report. New York: Peter G. Peterson Foundation. https://www.pgpf.org/sites/default/files/US-2050-Spatial-Patterns-of-Work-Poverty-and-Safety-Net-Provision-in-the-US.pdf

Atchison, B., & Diffey, L. (2018, February 5). Initiatives from preschool to third grade: A policymaker's guide. Retrieved from Education Commission of the States: https://www.ecs.org/initiatives-from-preschool-to-third-grade-a-policymakers-guide/

Barron, B., Cayton-Hodges, G., Boferding, L., Coople, C., Darling-Hammond, L., & Levine, M. (2011). Take a giant step: A blueprint for teaching children in a digital age. New York: The Joan Ganz Cooney Center at Sesame Workshop. Retrieved from http://joanganzcooneycenter. Org/upload_kits/jgcc_takeagiantstep.pdf

Blok, R. Oostdam, M. E. Otter, and M. Overmaat. University of Amsterdam. How effective are computer-assisted instruction (CAI) programs in support of beginning reading instruction: A review. Review of Educational Research, 72(1), 101-130.

Burchinal, M., Magnuson, K., Powell, D., & Hong, S. S. (2015). Early childcare and education. In M. H. Bornstein, T. Leventhal, & R. M. Lerner (Eds.), *Handbook of child psychology and developmental science: Ecological settings and processes* (pp. 223–267). John Wiley & Sons, Inc.

Charlson, C. (2017). *Early childhood education in South Dakota.* South Dakota Legislative Research Council. https://mylrc.sdlegislature.gov/api/Documents/IssueMemo/124660.pdf?Year=2017

Cohen, J., Cohen, P., West, S., & Aiken, L. (2003). *Applied multiple regression/correlation analysis for the behavioral sciences*. Hillsdale, NJ: Lawrence Erlbaum Associates.

Darling-Churchill, Kristen E., Lippman, Laura, (2016). Early childhood social and emotional development: Advancing the field of measurement. Journal of Applied Developmental Psychology. http://dx.doi.org/10.1016/j.appdev.2016.02.002

Daughtery, Lindsay, Dossani, Rafiq, Johnson, Erin Elizabeth, Oguz, Mustafa (2014). Using Early Childhood Education to Bridge the Digital Divide. Santa Monica, CA: RAND Corporation, 2014. https://www.rand.org/pubs/perspectives/PE119.html.

Fretchtling, J. (2007). *Logic Modeling Methods in Program Evaluation*. John Wiley & Sons, Inc.

Friedman-Krauss, A., Barnett, W. S., Garver K. A., Hodges K. S., Weisenfeld G.G., Gardiner B. A., & Jost T. M. (2022). *The State of Preschool 2021*. National Institute for Early Education Research. https://nieer.org/wp-content/uploads/2022/09/YB2021_Full_Report.pdf

Frey, J.R., Elliott, S.N., & Gresham, F.M. (2011). Preschoolers' Social Skills: Advances in Assessment for Intervention Using Social Behavior Ratings. School Mental Health, (3), 179-190.

Fyfe, E. R., Rittle-Johnson, B., & Farran, D. C. (2019). Predicting success on high-stakes math tests from preschool math measures among children from low-income homes. Journal of Educational Psychology, 111(3), 402.

Gresham, F. M., & Elliott, S. N. (2008). SSIS: Social skills improvement system. Minneapolis, MN: Pearson.

Hardy, J. K., & Hemmeter, M. L. (2019). Systematic instruction of early math skills for preschoolers at risk for math delays. *Topics in Early Childhood Special Education*, *38*(4), 234-247. https://doi.org/10.1177/0271121418792300

Hill, C. J., Bloom, H. S., Black, A. R., & Lipsey, M. W. (2007). *MDRC Working Papers on Research Methodology Empirical Benchmarks for Interpreting Effect Sizes in Research*. www.mdrc.org

Hoover, W. A., & Gough, P. B. (1990). The simple view of reading. Reading and Writing, 2, 127–160.

Hussar, B., Zhang, J., Hein, S., Wang, K., Roberts, A., Cui, J., & Dilig, R. (2020). *The Condition of Education 2020.* National Center for Education Statistics. https://nces.ed.gov/pubsearch/pubsinfo.asp?pubid=2020144

Jordan, N. C., Glutting, J., & Ramineni, C. (2010). The importance of number sense to mathematics achievement in first and third grades. Learning and individual differences, 20(2), 82-88.

Józsa, K., Amukune, S., Zentai, G., & Barrett, K. C. (2022). School readiness test and intelligence in preschool as predictors of middle school success: result of an eight-year longitudinal study. *Journal of Intelligence*, *10*(3), 66-85. https://doi.org/10.3390/jintelligence10030066

Kaufman, A. S., & Kaufman, N. L. (2014). Kaufman test of educational achievement, third edition. Bloomington, MN: NCS Pearson.

Kraft, M. A. (2020). Interpreting effect sizes of education interventions. *Educational Researcher*, *49*(4), 241–253. https://doi.org/10.3102/0013189x20912798

Lipsey, M., Puzio, K., Yun, C., Hebert, M., Steinka-Fry, K., Cole, M., Roberts, M., Anthony, K. and Busick, M. (2012). *Translating the statistical representation of the effects of education interventions into more readily interpretable forms*. Washington DC: Institute of Education Sciences.

Litkowski, E. C., Duncan, R. J., Logan, J. A., & Purpura, D. J. (2020). When do preschoolers learn specific mathematics skills? Mapping the development of early numeracy knowledge. *Journal of Experimental Child Psychology*, *195*, 1-25. https://doi.org/10.1016/j.jecp.2020.104846

Macaruso, P., Hook, P. E., & McCabe, R. (2006). The efficacy of computer-based supplementary phonics programs for advancing reading skills in at-risk elementary students. Journal of Research in Reading, 29(2), 162–172. https://doi.org/10.1111/j.1467-9817.2006.00282.x

Mejias S., Muller C., & Schiltz C. (2019). Assessing Mathematical School Readiness. *Frontiers In Psychology*, *10*, 1-11. https://doi.org/10.3389/fpsyg.2019.01173

Miller, P., & Votruba-Drzal, E. (2013). Early academic skills and childhood experiences across the urban-rural continuum. *Early Childhood Research Quarterly, 28*(2), 234-248. https://doi.org/10.1016/j.ecresq.2012.12.005

National Center for Education Statistics. (2023). Early childhood care and education programs in rural areas. *Condition of Education.* U.S. Department of Education, Institute of Education Sciences. https://nces.ed.gov/programs/coe/indicator/lca

National Early Literacy Panel. (2008). Developing early literacy: Report of the National Early Literacy Panel. Washington, DC: National Institute for Literacy. Retrieved from https://lincs.ed.gov/publications/pdf/NELPReport09.pdf

Nelson, G., & McMaster, K. L. (2019). The effects of early numeracy interventions for students in preschool and early elementary: A meta-analysis. *Journal of Educational Psychology, 111*(6), 1001-1022. https://doi.org/10.1037/edu0000334

Phillips, D. A., Lipsey, M. W., Dodge, K. A., Haskins, R., Bassok, D., Burchinal, M. R., (& Weiland, C. (2017). Puzzling it out: The current state of scientific knowledge on pre-kindergarten effects. A consensus statement. Washington, DC: The Brookings Institution. Retrieved from https://www.brookings.edu/research/puzzling-it-out-the-current-state-of-scientific-knowledge-on-pre-kindergarten-effects/

Raver, C. C., & Knitzer, J. (2002). Ready to Enter: What Research Tells Policy Makers about Strategies to Promote Social and Emotional School Readiness among Three and Four-Year-Olds. Washington DC: National Center for Children in Poverty.

Rogowsky, B.A., Terwilliger, C.C., Young, C.A, & Kribbs, E.E (2017): Playful learning with technology: the effect of computer-assisted instruction on literacy and numeracy skills of preschoolers, International Journal of Play, DOI: 10.1080/21594937.2017.1348324

Sheridan, S. M., Knoche, L. L., Edwards, C. P., Bovaird, J. A., & Kupzyk, K. A. (2010). Parent Engagement and School Readiness: Effects of the Getting Ready Intervention on Preschool Child’en's Social-Emotional Competencies. Early education and development, 21(1), 125–156. https://doi.org/10.1080/10409280902783517

Showalter, D., Klein, R., Johnson, J., & Hartman, S. (2017). *Why rural matters 2015-2016: Understanding the changing landscape.* Rural School and Community Trust. http://www.ruraledu.org/user_uploads/file/WRM-2015-16.pdf

Terry, S. D. (2021). *The relationship between prekindergarten and math achievement of first graders in a Title One school* (Publication No. 28768569). [Doctoral Dissertation, Trevecca Nazarene University]. ProQuest Dissertations Publishing.

U.S. Census Bureau (2022*). Enrollment Rates of Young Children.* https://nces.ed.gov/programs/coe/indicator/cfa/enrollment-of-young-children

Vogels, E. (2021). Some digital divides persist between rural, urban and suburban America. Report for Pew Research Center.

Waterford.org | Early Learning Software. (2022, April 20). Waterford.Org. Retrieved June 22, 2022, from https://www.waterford.org/

What Works Clearinghouse. (2022). *What Works Clearinghouse procedures and standards handbook, version 5.0.* U.S. Department of Education, Institute of Education Sciences, National Center for Education Evaluation and Regional Assistance (NCEE). https://ies.ed.gov/ncee/wwc/Handbooks

Wolf, Goodson, and Boulay (2023). Template for Reporting on Effects of Education Programs and Practices and on Fidelity of Implementation with a Focus on EIR Grant Evaluations. Prepared for the Institute of Education Sciences, U.S. Department of Education, Washington, DC., February. Abt Associates Inc.

Wood, E., Gottardo, A., Grant, A., Evans, M., Phillips, L., & Savage, R. (2012). Developing tools for assessing and using commercially available reading software programs to promote the development of early reading skills in children. NHSA Dialog, 15(4), 350-354.

Yoshikawa, H., Kholoptseva, J., & Suárez-Orozco, C. (2013). The role of public policies and community-based organizations in the developmental consequences of parent undocumented status. Social Policy Report, 27(3), 1-17. from http://www.srcd.org/sites/default/files/documents/spr_27_3.pdf

Appendix A. Regression Results

Table 8. Regression Summary of Predictors of Reading Comprehension

Variable	*B*	*SE B*	*β*
(Constant)	148.38	13.21	
Reading Treatment**	5.11	1.14	0.12
Pre-Test**	0.64	0.35	0.49
Cohort 1	0.49	1.89	0.10
Cohort 2	0.84	1.54	0.20
Idaho	0.42	2.12	0.01
Wyoming	-0.84	1.84	-0.18
North Dakota	-0.35	2.19	-0.006
South Dakota	-1.83	-0.31	-0.031
R^2	0.25		
F	44.08		

* p<0.05; ** p<0.01 (N= 1,053; Reading Treatment n=507; Math/Science Control n=546)

Table 9. Regression Summary of Predictors of Phonological Processing

Variable	*B*	*SE B*	*β*
(Constant)	204.71	11.84	
Reading Treatment**	5.19	0.69	0.19
Pre-Test**	0.58	0.26	0.56
Cohort 1*	-0.23	1.14	-0.07
Cohort 2*	-2.08	0.93	-0.75
Idaho	-1.54	1.277	-0.47
Wyoming	-2.30	1.11	-0.75
North Dakota*	-3.70	1.32	-0.10
South Dakota	-3.79	1.24	-0.10
R^2	0.36		
F	73.30		

* p<0.05; ** p<0.01 (N= 1,053; Reading Treatment n=507; Math/Science Control n=546)

Table 10. Regression Summary of Predictors of Letter and Word Recognition

Variable	*B*	*SE B*	*β*
(Constant)	41.33	8.44	
Reading Treatment**	14.82	1.59	0.187
Pre-Test**	0.97	0.26	0.738
Cohort 1	1.51	2.64	0.168
Cohort 2	-1.26	2.15	-0.159
Idaho	-0.94	2.96	-0.009
Wyoming	-2.63	2.57	-0.030
North Dakota	-3.73	3.05	-0.362
South Dakota	-0.16	2.88	-0.0015
R^2	0.58		
F	179.40		

* p<0.05; ** p<0.01 (N= 1,053; Reading Treatment n=507; Math/Science Control n=546)

Table 11. Regression Summary of Predictors of Math Concepts and Applications

Variable	*B*	*SE B*	*β*
(Constant)	175.68	10.43	
Math/Science Treatment**	4.97	1.06	0.12
Pre-Test**	0.63	0.026	0.65
Cohort 2	-0.91	1.27	-0.22
Idaho	-1.24	1.68	-0.22
Wyoming	0.76	1.30	0.02
North Dakota	-1.74	1.74	-0.30
R^2	0.45		
F	105.94		

* p<0.05; ** p<0.01 (N= 1,053; Reading Treatment n=387; Math/Science Control n=387)

Table 12. Subscale Correlation Table

	LWR GSV Pre-Test	PP GSV Pre-Test	RC GSV Pre-Test	LWR GSV Post-Test	PP GSV Post-Test	RC GSV Post-Test
LWR GSV Pre-Test	1.00**					
PP GSV Pre-Test	0.432**	1.00**				
RC GSV Pre-Test	0.4936**	0.3746**	1.00**			
LWR GSV Post-Test	0.7361**	0.4037**	0.4381**	1.00**		
PP GSV Post-Test	0.4843**	0.5560**	0.3446**	0.6143**	1.00**	
RC GSV Post-Test	0.6594**	0.3940**	0.4861**	0.7996**	0.5507**	1.00**

* p<0.05; ** p<0.01

Note: LWR (Letter & Word Recognition), PP (Phonological Processing), RC (Reading Comprehension), GSV (Growth Scale Value)

Appendix B. Baseline Equivalence

Baseline equivalence was established between the treatment and control groups on pre-test measures of achievement using independent samples *t*-tests. Mean differences are presented, along with *t*-values and significance levels.

Table 13 presents pre-test mean scores on the outcome variables of interest; Reading Comprehension, Letter and Word Recognition, Phonological Processing, and social-emotional constructs. Initial results from *t*-tests indicate pre-program differences between children assigned to treatment and control conditions were not significant, indicating comparable levels of early literacy, and social emotional skills between the two experimental groups prior to beginning the UPSTART program.

Table 13. Baseline Equivalence Assessment

	Control Group (Math/Science)			Treatment Group (Reading)					
Measure	Sample Size	Mean	St. Dev.	Sample Size	Mean	St. Dev.	*t*-value	Tr – Ctrl Diff	Standardized Diff*
Letter and Word Recognition	546	303.08	29.20	507	302.82	31.03	0.14	0.262	-0.009
Reading Comprehension	546	369.34	15.98	507	369.70	16.47	-0.36	0.362	0.022
Phonological Processing	546	453.26	13.24	507	452.42	13.19	1.02	0.834	-0.063
SSIS Rating Scales (RS)	545	96.56	16.05	506	96.79	15.55	-0.25	-0.240	0.015

* $p<0.05$; ** $p<0.01$
**Note. Hedges G Effect Size: Treatment minus control divided by pooled standard deviation.*

Table 14 presents the baseline equivalence for participant characteristics across research condition. We used Fisher's Exact statistical test to determine if there were nonrandom associations between the demographic characteristic and the research group. As shown, there

were no significant differences between the treatment and control conditions for any of the participant characteristics.

Table 14. Baseline Equivalence of Treatment-Control Groups by Participant Characteristic

Baseline Characteristic	N	Percent	*p-value*
Female			
Reading	254	50%	0.58
Math/Science	283	52%	
White			
Reading	438	86%	0.26
Math/Science	458	84%	
Over 200% Poverty			
Reading	188	38%	0.44
Math/Science	214	40%	
*p<0.05; ** p<0.01			

Appendix C. Attrition Analysis

Attrition is the loss of sample during the duration of a study and is a common issue in educational research. It is important to ensure attrition numbers fall within acceptable boundaries and do not compromise the integrity of the randomization process. Of the 1,348 students who were randomized for participation in the evaluation study, 652 were placed in the reading condition and 696 were placed in the math condition.

According to the WWC Standards and Procedures, overall individual-level attrition is the "the number of individuals in an allowable reference sample minus the number in the analytic sample" (5.0; 2022). Table 15 outlines the sample size attrition across time and outcome measure among the total research sample and each assigned research condition.

Table 15. Sample Sizes at Randomization and Analytic Sample Needed to Assess Attrition for an RCT with Individual-Level Assignment

Outcome Measure	Control Group (Math/Science)		Treatment Group (Reading)	
	# Randomized	# Analytic Sample	# Randomized	# Analytic Sample
Reading Comprehension	696	546	652	507
Phonological Processing	696	546	652	507
Letter and Word Recognition	696	546	652	507
SSIS Rating Scales	693	542	649	503

The current study had an overall individual-level attrition rate of 22% and a differential attrition rate of 0%, difference in attrition between the intervention group and control group (see Table 21). According to standards set by What Works Clearinghouse, an overall attrition rate of 22% must have a differential attrition rate of less than 5.1% to have a tolerable threat of bias

under both optimistic (i.e., attrition is exogenous or unrelated to the intervention) and cautious (i.e., attrition is endogenous or related to the intervention) assumptions regarding the relationship between attrition and outcomes. Attrition rates in Table 16 show that overall and differential attrition between math/science and reading students were within acceptable levels of bias set by What Works Clearinghouse standards (5.0; 2022) and that our RCT design remained intact.

Table 16. Number of Participants in Research Sample

Group	Control (Math/Science)	Treatment (Reading)	Total Sample
Cross Cohort (Cohort 1, 2 and 3)	696	652	1,348
KTEA-3			
Children Post-Tested	546	507	1,053
% Attrition	22%	22%	22%
% Differential Attrition	0%		
SSIS			
Children Post-Tested	542	503	1,047
% Attrition	22%	23%	24%
% Differential Attrition	1%		

Appendix D. Virtual Assessment Procedures

The format of the school readiness assessments changed as a result of safety measures instituted to prevent the spread of the novel coronavirus (SARS-CoV-2). In this Appendix, we describe the testing procedures implemented for in-person assessments at pre-test for Cohort 1 and the virtual testing assessments implemented for all subsequent data collection periods, from 2020 through 2022. Our approach focused on maintaining the integrity of the assessments and keeping the procedures similar despite the change in testing format.

In-Person Testing. At the beginning of in-person testing, parents were provided a paper consent form to complete by the test administrator. Once the consent form was signed, parents were asked to complete the parent survey and SSIS Rating Scales online using a tablet provided by the researchers. While parents were working on the survey, the test administrator conducted the assessment with the child using a set of KTEA-III instruments and Stimulus Books' 1 and 2. Parents received an Amazon gift card and children received stickers as a thank you for their participation at the end of the testing session.

Virtual Testing. Researchers modified data collection procedures during the summer of 2020 to complete all testing online. All measures used to collect parent and student data were moved to online systems that were remotely managed by research staff. A link to complete the informed consent form, parent survey and the SSIS Rating Scales survey were emailed to parents for completion prior to their child's online assessment. For students, a combination of video conferencing, desktop sharing and computer co-browsing software was used to create an online interactive testing platform that mimicked in-person testing using the KTEA-III. Test administrators had visual contact and audio capabilities with the student through video

conferencing and the student could interact with the test stimuli through a custom co-browsing software application. These features allowed for an interactive "virtual KTEA-III" test, and the student could point to test stimuli using her/his computer mouse.

Test administrators assisted parents with the technical setup prior to the child assessment to ensure the appropriate technology was present and functioning properly. Once the tech check was completed, the test administrator could then conduct the assessment with the child. Gift cards were emailed to parents following completion of all parent forms and the virtual child assessment.

During the assessments, the test administrators (TAs) followed precise testing protocols that included language, prompts and scoring. TAs scored each testing item in real time as they delivered each test question. Each TA had the ability to make note of relevant observations and return to their scoring following the completion of the assessment if adjustments were needed. The researchers observed each TA during each data collection period and used a structured protocol to provide feedback about the administration of each subscale. Additionally, research staff compared mean scores of all TAs, utilizing bi-variate statistical analyses, at three points during each data collection effort. At the end of the observation period, all staff were implementing the testing protocol with fidelity. Researchers also reviewed all data in the database to check for scoring or coding related errors.

Blinding. After children were pre-tested and randomly assigned to their research condition, families were made aware of their Upstart program assignment (Reading or Math/Science). At no time, however, were TAs made aware of the research condition of the child they were assessing. The research staff successfully ensured that any TA responsible for

conducting outcome assessments did not have access to any documentation containing the child's assigned condition.

CHAPTER TWO

EIR Great Plains Task Force Grant

Taking All to Success in Kindergarten

Implementation Study

Program Implementation Study

The focus of the implementation study was to determine if the key activities of the program were delivered to the participating families as planned. Studying program implementation was important because it provided information about the key program components and the extent to which the implementation was executed with fidelity. In accordance with the federal guidelines, fidelity was defined as "adherence to the planned delivery of the key program activities that were hypothesized to lead to the targeted outcomes of the program." This chapter presents our findings from three years of Upstart program implementation for the EIR Great Plains Task Force Grant spanning Year 2[5] (2019-2020), Year 3 (2020-2021) and Year 4 (2021-2022), including information about the cohorts and data sources used for assessing the success of the program.

Upstart Registration and Enrollment. Waterford, the program provider, registered over 5,500 children in the Upstart program across the three years from 2019-2022 within South Dakota, North Dakota, Wyoming, Idaho and Montana. Each program year represented a different cohort of pre-k children from a different combination of participating states:

Cohort 1: North Dakota, Idaho, Wyoming
Cohort 2: North Dakota, Idaho, Wyoming, South Dakota, Montana
Cohort 3: Wyoming, South Dakota, Montana[6]

Data Sources. To evaluate the implementation of the Upstart program, Waterford transferred several comprehensive datasets to ETI containing Upstart enrollment, provisioned

[5] Year 1 represented planning; Year 5 reporting

[6] North Dakota and Idaho were added to Cohort 3 to benefit more children as a result of the pandemic

educational technology, Upstart program usage, and information on the parent support and training activities and communications for each of the three program years spanning (2019-2022). The records provided data across all usage variables including average weekly use, average number of days per week, average session duration, average number of weeks and overall average total time with the program. Waterford provided program graduation data indicating whether or not graduation requirements were successfully met. Data included information about the type of technology provisioned to the family by Waterford. Waterford additionally provided information and documentation for the types of parental support and trainings offered to parents throughout their child's program year. Demographic information about the child and household were provided through the program registration process. These datasets were analyzed by ETI to produce the findings related to the success of the program's implementation goals.

Key Program Implementation Components. The critical inputs and activities that were essential for implementing the Upstart intervention were under the direct control of the program provider. As part of the evaluation, our logic model outlined all program inputs that reflected the components necessary to replicate the treatment or strategies undertaken to reach the desired outcome. Below is a list of the components represented in the logic model displayed below, each of which is described in further detail.

Upstart Program Usage. The primary component of Waterford's Upstart program, was an in-home computer-based preschool program that used software to provide children with reading, math, and science curriculum, with a focus on reading instruction. The program was designed to promote the development of literacy and math skills that prepare young children for

entry into school by providing an individualized learning experience that adapted to children's skill level. Content was delivered online through adaptive lessons, digital books, songs, and activities.

While the complete Upstart program typically consists of curriculum in both reading and math/science, in order to craft a rigorous evaluation as required by EIR standards, the program was split into two components: *Reading only* for children randomly assigned to the treatment group, and *Math/Science only* for children assigned to the control group. The format, delivery, and requirements of the software program were identical between the two groups, and the only difference consisted of curriculum content.

Program Usage Requirements. While the recommended use for either the Reading or Math/Science Upstart program is 20 minutes a day for 5 days a week, children are required to use the program at home for 15 minutes a day, 5 days a week throughout the program year.

Education Technology. The Upstart educational software was delivered over the internet in a web-based format that allowed access from any computer with high-speed internet that met minimum operating requirements. To determine need, Waterford coaches assessed families' existing computer hardware (e.g., operating system, RAM data storage, CPU processor) during the initial intake call and determined whether or not families had access to high-speed internet. Some families recruited to participate in the program did not have the necessary technology or suitable internet access. Fortunately, grant funds allowed for the provision of compatible computers, modems, wireless cards, and/or internet service to access Upstart software.

Parent Training and Support. In order to facilitate academic supervision and encourage the use of program software, families were provided with several forms of technical, motivational and curriculum support. A description of the various family support resources are outlined below.

Parent Training and Ongoing support. Before beginning the program, participating parents attended a comprehensive orientation where they reviewed state kindergarten preparedness guidelines alongside Upstart curricular content, discussed strategies for motivating children to use the program consistently, learned about software features, and reviewed available resources.

Social Emotional Learning. Waterford staff additionally sent parents Child Social and Emotional Workbooks and tips on how to foster social and emotional learning with their four-year old children. Tips and additional training were given in weekly mentor emails, which covered topics in the SEL workbook.

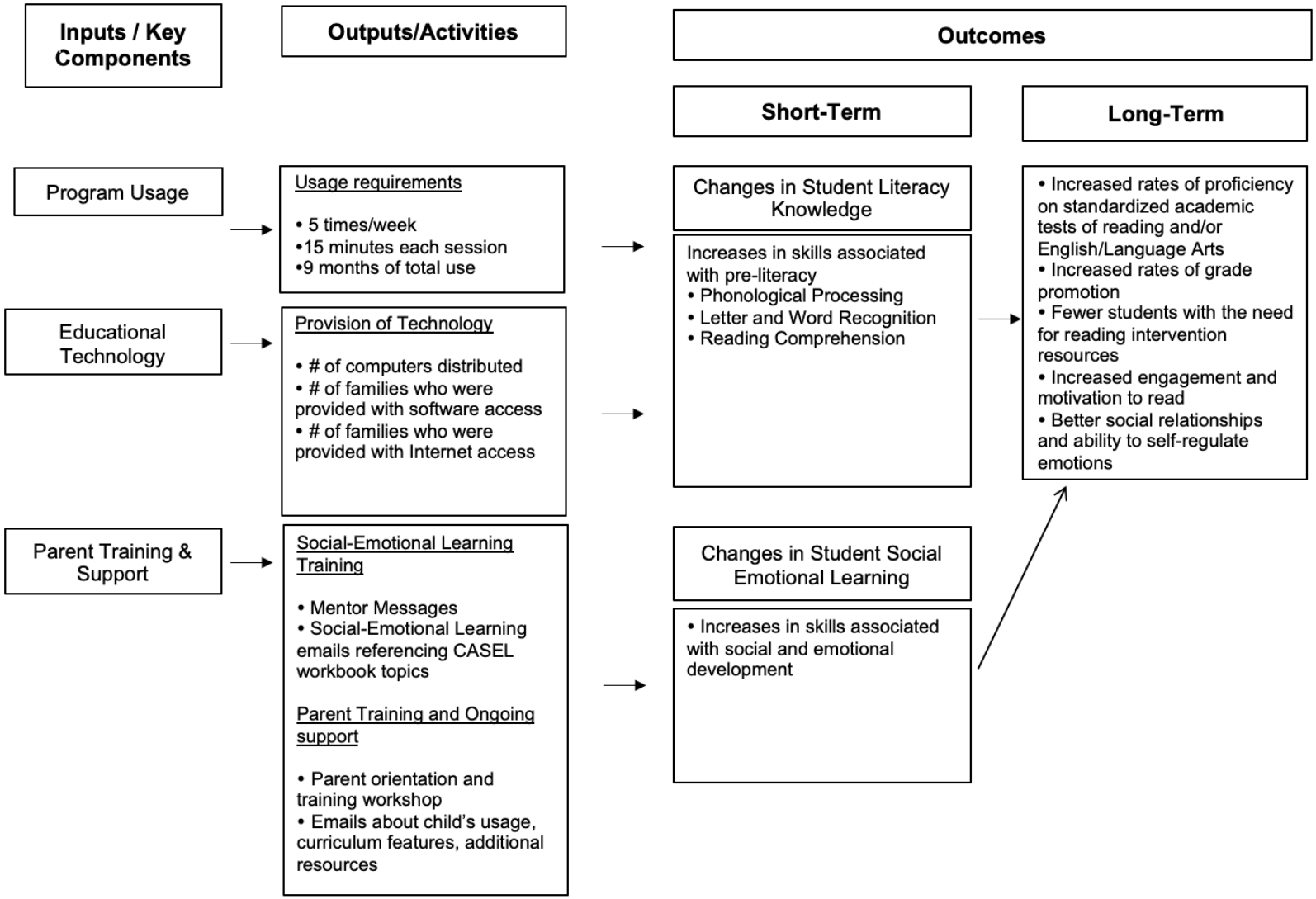
Coaching to Enhance Program Engagement. Families also received ongoing support from their assigned Waterford coach. These specialized staff members served as partners with families to provide technical and motivational support. They monitored children's program usage and contacted parents if usage dropped below the minimum requirement. They provided families with individualized strategies to encourage consistent use and served as the primary contact if parents encountered technical difficulties or challenges when using the computer program.

Parent Manager. A password-protected Parent Manager portal in the Upstart software program allowed parents to monitor children's usage on a daily basis, as well as review children's assessment scores and progress. In the Manager portal, parents downloaded reports

that showed children's placement results, current lessons, and unit lesson scores. In addition, there were specific activity recommendations and enrichment materials in the Manager that parents and children could do together to supplement learning.

Weekly Messages. Weekly emails that contained graphs of children's weekly usage, features of program curricular content, and suggestions for supplementary educational activities were sent to parents throughout the program year. Families received either a Reading or Math/Science email, depending on their experimental condition. In addition, regular Mentor messages were sent to the Upstart site to provide social-emotional development details for parents to access.

EIR Logic Model



Measuring Program Implementation Fidelity

Measuring Fidelity. Program implementation fidelity was defined as adherence to the planned delivery of the key program activities that were hypothesized to lead to the targeted outcomes of the program. Our evaluation measured the extent to which the key components of the program were delivered as planned by the program provider. **Table 17** provides how fidelity was measured annually for each of the 3 key program components including how adequate implementation was defined and what sample-level threshold was required to meet program implementation fidelity.

Table 17. Scoring Definitions for Adequate Implementation and Sample-Level Thresholds for Fidelity of Each Key Program Component

Indicators	Definition	Unit	Indicator Scoring at Unit Level	Indicator Scoring at Sample Level
Key Component 1. Program Usage				
Minutes of Use (per cohort)	Total number of minutes of Upstart use	Student	Low (0): ≤ 1350 minutes Moderate (1): = 1351-2159 minutes High (2): ≥ 2160 minutes	
Weeks of Use (per cohort)	Total number of weeks of Upstart use	Student	Low (0): ≤ 18 weeks Moderate (1): = 19-28 weeks High (2): ≥ 29 weeks	
Average Weekly Use (per cohort)	Average number of weekly minutes of Upstart use	Student	Low (0): ≤ 38 minutes Moderate (1): = 39-59 minutes High (2): ≥ 60 minutes	
Key Component 1 Total Score Program Usage			Sum of Student-level indicator scores (range = 0-6) Adequate student-level score = ≥5	Sample-level threshold for fidelity = at least 75% of students with a score of ≥ 5

Indicators	Definition	Unit	Indicator Scoring at Unit Level	Indicator Scoring at Program Level
Key Component 2. Provision of Technology				
Provision of hardware (per cohort)	% families in need who were provided with computers	Program	Low (0): ≤ 50% Moderate (1): = 51-79% High (2): ≥ 80%	
Provision of software (per cohort)	% families who were provided with software access	Program	Low (0): ≤ 90% Moderate (1): = 91-94% High (2): ≥ 95%	
Provision of internet access (per cohort)	% families in need who were provided with internet access	Program	Low (0): ≤ 50% Moderate (1): = 51-79% High (2): ≥ 80%	
Key Component 2 Total Score Education Technology				Sum of Program-level indicator scores (range = 0-6) Adequate Program-level = score of ≥5

Indicators	Definition	Unit	Indicator Scoring at Unit Level	Indicator Scoring at Sample Level
Key Component 3. Parent Training and Support				
Mentor messages (per cohort)	Messages were sent to the Upstart site to provide social-emotional development details for four-year-old's	Family	Low (0): ≤ 20 messages Moderate (1): = 21-30 messages High (2): ≥ 31 messages	
Social-emotional learning emails (per cohort)	Emails sent to guide family learning about social-emotional topics and tips	Family	Low (0): ≤ 20 emails Moderate (1): = 21-30 emails High (2): ≥ 31 emails	
Parent orientation and training workshop (per cohort)	Two-hour training and introductory session	Family	Low (0): 0 hours Moderate (1): = 1 hour High (2): 2 hours	
Curriculum support (per cohort)	Emails to parents about UPSTART curriculum features	Family	Low (0): ≤ 4 emails Moderate (1): = 5-7 emails High (2): ≥ 8 emails	
Key Component 3 Total Score Parent Training and Support			Sum of Family-level indicator scores (range = 0-8) Adequate family-level = score of ≥7	Sample-level threshold for fidelity = 75% of families with a score of ≥ 7

Fidelity Findings

The results from our implementation analysis are provided in **Table 18**. The three key program components were evaluated based on the aforementioned measurement approach and results were broken out for each year the program was implemented. As shown in **Table 18**, the sample-level threshold for fidelity was achieved in all years for educational technology and for parent training and support. Program usage met the sample-level threshold in Year 1.

Table 18. Findings on Fidelity of Implementation by Key Components in Multiple Years

Key Components, Number of Indicators, Units, and Threshold				Year 1 Results (2019-20 School Year)			Year 2 Results (2020-21 School Year)			Year 3 Results (2021-22 School Year)		
Key Component	**Total # of Measurable Indicators**	**Unit of Implementation**	**Sample-Level Threshold for Fidelity of Implementation**	**Number of Units in Which Component was Implemented**	**Number of Units in Which Fidelity of Component was Measured**	**Achieved Fidelity Score and Whether Program Met Sample-Level Threshold**	**Number of Units in Which Component was Implemented**	**Number of Units in Which Fidelity of Component was Measured**	**Achieved Fidelity Score and Whether Program Met Sample-Level Threshold**	**Number of Units in Which Component was Implemented**	**Number of Units in Which Fidelity of Component was Measured**	**Achieved Fidelity Score and Whether Program Met Sample-Level Threshold**
1.Program Usage	3[i] Each with a score of 0-2, summed to a component score ranging from 0-6	Student	75% of students are categorized as high fidelity (i.e. score ≥ 5)	723 Students	723 Students	Fidelity score = 78% ***Met program fidelity = Yes***	2590 Students	2590 Students	Fidelity score = 53% ***Met program fidelity = No***	2277 Students	2277 Students	Fidelity score = 59% ***Met program fidelity = No***
2.Educational Technology	3[ii] Each with a score of 0-2, summed to a component score ranging from 0-6	Program	Program-level score of 5 or higher	723 Families	723 Families	Fidelity score = 6 ***Program fidelity = Yes***	2590 Families	2590 Families	Fidelity score = 6 ***Program fidelity = Yes***	2277 Families	2277 Families	Fidelity score = 6 ***Program fidelity = Yes***
3.Parent Training and Support	4[iii] Each with a score of 0-2, summed to a component score ranging from 0-8	Family	75% of families are categorized as high fidelity (i.e. score ≥ 7)	723 Families	723 Families	Fidelity score = 100% ***Program fidelity = Yes***	2590 Families	2590 Families	Fidelity score = 100% ***Program fidelity = Yes***	2277 Families	2277 Families	Fidelity score = 100% ***Program fidelity = Yes***

[i] Indicators include: minutes of total use, number of weeks of use, average weekly minutes of use
[ii] Indicators include: provision of hardware, software, internet access
[iii] Indicators include: mentor messages, email SEL messages, parent orientation, curriculum support

Implementation Summary of Findings

Upstart Software Program Usage: Met fidelity of implementation Year 1 only

Program usage was defined by the following indicators: (1) total number of minutes, (2) total number of weeks, and (3) average weekly minutes that the child used the program. For each indicator, the student data was categorized into 3 levels: Low, Moderate and High. The percentage of students able to attain the highest level of usage dipped in the second year of implementation. Implementation results show that 78% of the students achieved the highest level for program usage in Year 1, approximately 53% in Year 2, and 59% in Year 3. The onset of the pandemic just prior to Year 2 likely influenced the decline among the usage indicators. In-person orientations and family onboarding was no longer possible in Year 2; all Waterford parent partners and Upstart participant coorespondence transitioned to virtual platforms. Some participating EIR families experienced upheaval and disruption to daily life throughout Year 2 and possibly into Year 3, however, lower levels of program usage may have been more directly influenced by the lack of in-person parent engagement and an absence of program staff serving as 'boots on the ground.'

Educational Technology: Met fidelity of implementation Year 1, 2 and 3

As part of the implementation of the program, Waterford sought to provide families in need with the appropriate hardware, software or internet access. For each new cohort, families were assessed for their existing technology and internet access. Implementation findings show that all families with educational technology needs were provisioned equipment and/or access in order to participate in the program during every program year.

Parent Training and Support: Met fidelity of implementation Year 1, 2 and 3

Waterford engaged with participating families in a number of different ways and adapted over the course of the grant. For Cohort 1, the Waterford staff held in-person orientation workshops in order to train and onboard families. For Cohorts 2 and 3, Waterford prepared families in the beginning with virtual online orientation workshops and trainings. Additionally, parents received emails about software content and curriculum support throughout the duration of the program. Waterford coaches partnered with parents and monitored children's program usage and provided motivation to families. Parents were given a SEL Workbook to foster social and emotional learning with their four-year old. Waterford continued to track successfully within the parent training and support implementation indictors among all 3 Cohorts of participants.

CHAPTER THREE

EIR Great Plains Task Force Grant

Taking All to Success in Kindergarten

Scale-Up Study

Upstart Program Scale Up Study

Effective interventions need a plan to scale up and develop the funding, partnerships and resources required to continue their mission and serve a broader population after their grant is over. The Education and Innovation Research (EIR) Expansion Grant: *The UPSTART Great Plains TASK Force: Taking All to Success in Kindergarten*, had several big picture scale up goals beyond strictly serving students. Establishing a partnership network of support was critical for Waterford to exceed yearly recruitment targets and serve more students throughout the five partner states. The following chapter presents our findings from a scale up analysis, assessing the grantee's ability to create a foundation within the EIR states from which the UPSTART program could continue beyond the federal grant funded years.

The Evaluation and Training Institute (ETI) analyzed the scale up strategies and evaluated the level of success of each scale up effort. Data from quarterly Task Force meetings throughout 2019-2022, annual grantee progress reports, and interviews with program staff were used to evaluate each scale up effort. These data sources provided information on Task Force meeting topics and strategies, Waterford yearly implementation updates, and details on how families were recruited to participate in the program. From interviews we learned more about the logistical challenges of building relationships and disseminating information about the program in this specific geographic region of the country.

Table 19 lists the main goals for creating program spread and sustainability, the known challenges to the stated goals and the strategies intended to address those challenges.

Table 19. Strategy to Scale

Scale Up Goal	Challenge to Meeting Goal	Strategy to Address Challenge
Goal #1: Task Force Development-Develop inroads into state preschool and early childhood stakeholder groups and government agencies to enact state-wide preschool initiatives**.**	Stakeholder groups in partner states might be difficult to engage and may be resistant to participating in a cross-state consortium. Educational professionals might consider enacting statewide preschool initiatives secondary to the more immediate need to respond to COVID-19, and the time and effort directed toward COVID-19 might leave little time available for other efforts.	Extensive outreach to high-level educational professionals in each state to form a Task Force, intended to guide program buy-in efforts in the state and assist in developing implementation strategies. Develop a committee of stakeholders that meet one of the following qualifications: early childhood education (ECE) expertise, knowledge of state ECE policy, including members who can identify ECE challenges in their state. Task Force meetings with members from all five states, convened in person and virtually on a quarterly basis.
Goal #2: Liaison Development[7]**-** Develop state-wide network of field staff to develop relationships and support home-based preschool activities.	States often do not have the capacity to hire qualified professionals to serve as liaisons. Establishing a liaison network is especially challenging in rural areas, because it is more difficult for liaisons to cover the larger geographic area and work with the larger number of smaller school districts that exist in rural areas as compared to suburban and urban areas. Challenges connected to Covid-19 limiting in-person interactions and travel restrictions.	*Intended:* Develop a state-wide model for finding, training, and employing qualified Program Liaisons to work with program families. *Intended:* Find local educational professionals to serve as Liaisons, skilled at adapting and working within their state's preschool and early childhood environment.
Goal #3: Outreach and Enrollment Team- Create team of professionals who are skilled at navigating each state's preschool and early childhood environment and who will assist with program recruitment.	Difficulty assembling a team to manage the facilitation of program awareness, interest, and enrollment with a large number of smaller school districts.	Build relationships and hire employees deeply rooted in each state's preschool and early childhood environment who can develop recruitment strategies by creating partnerships with individuals and organizations. Partnerships will help to guide program buy-in efforts within the state.

[7]Local Liaisons were part of the original design but shifted to Outreach and Enrollment team as a result of pandemic restrictions.

Measurement of Scaling Strategies

In order to evaluate the scaling strategies, each strategy's intended implementation was defined, a threshold for a successful implementation was established, and a plan for collecting data and reporting results was created. This section outlines the measurement of 3 main scale up goals, two of which were part of the original design and one that evolved as a necessary alternative for addressing the pandemic. Several methods were established to measure the success of these scale up strategies as defined in more detail in **Table 20**.

Goal 1: Assembling a Task Force of early childhood education professionals and experts in EIR state policies was the scale up strategy designed to provide state-level implementation knowledge and buy-in during the grant funded years and beyond. In order to measure the success of this strategy, data were collected from all quarterly task force meetings from 2019-2022.

Goal 2: Similarly, the liaison approach was designed to develop a statewide network of local professionals from school districts and other early childhood organizations within each partner state, building relationships to be leveraged for future program recruitment.

Goal 3: The work conducted by the new Outreach and Enrollment Team was more centralized, minimizing the challenges inherent in a 'boots on the ground' approach, making it easier to assemble a team to manage the facilitation of program awareness, interest, and enrollment. In this necessary adaptation, team members were not bound by geography but rather hired to form and develop local partnerships, making canvassing the state more efficient.

Table 20. Measurement of Scaling Strategy

Scale-up Goal	Challenge to Meeting Goal	Strategy to Address Challenge	Definition of Full Implementation of Strategy (in measurable units)	Threshold for Level of Implementation Defined as "Successful"	Data Collection and Reporting Plan for Measuring Implementation of Strategy
Goal #1- Task Force Development	Stakeholder groups in partner states might be difficult to engage and may be resistant to participating in a cross-state consortium. Educational professionals might consider enacting statewide preschool initiatives secondary to the more immediate need to respond to COVID-19, and the time and effort directed toward COVID-19 might leave little time available for other efforts.	Extensive outreach to high-level educational professionals in each state to form a Task Force, to guide program buy-in efforts the state and assist in developing implementation strategies in each state. Develop a committee of stakeholders that meet one of the following qualifications: early childhood education (ECE) expertise, knowledge of state ECE policy, including members who can identify ECE challenges in their state and determine solutions best suited for their constituents. Task Force meetings with members from all five states, convened in person and by phone on a quarterly basis.	Quarterly TASK force meetings to share overall project updates, state specific updates and evaluation progress. Meeting minutes per year (3 total per year for 4 years)	Implementation was considered a success if quarterly TASK force meetings took place and if planning documents (meeting minutes) for all TASK Force meetings were distributed via email.	Data were collected at the conclusion of each Task Force meeting, when meeting minutes were distributed via email to all participants. Data were also reported in the Waterford annual report and analyzed at the end of each year of the grant.

Scale-up Goal	Challenge to Meeting Goal	Strategy to Address Challenge	Definition of Full Implementation of Strategy (in measurable units)	Threshold for Level of Implementation Defined as "Successful"	Data Collection and Reporting Plan for Measuring Implementation of Strategy
Goal #2- Liaison Development	States often do not have the capacity to hire qualified professionals to serve as liaisons. Establishing a liaison network is especially challenging in rural areas, because it is more difficult for liaisons to cover the larger geographic area and work with the larger number of smaller school districts that exist in rural areas as compared to suburban and urban areas. Challenges connected to Covid-19 limiting in-person interactions and travel restrictions.	Develop a state-wide model for finding, training, and employing qualified Program Liaisons to work with program families. Find local educational professionals to serve as Liaisons, skilled at adapting and working within their state's preschool and early childhood environment.	*Intended*: Meeting minutes, distributed yearly	*Intended:* Implementation was considered a success if at the end of the grant, planning documents (meeting minutes) for all hiring program liaisons were distributed.	*Intended:* Data collected from meeting minutes and from the Waterford annual report and analyzed at the end of each year of the grant.

Scale-up Goal	Challenge to Meeting Goal	Strategy to Address Challenge	Definition of Full Implementation of Strategy (in measurable units)	Threshold for Level of Implementation Defined as "Successful"	Data Collection and Reporting Plan for Measuring Implementation of Strategy
Goal #3: Outreach and Enrollment Team	Difficulty assembling a team to manage the facilitation of program awareness, interest, and enrollment with a large number of smaller school districts.	Build relationships and hire employees deeply rooted in each state's preschool and early childhood environment who can develop recruitment strategies by creating partnerships with individuals and organizations. Partnerships will help to guide program buy-in efforts within the state.	Waterford will hire 40 employees by the final year of grant to serve on the Outreach and Enrollment Team who can develop recruitment strategies and partnerships.	Implementation was considered a success if 40 employees were hired, and the team created local partnerships in order to meet yearly recruitment goals.	Data collected directly from outreach team members and reported in annual progress reports and analyzed at the end of the grant.

Findings on Implementation of Scaling Strategies

The scale up approaches were intended to provide state-level connections and partnerships for Waterford to leverage with UPSTART implementation beyond the federal grant funded years. The strategies put forth at the beginning of the project proved to be both successful and flexible given the onset of a global health crisis.

Task Force Development: A committee of high-level educational professionals representing each EIR partner state, met consistently each quarter to share knowledge, state-level policy information, and relevant early childhood education updates. These collaborative gatherings took place throughout the duration of the grant from 2019-2022 (transitioning from in-person to virtual). Despite the necessary adjustments, the Task Force development was deemed successfully executed as intended.

Liaisons/Outreach and Enrollment Team: The strategy to create a team of people to assist with local partnerships in the early childhood education space and encourage the adoption of the Upstart program, took a couple of different forms over the course of the grant. The original liaison model was changed into a more centralized outreach and enrollment team, largely created to achieve similar goals. Despite the change to this scale up strategy, it was executed successfully and resulted in a large outreach team developing early education partnerships and fostering state-wide program enrollment efforts. Overall findings are presented in **Table 21**.

Table 21. Overall Findings on Implementation of Scale-up Strategies

Scale-up Strategy	Threshold for Successful Implementation	Findings on Actual Level of Implementation	Implementation of Strategy Met or Exceeded Threshold (Yes/No)	If Implementation of Strategy Did not Meet Threshold, Possible Reasons
Goal #1: Task Force Development	Implementation was considered a success if quarterly TASK force meetings took place and if planning documents (meeting minutes) for all TASK Force meetings were distributed via email.	TASK force meetings took place one time per quarter (over all program implementation years). Meeting minutes for all quarterly meetings were distributed via email.	Yes	
Goal #2: Liaison Development	Implementation was considered a success if at the end of the grant, planning documents (meeting minutes) for all hiring program liaisons were distributed.	Strategy abandoned due to Covid-19.	No	There were a few reasons why this scale up strategy was not met: • The geographical vastness and the volume of school districts across the five EIR states made it difficult to implement a liaison model with impact. These factors made it difficult to find and retain employees for the liaison position. • As a result of the Covid-19 pandemic, Waterford pivoted away from the liaison model and created a larger Outreach and Enrollment team.
Goal #3: Outreach and Enrollment Team	Implementation was considered a success if 40 employees were hired, and the team created local partnerships in order to meet yearly recruitment goals.	Waterford hired 45 employees who created local partnerships and assisted with program recruitment.	Yes	

Summary

Throughout the grant years, there were several challenges that impacted the scale up strategies, ranging from the Covid-19 pandemic to the reality of living and working in rural EIR partner states. Waterford was successful in meeting their scale up goals and employed a flexible approach to working in remote regions of the Great Plains states. While scale up strategies were successfully implemented, challenges surfaced during the program years that required solutions on a much higher level and should be addressed in future efforts. In one of the final meetings of the Task Force, representatives from all five partner states emphasized the importance of pre-kindergarten programs and the increased awareness due to grant activities but mentioned the lack of state funding for preschool as a dominant threat to scaling and sustainability. Stakeholders from several EIR states (Idaho, South Dakota, and Wyoming) communicated that legislative turnover and instability at the state government level led to uncertainty around future preschool funding. Task Force members recommended that Waterford attempt to regain lost momentum, as a result of high-level policy maker turnover, by connecting with and informing new legislators and superintendents moving forward. Stakeholders from the final two states, North Dakota and Montana, however, were not encumbered by political and policy changes, and discussed ongoing partnerships with Waterford beyond the grant.

CHAPTER FOUR

EIR Great Plains Task Force Grant

Taking All to Success in Kindergarten

Cost Effectiveness Study

Cost Effectiveness Study

During 2019-2022, pre-kindergarten children from South Dakota, North Dakota, Wyoming, Idaho and Montana were enrolled in the Upstart program as part of the Education and Innovation Research (EIR) Expansion Grant: *The UPSTART Great Plains TASK Force: Taking All to Success in Kindergarten*. The Evaluation and Training Institute (ETI) conducted a cost effectiveness analysis (CEA) to determine the interventions's cost-per-student and the relationship between the net costs and the effect on student literacy achievement, as measured across multiple literacy outcome variables. The CEA can help identify strategies for minimizing cost while maximizing students' early literacy gains. The remainder of this report includes a brief description of the types of activities included in program costs, a summary of cost-effectiveness by learning domain, and details about the calculations related to the cost-effectiveness estimates.

Activities Covered by Program Costs

The grant included an investment in outreach activities, allowing costs to be allocated to large-scale efforts to register families qualified to participate in the program, which occurred over five rural states. The program vendor counted two types of families served through the grant in their count of families (students) served: outreach families that did not enroll and takeup the program and registered families that enrolled and used the software program[8]. Both types of families went through an outreach and interview process to determine their eligibility prior to

[8] The program vendor was not able to breakout costs between outreach and registration activities and software program use alone.

enrollment, and their were costs associated with marketing and contacting the families with information about the program.

Summary of Cost-Effectiveness by Learning Domain

The cost effectiveness ratio measured how efficiently the Upstart intervention produced an increase in early literacy outcomes by subdomain, by assessing the net cost of the program divided by the effect on student outcomes[9]. Waterford's financial records provided all expenditures related to non-Federal and Federal spending. Annual spending data from Waterford's 2019-2022 yearly grant reports were used to calculate net program costs. A program impact study was used to determing the program's effect on student learning. The calculations presented in **Table 22** are based on statistically significant effect sizes for each subdomain following the intervention (during the preschool year). Letter and Word Recognition had the lowest cost-effectiveness, meaning that raising students ability to identify letters and words cost less than, for example, increasing their reading comprehension skills.

Table 22. Upstart Cost-Effectiveness

Learning Domain	Cost-Effectiveness
Letter and Word Recognition	$2,240.15 per student
Reading Comprehension	$4,650.14 per student
Phonological Processing	$2,758.22 per student

[9] The cost-effectiveness ratio is the average cost per-student divided by the effect size (measured in standard deviation units). Cost-effectiveness can be interpreted as the per-student cost associated with an increase in the student outcome of one standard deviation.

Cost-Effectiveness Calculations: Cost-per Student Rates and Program Impact Estimates

The cost-per student rate was calculated using the following equation:

Expenditures	=	The total of federal, non-federal, and other costs expended in reporting year **MINUS** evaluation costs expended in reporting year
Students served		Number of registered and enrolled students served during the reporting year

Expenditures per year varied based on the number of participating states and the number of students served within the annual cohort[10]. The overall average cost-per student was calculated by multiplying each individual yearly cost-per student estimate by the number of students served that year, including outreach students (who were screened, offered the program but did not enroll) and registered students (who participated in the program), and dividing by the total number of students enrolled from Years 2-4. The average cost-per-student was $1,288.09 across all enrollment years (see **Table 23**).

Table 23. Grant Reporting Per Student Cost by Year

	Cost-Per-Student in Dollars	Registered Students	Outreach but did not enroll Students	Total Students	Detailed Notes on Costs Included in Cost-per-Student[i]
Year 1	No Students				No cost-per student rate calculated
Year 2	1,171.90	723	1,890	2,613	**Net Spending: $3,062,176**
Year 3	1,371.99	2,590	1,228	3,818	**Net Spending: $5,238,268**
Year 4	1,282.91	2,277	958	3,235	**Net Spending: $4,150,222**[ii]
Year 5	No students				No cost-per student rate calculated
Average	1,288.09				

Notes: Net spending for Year 4 was readjusted to exclude supplemental costs used for a summer learning program (SLP) created to extend the program benefits to those impacted by the

[10] Year 2 (Cohort 1): North Dakota, Wyoming, Idaho; Year 3 (Cohort 2): North Dakota, Wyoming, Idaho, South Dakota, Montana; Year 4 (Cohort 3): North Dakota, Wyoming, Idaho, South Dakota, Montana

pandemic. The SLP was not originally planned as part of the grant. The Upstart program expenditures, therefore, included program components allocated for Waterford staff outreach and registration activities, full-year preschool participation such as the provision of computers, internet access, and program software for all families in need. Waterford staff resources were required for parent trainings, recruitment efforts, all virtual Upstart family events, and continued outreach and communication throughout the entire program from dedicated Waterford parent partners by email and phone.

The study had three confirmatory literacy outcomes, including Letter and Word Recognition, Reading Comprehension, and Phonological Processing. The effect sizes (Hedges' g) are listed in **Table 24**, and represent the average impact estimate for each of the literacy outcomes. The individual components of the cost-effectiveness analysis are itemized in **Table 24** to show the impact and cost per student. The cost per student estimates reflect the combined cohort average as defined in the previous section. The standardized effect sizes for each outcome were calculated using Hedges' g based on predicted outcome scores from ordinary least squares multiple regression[11].

Table 24. Student Impacts and Cost-Per-Student by Outcome

Student Outcome Measure	Impact Estimate (Hedges' g)	Cost-Per-Student
Letter and Word Recognition	0.575**	$1,288.09
Reading Comprehension	0.277**	$1,288.09
Phonological Processing	0.467**	$1,288.09

Note: The impact estimate (standardized effect size) was based on cross-cohort preschool year results.
*p<0.05; ** p<0.01

[11] Scales are based on growth score values, or GSVs, which are measures of change over time, or growth, to describe an examinee's absolute level of performance from pre-test to post-test.