Toward an Early Start for Head Start Children: Evidence from a Literacy Enrichment Starting at Age Three

Haiyan Zhang

Department of Psychology, State University of New York at Cortland, New York
https://orcid.org/0000-0002-6994-5059

Abstract. The purpose of this study is to evaluate the utility of providing comprehensive literacy instruction to preschool children starting at age three. The literacy development in two groups of four-year-old children (N = 1320) was compared during their enrollment in a literacy-enriched Head Start program. The first group entered the program at age four and participated in the program for only one year (Group 4-1); the second group had previously participated in the program at age three and at the time of the study were in their second year of the program (Group 4-2). Multilevel growth modelling procedures were used to compare literacy performance levels and growth rates between the two groups after controlling for relevant child, family, and educational variables. Results showed that Group 4-2 exhibited significantly higher literacy skills than Group 4-1 at both the beginning and end of the program year, although Group 4-2 had a slower growth rate than Group 4-1 over the program year. The theoretical and practical implications of the study were discussed.

Keywords: early literacy; Head Start; multi-level growth models; literacy enrichment; age four; age three

1. Introduction

The debate on whether explicit academic instruction in preschool programs, particularly the Head Start program, should be implemented received much attention in past decades and has resulted in changes in performance standards and instructional practices in early childhood settings (NAEYC, 2009; U.S. DDHS ACF, 2016). The central issue in this debate is the nature and substance of what has been called "developmentally appropriate" instruction (Harding et al., 2019; Elkind, 2001; Whitehurst, 2001). Traditional educational philosophies suggest that preschool children benefit most from a child-centered environment that encourages sustained play with concrete objects (Elkind, 2001). According to this view, early academic instruction would interfere with the child's self-

regulated engagement with the educational environment. Consequently, the teaching of preacademic skills such as knowledge of letters and numbers has been discouraged in preschool settings.

With the emergence of evidence-based experimental research, scholars began to examine the content-centered approach involving explicit teaching of early literacy skills in lieu of the traditional child-centered approach (Harding et al., 2019; Lonigan et al., 2000; Storch et al., 2001; Vellutino et al., 2000). The findings motivated changes in the definition of developmentally appropriate practice in early childhood settings. It is now considered developmentally appropriate and even necessary to familiarize preschool children with print concepts and conventions, the utility and various uses of printed words, at least a few letter names and letter sounds, and some high-utility sight words (NAEYC, 2009).

1.1 Literacy Performance and Practice in Head Start

Despite changes in federal mandates and guidelines, studies show that many early childhood educators still lack a comprehensive awareness of the concept of early literacy for children (Harding et al., 2019; Powell et al., 2008; Wasik & Hindman, 2011). Educators tend to hold on to traditional educational philosophies, believing that children naturally develop literacy and that play alone is a developmentally appropriate practice as opposed to implementing learning through play and engaging activities. These misconceptions translate to weak classroom support for literacy and language development in Head Start settings. Such limited support often targets alphabet knowledge whereas key skills such as phonological awareness skills receive limited attention (Castle et al., 2016; Hawken et al., 2005).

Thus, it is essential to provide professional development training in connection with a comprehensive academic curriculum informed by research that can be readily implemented in Head Start classrooms (Castle et al., 2016; Landry et al., 2017).

1.2 Supplementary Literacy Enrichment in Head Start

Several empirical studies have evaluated the utility of incorporating emergent literacy instruction into Head Start programs through curriculum modification and teacher professional development. These studies provide evidence that the explicit teaching of emergent literacy skills can be tailored to the preschool children's level of emotional, social, and cognitive development. For example, Whitehurst (1994) demonstrated that the introduction of a modest literacy component combining interactive book reading and phonological awareness training with the typical Head Start curriculum produced significant increases in children's language and emergent literacy skills. Similarly, Wasik and Hindman (2011) employed a variation of Whitehurst's dialogic reading strategies in the professional development of Head Start teachers using a randomized control design to find that students of teachers in the treatment group performed better on measures of receptive vocabulary and phonological sensitivity than students of teachers in the control group. Bierman (2008) adopted similar strategies in a study using a randomized experimental/control enrichment intervention that

targeted both emergent literacy skills and social-emotional skills to find that children in the intervention classrooms had higher vocabulary, emergent literacy, and learning engagement skills at the end of the program compared with children in the control classrooms. Moreover, Powell et al. (2010) trained teachers in the use of classroom strategies to improve children's oral language skills, phonological awareness skills, and letter knowledge. This resulted in positive effects on classroom support for literacy and language development, as well as increased performance seen in the measures of letter knowledge, blending skills, writing, and print concepts of children in those classrooms. In addition, in an enrichment study conducted by Zimmerman et al. (2008), Head Start children were provided with literacy-related materials and resources in their classrooms. Furthermore, classroom teachers were provided with professional development training to facilitate a better understanding of early literacy development along with the acquisition of strategies for the effective use of such materials. The Head Start children who participated in the enrichment program were found to score above children who received the traditional Head Start program at second and fifth grade standardized reading measures.

Finally, results from a dissertation study are relevant (Pietrangelo, 1999) as the study established the use of a comprehensive literacy curriculum at the Head Start site to provide the data for the present study. Pietrangelo (1999) implemented an emergent literacy curriculum that involved the systematic and direct teaching of a full range of literacy skills facilitated by initial teacher training and weekly support for 14 weeks. The training program provided the teachers with a strong rationale for adding a supplementary literacy component to the Head Start curriculum, along with literacy enrichment activities and daily guidance and mentoring to facilitate the development of emergent literacy skills. Targeted skills included promoting the motivation to read, letter identification, phoneme awareness, at least a rudimentary understanding of the alphabetic principle, print awareness, print conventions, and the ability to identify a few high-frequency words along with the sounds associated with those words. The intervention component of classroom activities consisted of large and small group reading, letter name and letter sound instruction, phoneme awareness activities, and word play activities. The instructional format ranged from formal and explicit instruction to informal games and musical activities. The materials used in the intervention classes consisted of the first 20 books and corresponding component skills and lesson plans provided by the Ready Readers series Stage 0 (Modern Curriculum Press, 1997). Children who received the supplementary literacy instruction were found to perform better on measures of phonological skills than children who received the regular Head Start curriculum at the end of the program.

1.3 Timing of Head Start Participation and Literacy Enrichment
Thus far, the empirical evidence suggests that explicit and comprehensive academic instruction can benefit preschool children in ways that address their developmental needs. However, whether there is added benefit associated with the level of exposure to such instruction remains unclear. Some recent findings from longitudinal studies using national Head Start survey data suggested

additional benefits for academic skills associated with longer exposure to Head Start preschool programs (Zaslow et al., 2010; Zhang & Dobbs-Oates, 2019). However, less is known about whether the same pattern is present with the use of comprehensive literacy instruction. As the development of adequate literacy skills is crucial for the success of children in early intervention programs, the knowledge of when to implement such instruction is important for policy makers, educational researchers, and practitioners alike.

The only two studies that attempted to address this question were conducted by Leow et al. (2015) and Domitrovich et al. (2013). The first study (Leow et al., 2015) analyzed longitudinal national data from the Head Start Family and Child Experiences Survey (FACES). Propensity score analysis was used to control for selection bias. It was found that children who participated in the Head Start program for two years performed significantly better on literacy and math skills compared with children who attended the program for only one year.

The second study (Domitrovich et al., 2013) introduced a comprehensive supplementary curriculum designed to facilitate the development of social-emotional skills and literacy skills in Head Start children. The researchers assessed the children's literacy and math skills once every year and compared the effects of one versus two years of Head Start on kindergarten literacy and math skills. Hierarchical linear modeling procedures were used to account for children nested in classrooms while propensity scores were computed to create two matched groups for comparison. The researchers found that Head Start children who participated in the program starting from age three scored significantly higher on measures of literacy and math skills administered in the fall of kindergarten year than Head Start children who participated in the program starting from age four. However, because assessments occurred once every year, the authors were not able to compare growth trajectories between the two groups; neither were they able to examine the children's academic performance closely during preschool time. In addition, the sample was relatively small (n = 116 per group).

1.4 Present Study
Thus, it would be useful to focus on the preschool period with data collected between shorter time intervals to model children's performance more reliably. Therefore, the question of interest can be formulated: will children who initially entered the Head Start program at age three and received two years of comprehensive literacy instruction perform better on measures of literacy development compared with children who initially entered the program at age four and received only one year of such instruction?

This question was addressed using hierarchical linear modeling procedures (Raudenbush & Bryk, 2002) after controlling for the effects of language and social skills, as well as the effects of other related child, family, and educational factors.

2. Methods

2.1 Participants

Participants in this study included 16 cohorts of children (n = 1320) attending an urban Head Start center-based program between 2005 and 2013. Some of these children entered the program at age four and attended for only one year (Group 4-1; n = 790); others had previously enrolled in the program at age three and were in their second year of the program (Group 4-2; n = 530). The 4-1s and 4-2s were taught in the same classrooms.

Table 1 presents the number of children in each group by year. Cohort analysis identified no significant differences in any of the demographic variables. Table 2 summarizes the results of these variables across the years. The 4-1s and 4-2s had similar demographics (see Table 2). However, because of the selection criteria employed at this Head Start site, the 4-2s came from families that, on average, had lower incomes compared with the 4-1s (t = 2.41, df = 1318, p < .05).

Participants also included 25 teachers. There were 10 classes each year between 2005-08, 2009-10, and 2011-13, and nine classes each year between 2008-09 and 2010-11.

Table 1: Number of participants by age, experience groups and year of assessment

| | Year of Assessment | | | | | | | | |
	2005-06	2006-07	2007-08	2008-09	2009-10	2010-11	2011-12	2012-13	Total
Group 4-1	113	110	113	90	107	78	86	93	790
Group 4-2	51	55	44	70	56	82	92	80	530

Table 2: Descriptive statistics for child level demographic variables

| | | Group 4-2 | | Group 4-1 | |
		N	%	N	%
Ethnicity	Caucasian	147	19.3	106	20.7
	African American	271	35.5	179	34.9
	Hispanic	164	21.5	104	20.3
	Biracial	142	18.6	106	20.7
	Others	39	5.1	18	3.5
Language	English	657	85.8	440	85.3
	Spanish	71	9.3	43	8.3
	Others	38	4.9	33	6.4
Primary caregiver is mother		666	89.3	462	90.1
Single adult in family		348	46.8	249	48.5
More than 2 children in family		385	51.7	264	51.8
Female		412	52.2	276	52.1
		Mean (N)	SD	Mean (N)	SD
Child and Family Risk		6.68(790)	2.92	7.07(530)	2.87
Annual Family Income		11,347.36 (721)	6726.13	10903.58 (488)	7001.77

Note: The maximum possible score for the Child and Family Risk Status is 30.

2.2 Measures

2.2.1 Child and Family Risk Measure

The child and family risk measure included nine items denoting risks inherent to the child (e.g., whether the child was born prematurely, had significant injury or illness, was taking regular medication) and 10 items capturing relevant risks related to the family environment (e.g., single-parent home, disability in members of the family). The raw scores of these items were coded according to a rubric created by two early childhood experts. The rubrics assigned weighted scores representing the estimated levels of risk, with higher numbers indicating higher risk status. The weighted scores were then added to calculate the total scores of the risk measure (ranges 0-30).

2.2.2 Literacy Skills

Literacy skills were assessed using the Phonological Awareness Literacy Screening test (PALS) (Invernizzi et al., 2004) with additional items measuring alphabet knowledge, word identification, and spelling. The Cronbach's alpha figures for the component measures in PALS were: $r = .93$ for Beginning Sound Awareness; $r = .75$ for Print and Word Awareness; $r = .84$ for Rhyme Awareness; and $r = .77$ for Nursery Rhyme Awareness.

The additional items were added to enhance the literacy development spectrum, insufficiently addressed by the PALS for this population. The alphabet knowledge subtest (ranges 0-26) assessed the extent to which children could say or sing the 26 letters of the alphabet. The word identification subtest (ranges 0-20) assessed the children's knowledge of 20 high-frequency printed words. The spelling subtest (ranges 0-21) assessed the children's ability to write letters representing the phonemes constituting five words (*LAP*, *SICK*, *ELEPHANT*, *PRETTY*, and *TRAIN*). One point was given for each phoneme the child was able to produce.

The reliability coefficients for the combined literacy measure were as follows: for 4-1s, $r = .80$ for the first and second assessments and $r = .92$ for the second and third assessments; for 4-2s, $r = .83$ for the first and second assessments and $r = .94$ for the second and third assessments.

2.2.3 Language Skills

The language skills measure assessed each child's ability to comprehend and produce spoken English. It included items assessing the ability to follow classroom routines and answer simple Yes/No questions about self and environment (e.g., "Is it raining outside?", ranges 0-7), follow two- to three-step directions (e.g., "Pick up the pencil. Bounce the ball. Turn over the cup"; ranges 0-12), and answer "wh-questions" (e.g., "What do you wash your hands with?", ranges 0-5). The measure also included items assessing the child's knowledge of the appropriate use of objects (e.g., "You use a chair to____"; ranges 0-3), as well as items assessing the child's ability to communicate with others (e.g., "Does the child start conversations with statements, question, seek attention?"; ranges 0-6). The language skills measure also included items assessing the length of sentences in terms of words used in spontaneous language (ranges 0-20).

Test-retest reliability coefficients for the 4-1s on the language measure were as follows: $r = .77$ for the first and second assessments and $r = .85$ for the second and third assessments; for the 4-2s, $r = .74$ for the first and second assessments and $r = .80$ for the second and third assessments.

2.2.4 Social Skills

The social skills measure consisted of two components. The first component assessed the child's concept of self and family by counting the number of non-repetitive statements he or she was able to make: "Tell me something about you" and "Tell me something about your family" (ranges 0-12). The second component used a five-point scale to measure the frequency of occurrence of a variety of behaviors reflecting specific social-emotional skills. The scale ranged from 0 ("The behavior does not occur") to 5 ("The behavior occurs most of the time"). The categories of behaviors assessed are as follows: independence in classroom activities (ranges 0-5); pride in one's own accomplishments (ranges 0-5); self-control in classroom situations (e.g., "Is able to withhold an angry response and seek help?"; ranges 0-25); cooperation with peers (e.g., "Shows ability to use compromise and discussion in interactions with peers"; ranges 0-15); appreciation of social relationships (e.g., "Responds sympathetically to those who are hurt/upset"; ranges 0-15); knowledge of oneself and one's community (e.g., "Growing awareness of different jobs and who does them and what is required"; ranges 0-15); tendency to initiate activities and discussions (e.g., "Shows eagerness to learn and discuss a growing range of topics, ideas, tasks"; ranges 0-15); tendency to persist in completing tasks (e.g., "Is able to develop goals and follow through on plans"; ranges 0-15); and problem-solving skills (e.g., "Is learning to classify, compare, contrast objects, events, and experiences"; ranges 0-15).

Test-retest reliability coefficients for this combined measure were as follows: for the 4-1s, $r = .86$ for the first and second assessments and $r = .87$ for the second and third assessment; for the 4-2s, $r = .89$ for the first and second assessments and $r = .83$ for the second and third assessments.

2.2.5 Educational Program

The educational program denotes the type of Head Start classrooms. The Head Start classrooms that participated in a district-funded Universal Pre-Kindergarten (UPK) program were coded as 1; the regular Head Start classrooms that did not participate in the UPK program were coded as 0. Upon entering the program for four-year-olds, children were randomly assigned to either a UPK classroom or a regular Head Start classroom. The teachers in the UPK classrooms received teacher professional development training and additional materials not available to the teachers in non-UPK classrooms on a yearly basis.

2.3 Procedures

Eligibility for participation in the Head Start program was initially assessed based on family income and proof of the child's age. In addition, a screening inventory developed by Head Start staff was completed by applicants for the program to help identify children and families with the greatest need. Thus, the

three-year-olds identified with the greatest need were accepted into the Head Start program and participated in the program for two consecutive years (4-2s for the second year). Among the four-year-olds who were new to the program (4-1s), those with the greatest need were accepted into the program and placed in the same classrooms with the four-year-olds who had been enrolled in the program the previous year (4-2s).

All parents of enrolled children were interviewed at the beginning of each program year to establish their risk status. All children were assessed by the program consultant with assistance from the classroom teachers at the beginning (mid-October), middle (mid-February), and end (mid-May) of each program year.

2.4 Data Analysis

Multilevel individual growth models (Raudenbush & Bryk, 2002; Singer & Willett, 2003) were used to analyze the longitudinal data collected from individuals for the purpose of assessing change in outcome variables over time. Estimation of such a model is implemented in two steps. The first step is to choose the Level-1 or within-person component of the model by comparing different growth structures when other time-varying and time-invariant covariates are not included (referred to as the "unconditional growth model"). The second step adds theoretically and empirically grounded covariates at different levels, including time-varying covariates at Level 1 and time-invariant covariates at Level 2 (individual level) and Level 3 (classroom level) to determine whether the predictors are able to forecast both the level of performance (the "intercept") and the rate of growth on the outcome variable (referred to as the "conditional growth model").

3. Results

Results from preliminary analyses (Table 3) showed that the 4-2s performed significantly better than the 4-1s on all three assessments implemented at the beginning ($t = 9.92$, $p < .001$), middle ($t = 6.14$, $p < .001$), and end of the year ($t = 3.75$, $p < .001$). Separate cohort analyses were conducted prior to aggregating the data presented here. No significant differences were identified.

Table 3: Descriptive statistics for the dependent variable (Literacy Skills) and the time-variant covariates (Language and Social Skills) in the growth models separated by experience groups

	Group 4-1(n=790)		Group 4-2 (n=530)		t-tests
	Mean	SD	Mean	SD	4-2s vs 4-1s
Literacy Skills					
Time 1	43.35	26.77	60.07	32.03	9.92***
Time 2	76.92	35.97	89.74	37.97	6.14***
Time 3	99.35	39.89	107.76	39.87	3.75***
Language Skills					
Time 1	37.10	9.64	39.08	8.86	3.78***
Time 2	43.76	7.90	45.10	6.73	3.20**
Time 3	47.15	6.24	47.60	5.82	1.34

			Social Skills		
Time 1	84.21	23.29	88.83	25.04	3.38**
Time 2	100.35	21.06	103.24	21.37	2.43*
Time 3	114.12	19.08	115.13	18.76	.95

Note. The maximum possible score on the measures of literacy, language, and social skills were 192, 53, and 137 respectively. Times 1, 2, and 3 correspond with the beginning, middle, and end of one program year in Head Start.
*** $p < .001$.

3.1 Fitting the Unconditional Growth Models
The unconditional growth models were specified and evaluated to provide baseline measures of literacy skills and literacy growth rates. Results from the unconditional growth models are presented in the first two columns in Tables 5 and 6. First, the mean on the measure of beginning literacy skills was 50.50 ($p < .001$) and children gained an average of 7.48 points per month ($p < .001$) on this measure (see Table 5). In addition, children were found to vary significantly in their ability to acquire literacy skills ($\chi^2 = 14247.60$; $df = 1242$, $p < .001$) and there were substantial differences among their growth rates in acquiring these skills ($\chi^2 = 4440.96$; $df = 1242$, $p < .001$). Similar results were obtained when time was centered at the end of the year (see Table 6). These results suggested that the effects of Level 2 predictors would be usefully evaluated. At the beginning of the program year, it was found that only 6.6% of the variance on literacy performance was between classrooms, leaving 93.4% of the variance within classrooms. However, during the program year, about one-third (31.7%) of the variance on literacy growth was found between classrooms. At the end of the program year, 15.4% of the variance on literacy performance was found between classrooms. These results confirmed the use of a three-level model and suggested that educational characteristics specific to classrooms such as classroom type and teacher characteristics at Level 3 may be explored by explaining the between-class variance on literacy growth.

3.2 Fitting the Conditional Growth Models
Three conditional growth models were examined when time was centered at either the beginning or the end of the program year. The first conditional growth model included two time-varying control variables measuring the child's language and social skills. This model provides baseline measurements. The second conditional growth model added additional control variables at Level 2 and Level 3 of the model, including a child's gender and child and family risk measure at Level 2, and educational program type at Level 3. The third conditional growth model added the variable denoting prior program experience. This is the independent variable of special interest and was coded 1 for 4-2s and 0 for 4-1s. Cohort effects across years were tested. No significant differences were identified across years on either of the two intercepts or growth rates.

Results from preliminary analyses (Table 3) showed that the 4-2s scored significantly higher than the 4-1s on entry-level and mid-year language skills ($t = 3.78$, $p < .001$ and $t = 3.20$, $p < .01$ respectively) as well as entry-level and mid-year social skills ($t = 3.38$, $p < .01$ and $t = 2.43$, $p < .05$ respectively). However, at

the end of the program year, the two groups had comparable scores on both measures ($t = 1.34$, $p > .05$ for language skills; $t = .95$, $p > .05$ for social skills). As shown in Table 4, there were slightly more girls than boys in both groups (52.2% in the 4-1 group and 52.1% in the 4-2 group) and relatively more children in classrooms with UPK affiliation (62% in the 4-1 group and 53.6% in the 4-2 group). The 4-2s had significantly higher levels of child and family risk than the 4-1s ($t = 2.41$, $p < .05$).

Table 4: Descriptive statistics for the time-invariant child, family, and educational variables included in the conditional growth models

		Group 4-1 (n=790)		Group 4-2 (n=530)	
		N	%	N	%
Gender	Female	412	52.2	276	52.1
	Male	378	47.8	254	47.9
Program	UPK	490	62.0	284	53.6
	RHS	300	48.0	246	46.4
		Mean	SD	Mean	SD
Child and Family Risk		6.68	2.92	7.07	2.87

Note. UPK = Head Start with Universal Pre-K affiliation; RHS = Regular Head Start

Results from the third (full) conditional models are presented in the last two columns in Tables 5 and 6. Compared with the baseline model, the variables in the conditional model accounted for an additional 10.6% of the variance on the measure evaluating entry-level literacy skills, 5.7% of the variance on the measure evaluating end-of-year literacy skills, and 3.3% of the variance in literacy growth rates.

Compared with the 4-1s, the 4-2s averaged 16.36 points higher ($p < .001$) on the measure evaluating entry-level literacy skills and 10.53 ($p < .001$) points higher on the measure evaluating end-of-year literacy skills. In addition, all Level 2 variables explained significant variance on the measure evaluating both beginning- and end-of-year literacy skills, as well as the measure evaluating literacy growth rates. Moreover, the Level 3 variable educational program (UPK vs. regular Head Start) explained the significant variance in individual literacy growth rates and end-of-year literacy skills. Compared with children in regular Head Start programs, those who participated in UPK Head Start programs scored 1.85 points ($p < .001$) higher on growth rates over the program year and finished 16.68 points ($p < .001$) higher on end-of-year literacy skills. However, contrary to expectations, the 4-2s had an average literacy growth rate that was 0.83 points per month lower ($p < .05$) than the score of the 4-1s.

Table 5: Growth parameters for literacy skills examining the effects of prior program experience on beginning-of-year scores and growth rates (time centered at the first assessment)

	Unconditional Model		Full Conditional Model	
	Coefficient	SE	Coefficient	SE
Fixed Effects				
Initial Status				
Mean Status†	50.50***	1.19	44.95***	1.79
Program (UPK=1)			3.73	2.11
Gender (Female=1)			6.89***	1.35
Child and Family Risk			-4.14***	0.71
Group 4-2 vs. 4-1			16.36***	1.62
Time Slope				
Mean Time Slope	7.48***	0.25	4.68***	0.34
Program (UPK=1)			1.85***	0.39
Gender (Female=1)			0.49*	0.21
Child and Family Risk			-0.23*	0.11
Group 4-2 vs. 4-1			-0.83***	0.21
Language Slope			0.62***	0.13
Social Slope			0.25***	0.04
	Variance	χ2 (df)	Variance	χ2 (df)
Random Effects				
Level 1 and Level 2				
Individual Mean	808.09	14247.60***(1242)	684.06	13001.53***(1239)
Time Slope	9.17	4440.96***(1242)	8.51	4412.07***(1239)
Level 1 Effect	87.82		82.30	
Level 3				
Class Mean	56.98	162.87***(77)	44.38	155.20***(76)
Class Time Slope	4.25	519.13***(77)	2.27	333.37***(76)
Deviance (df)	35344.46 (9)		34916.92 (19)	

Note. Program = Universal Pre-Kindergarten (1) vs. Regular Head Start (0)

* p < .05; ** p < .01; *** p < .001.

Table 6: Growth parameters for literacy skills examining the effects of prior program experience on end-of-year scores and growth rates (time centered at the third assessment)

	Unconditional Model		Full Conditional Model	
	Coefficient	SE	Coefficient	SE
Fixed Effects				
Initial Status				
Mean Status	102.84***	2.08	77.68***	3.14
Program (UPK=1)			16.68***	3.56
Gender (Female=1)			10.33***	2.08
Child and Family Risk			-5.72***	1.07
Group 4-2 vs. 4-1			10.53***	2.15
Time Slope				
Mean Time Slope	7.48***	0.25	4.68***	0.34
Program (UPK=1)			1.85***	0.39
Gender (Female=1)			0.49*	0.21
Child and Family Risk			-0.23*	0.11
Group 4-2 vs. 4-1			-0.83***	0.21
Language Slope			0.62***	0.13

	Unconditional Model		Full Conditional Model	
	Coefficient	*SE*	Coefficient	*SE*
Social Slope			0.25***	0.04
	Variance	χ2 (*df*)	Variance	χ2 (*df*)
Random Effects				
Level 1 and Level 2				
Individual Mean	1382.88	26186.44***(1242)	1336.50	26981.96***(1239)
Time Slope	9.17	4440.96***(1242)	8.51	4412.07***(1239)
Level 1 Effect	87.82		82.30	
Level 3				
Class Mean	251.07	306.76***(77)	161.42	230.52***(76)
Class Time Slope	4.25	519.13***(77)	2.27	333.37***(76)
Deviance (*df*)	35344.46 (9)		34916.92 (19)	

Note. Program = Universal Pre-K (1) vs. Regular Head Start (0)

$^{*} p < .05$; $^{**} p < .01$; $^{***} p < .001$.

4. Discussion

The present study investigated the utility of providing comprehensive and explicit literacy instruction for children starting at age three. The effect of program experience on literacy development at age three was evaluated by comparing the literacy performance of two groups of four-year-old Head Start children with and without prior Head Start experience at age three. The beginning and end-of-year performance levels and the growth trajectory of each child during the program year were assessed to obtain these results. Three-level growth models were used, and a set of relevant covariates were controlled, including individual child characteristics, family background, and educational characteristics.

4.1 Summary of Findings

Consistent with existing literature indicating additional benefits on academic skills associated with longer exposure to Head Start preschool programs (Zaslow et al., 2010; Zhang & Dobbs-Oates, 2019), the results showed that prior program experience helped promote children's literacy skills. Specifically, after controlling for relevant covariates, the 4-2s were found to perform significantly better than the 4-1s on the literacy measure at the beginning and end of the program year. In contrast to the expectation that the 4-2s would have similar or higher rates of literacy growth compared with the 4-1s, the 4-1s were found to have higher growth rates than the 4-2s. This finding may be because the two groups were taught by the same teachers in the same classrooms and that the 4-1s had much more room to grow and possibly received more instructional attention due to lower levels of performance throughout the program year, when compared with the 4-2s.

Among the covariates examined in the study, the child and family risk variable was found to be negatively related to the beginning- and end-of-year literacy performance and literacy growth. These findings indicate that literacy growth in higher-risk children was significantly slower than literacy growth in lower-risk children during their participation. This is consistent with previous research findings that children who experience more risk would likely experience

increased challenges in early literacy acquisitions (Gerstein et al., 2021; Hooper et al., 2020). Therefore, the variable denoting risk status would serve as a useful control variable in analyses of literacy performance.

Regarding the other variables used to predict literacy performance, it is interesting to note that the measures of language and social skills were found to serve as significant predictors of literacy growth rates over the program year. Home background and environmental stimulation are factors contributing to the acquisition of language, social, and early literacy skills (Gerstein et al., 2021; Lonigan et al, 2000; Storch et al., 2001; Whitehurst et al, 1994), and it is possible that variability in these factors contributed to the degree of variability and intercorrelation among the language, social, and literacy skills measures that facilitated the prediction of literacy skills.

It is of particular importance to note that the educational program (UPK vs. regular Head Start) explained significant variances in individual literacy growth rates and end-of-year literacy performance. Compared with children in regular Head Start programs, those who participated in UPK programs showed higher growth rates and end-of-year literacy performance. This may be because teachers in the UPK program had higher qualifications and received more technical support in terms of educational materials and teacher training opportunities than teachers in regular Head Start classrooms. This finding is consistent with recent literature highlighting the need for professional development of in-service teachers in Head Start programs (Castle et al., 2016; Landry et al., 2017).

4.2 Theoretical and Practical Implications

Results from the present study have both theoretical and practical implications for researchers, educators, and policy makers. First, the findings support the view that early exposure to comprehensive and explicit literacy instruction is pivotal for children from disadvantaged backgrounds (Gerstein et al., 2021; Landry et al., 2017; Whitehurst, 2001). In addition, the finding that the immediate benefits generated by early exposure to the Head Start program are sustained a year later is consistent with findings from previous studies (Domitrovich, 2013; Leow et al., 2015) and can be taken as support for the implementation of Head Start programs for three-year-old children.

Second, the findings support the recruitment of teachers with higher educational degrees and credentials as well as the provision of in-service professional development for teachers in Head Start programs (NAEYC, 2009). It is widely accepted that the competencies of children are dependent on the quality of their experiences in educational settings, which is in turn dependent on the competencies and characteristics of their teachers (Castle et al., 2016; Landry et al., 2017; Pianta et al., 2008). The UPK versus regular Head Start program affiliation variable was found to be positively associated with literacy growth rates and the end-of-year literacy performance. These findings are quite likely because those who participated in the school district-funded UPK programs were taught by more teachers with graduate degrees, the New York State

certification, and who had received yearly professional development support, compared with the classrooms that did not participate in the UPK program. Thus, this pattern of results may be taken as evidence in support of in-service professional development training for Head Start teachers.

4.3 Limitations and Future Research

Despite the strengths outlined in the preceding sections, the present study has several limitations. First, the study was not designed as an experimental study because it was not possible to assign children randomly to treatment groups for purposes of comparison. Therefore, observed differences between Head Start experience groups can only be generalized to populations of children such as those assessed in the present study. Similarly, there was no data available on measures of relevant control variables for 4-1s at three years of age. The control variables used in the study were measured at the beginning of the 4-1s' first and only year in the Head Start program and at the beginning of the 4-2s' second year in the program. Because the scores for the 4-2s were influenced by prior participation in the program, and thus were relatively higher compared with those for the 4-1s, group differences in literacy development, as a function of prior program experience, might have been larger if performance on the control variables had been assessed when the 4-1s and 4-2s were both three years old.

Second, it was necessary to compromise on the control for educational characteristics. More specifically, the UPK versus regular Head Start program affiliation indicator was a useful estimate; however, it did not sufficiently capture variability in teacher effectiveness. Thus, it would be useful to conduct individual teacher interviews in the future to obtain more information about Head Start teachers' beliefs and willingness to teach early literacy skills explicitly in their classrooms (Powell et al., 2008). Classroom observations could also be conducted using a comprehensive instrument to evaluate classroom climate (e.g., the Classroom Assessment Scoring System, or CLASS- Pre-K) (Pianta et al., 2008) regarding how and to what extent teachers' beliefs transfer to teaching practices. Such data could be cross-validated with children's experience in Head Start programs (e.g., second year vs. first year) to evaluate the differential effects of these variables on measures of literacy development.

Finally, all the child, family and educational variables combined accounted for a relatively small portion of the variance in literacy performance and growth rates. One possible source of this relatively low-level of explained variance is the reduced variability on the risk measures, possibly due to an established health system in this Head Start program. The majority of children who participated in the program had access to regular medical care and their families received quality community services. Thus, it may be reasonable to suggest that in a Head Start program with a less established health system, these risk variables might explain more of the variance among individual children's literacy growth rates. In addition, other variables not yet identified may also influence the rate of growth at which children acquire early literacy skills. For example, home and background factors, two potentially important contributors to such growth,

were not included in the model and quality of instruction (Gerstein, 2021; Zaslow et al., 2010).

5. Conclusions

In summary, the present study assessed the utility of providing comprehensive and explicit literacy instruction to children starting at age three. Consistent with previous research findings (Domitrovich et al., 2013; Leow et al., 2015), the results support the likelihood of additional benefits associated with early exposure to the program. It was found that four-year-olds who had received an additional year of the program at age three exhibited significantly higher literacy skills than their four-year-old age mates who entered the Head Start program at age four. The results also support the view that it is appropriate (NAEYC, 2009) and even necessary to provide explicit and comprehensive literacy instruction to both three- and four-year-old children in Head Start programs. In addition, the study extends previous research on curriculum modifications and teacher professional development in Head Start programs (Landry et al., 2017; Powell et al., 2010; Wasik & Hindman, 2011; Zimmerman et al., 2008) by providing a tested model for the gradual integration of a research-based supplementary curriculum through continued support for teacher professional development that features the explicit and comprehensive teaching of emergent literacy skills to disadvantaged preschool children.

6. References

Bierman, K. L., Domitrovich, C. E., Nix, R. L., Gest, S. D., Welsh, J. A., Greenberg, M. T., & Gill, S. (2008). Promoting academic and social-emotional school readiness: The Head Start REDI program. *Child Development, 79*, 1802-1817. https://doi.org/10.1111/j.1467-8624.2008.01227.x

Castle, S., Williamson, A. C., Young, E., Stubblefield, J., Laurin, D., & Pearce, N. (2016). Teacher–child interactions in Early Head Start classrooms: Associations with teacher characteristics. *Early Education and Development*, 27, 259–274. https://doi.org/10.1080/10409289.2016.1102017

Domitrovich, C. E., Morgan, N. R., Moore, J. E., Cooper, B. R., Shah, H. K., Jacobson, L., & Greenberg, M. T. (2013). One versus two years: Does length of exposure to an enhanced preschool program impact the academic functioning of disadvantaged children in kindergarten? *Early Childhood Research Quarterly, 28*, 704–713. https://doi.org/10.1016/j.ecresq.2013.04.004

Elkind, D. (2001). Much too early. *Education Next, 12*, 8-14.

Gerstein, E. D., Cherry, K. E., & Diemer, M. C. (2021). Identifying developmental profiles for children in Early Head Start: Associations with early risk and protective factors. *Early Childhood Research Quarterly*, 57, 191–201. https://doi.org/10.1016/j.ecresq.2021.06.007

Hawken, L. S., Johnston, S. S., & McDonnell, A. P. (2005). Emerging literacy views and practices: Results from a national survey of Head Start preschool teachers. *Topics in Early Childhood Special Education, 25*, 232-242. https://doi.org/10.1177/02711214050250040401

Hooper, S. R., Costa, L.-J. C., Green, M. B., Catlett, S. R., Barker, A., Fernandez, E., & Faldowski, R. A. (2020). The relationship of teacher ratings of executive functions to emergent literacy in Head Start. *Reading and Writing: An Interdisciplinary Journal*, 33, 963–989. https://doi.org/10.1007/s11145-019-09992-1

Invernizzi, M., Sullivan, A., Meier, J. D., & Swank, L. (2004). *PALS: Phonological awareness literacy screening*. University Printing Services.

Landry, S. H., Zucker, T. A., Williams, J. M., Merz, E. C., Guttentag, C. L., & Taylor, H. B. (2017). Improving school readiness of high-risk preschoolers: Combining high quality instructional strategies with responsive training for teachers and parents. *Early Childhood Research Quarterly*, 40, 38–51. https://doi.org/10.1016/j.ecresq.2016.12.001

Leow, C., Wen, X., & Korfmacher, J. (2015). Two-year versus one-year Head Start program impact: Addressing selection bias by comparing regression modeling with propensity score analysis. *Applied Developmental Science, 19*, 31–46. https://doi.org/10.1080/10888691.2014.977995

Lonigan, C. J., Burgess, S. R., & Anthony, J. L. (2000). Development of emergent literacy and early reading skills in preschool children: Evidence from a latent-variable longitudinal study. *Developmental Psychology, 36*, 596-613. https://doi.org/10.1037/0012-1649.36.5.596

Modern Curriculum Press. (1997). *Ready readers stage 0*. Modern Curriculum Press.

National Association for the Education of Young Children (NAEYC). (2009). *Developmentally appropriate practice for early childhood program serving children from birth to age 8*. NAEYC.

Pianta, R. C., La Paro, K., & Hamre, B. K. (2008). *Classroom assessment scoring system*. Paul H. Brookes Publishing.

Pietrangelo, D. J. (1999). *Outcomes of an enhanced literacy curriculum on the emergent literacy skills of Head Start preschoolers* (Doctoral dissertation). University at Albany – SUNY.

Powell, D. R., Diamond, K. E., Bojczyk, K. E., & Gerde, H. K. (2008). Head Start teachers' perspectives on early literacy. *Journal of Literacy Research, 40*, 422-460. https://doi.org/10.1080/10862960802637612

Powell, D. R., Diamond, K. E., Burchinal, M. R., & Koehler, M. J. (2010). Effects of an early literacy professional development intervention on Head Start teachers and children. *Journal of Educational Psychology, 102*, 299-312. https://doi.org/10.1037/a0017763

Raudenbush, S. W., & Bryk, A. S. (2002). *Hierarchical linear models: Applications and data analysis methods*. Sage.

Singer, J. D., & Willett, J. B. (2003). *Applied longitudinal data analysis: Modeling change and event occurrence*. Oxford University Press.

Storch, S. A., Whitehurst, G. J., Britto, P. R., & Brooks-Gunn, J. (2001). The role of family and home in the literacy development of children from low-income backgrounds. In P. R. Britto & J. Brooks-Gunn (Eds.), *The role of family literacy environments in promoting young children's emerging literacy skills* (pp. 53-71). Jossey-Bass.

United States Department of Health and Human Services, Administration for Childen and Families (US DHHS, ACF). (2016). *Head Start program performance standards*. US DHHS, ACF.https://eclkc.ohs.acf.hhs.gov/sites/default/files/pdf/hspps-appendix.pdf

Vellutino, F. R., Scanlon, D. M., & Lyon, G. R. (2000). Differentiating between difficult-to-remediate and readily remediated poor readers: More evidence against the IQ-achievement discrepancy definition of reading disability. *Journal of Learning Disabilities, 33*, 223-238. https://doi.org/10.1177/002221940003300302

Wasik, B. A., & Hindman, A. H. (2011). Improving vocabulary and pre-literacy skills of at-risk preschoolers through teacher professional development. *Journal of Educational Psychology, 103*, 455-469. https://doi.org/10.1037/a0023067

Whitehurst, G. J. (2001). Much too late. *Education Next, 12*, 15-21.

Whitehurst, G. J., Epstein, J. N., Angell, A. L., Payne, A. C., Crone, D. A., & Fischel, J. E. (1994). Outcomes of an emergent literacy intervention in Head Start. *Journal of Educational Psychology, 86*, 542-555. https://doi.org/10.1037/0022-0663.91.2.261

Zaslow, M., Anderson, R., Redd, Z., Wessel, J., Tarullo, L., & Burchinal, M. (2010). *Quality dosage, thresholds, and features in early childhood settings: A review of the literature, OPRE 2011-5.* Office of Planning, Research and Evaluation, US DHHS, ACF.

Zhang, C., & Dobbs-Oates, J. (2019). The relations between American children's Head Start experience and pre-academic skills: A comparison with children from a community group. *Journal of Research in Childhood Education*, *33*, 623-636. https://doi.org/10.1080/02568543.2019.1649769

Zimmerman, S. S., Rodriguez, M. C., Rewey, K. L., & Heidemann, S. L. (2008). The impact of an early literacy initiative on the long-term academic success of diverse students. *Journal of Education for Students Placed at Risk, 13*, 452-481.