Published in final edited form as:
Child Dev. 2014 November ; 85(6): 2317–2338. doi:10.1111/cdev.12278.

Who Benefits Most from Head Start? Using Latent Class Moderation to Examine Differential Treatment Effects

Brittany Rhoades Cooper and
Human Development, Washington State University

Stephanie T. Lanza
Methodology Center, Pennsylvania State University

Abstract

Head Start is the largest federally funded preschool program for disadvantaged children. Research has shown relatively small impacts on cognitive and social skills; therefore some have questioned its effectiveness. Using data from the Head Start Impact Study (3-year-old cohort; N = 2,449), we used latent class analysis to (1) identify subgroups of children defined by baseline characteristics of their home environment and caregiver and (2) test whether the effects of Head Start on cognitive, and behavioral and relationship skills over two years differed across subgroups. The results suggest that the effectiveness of Head Start varies quite substantially. For some children there appears to be a significant, and in some cases, long-term, positive impact. For others there is little to no effect.

Childhood poverty rates in the United States (U.S.) have increased significantly in the past few decades. In 2010, 21.6% were living at or below the poverty threshold (Macartney, 2011; McLoyd, Kaplan, Purtell, & Huston, 2011; McLoyd, 1998). The detrimental effects of poverty are well documented; children who experience poverty are at increased risks for poor social and academic outcomes (Alexander & Entwisle, 1988; Belsky & MacKinnon, 1994; Brooks-Gunn & Duncan, 1997; Campbell, March, Pierce, Ewing, & Szumowski, 1991; Raver, Blair, & Willoughby, 2013; McLoyd et al., 2011; Patterson, Forgatch, Yoerger, & Stoolmiller, 1998; Ryan, Fauth, & Brooks-Gunn, 2006; Smith, Brooks-Gunn & Klebanov, 1997; Yoshikawa, Aber, & Beardslee, 2012). Low-income children as young as 2 perform significantly worse on cognitive andlanguage tasks than their higher income peers (National Institute of Child Health and Human Development Early Child Care Research Network, 2005). This gap persists into preschool and places low-income children at a disadvantage for the demands of school (McClelland, Acock, & Morrison, 2006). Therefore, preschool programs have been highlighted as a critical, cost-effective early intervention strategy, especially for disadvantaged children (Belfield, Nores, Barnett, & Schweinhart, 2006).

Ecological systems theory (Bronfenbrenner & Morris, 1998) outlines multiple interacting contexts that influence children's development. In early childhood, the primary influences

Correspondence concerning this article should be addressed to Brittany Rhoades Cooper at 510 Johnson Tower, Pullman, WA 99164. brittany.cooper@wsu.edu.

are the home environment (including the characteristics and experiences of primary caregivers), and caregiving and educational settings. In line with this theoretical perspective, extant research indicates that quality, responsive preschool programs can reduce the achievement gap by improving cognitive, academic, and long-term adjustment outcomes (Campbell, Pungello, Miller-Johnson, Burchinal, & Ramey, 2001; Ramey & Ramey, 2004; Reynolds & Temple, 1998). Recent meta-analyses found that early education programs have small to moderate positive effects on cognitive skills, school progress, social-emotional development, and behavior (Burger, 2010; Camilli, Vargas, Ryan, & Barnett, 2010). Head Start (HS) was first launched as an anti-poverty effort in 1965 to improve the social competence of low-income children and help prepare them for elementary school (Zigler & Styfco, 1994). HS takes a comprehensive, "whole child" approach: it provides physical health, mental health, and social services for parents and children. However, the efficacy of HS remains unclear. In 1998, Congress mandated the first nationally representative, randomized evaluation of HS, the HS Impact Study, implemented in 2002–2006. Data from this study have recently become available to researchers, providing a unique opportunity to investigate efficacy. The present study builds upon existing research and the growing literature on differential treatment effects by examining variance in HS efficacy across subgroups of children with different combinations of socio-demographic characteristics.

Rather than examining individual characteristics separately with traditional moderation analyses (e.g., Aiken & West, 1991), we use a latent class framework first to empirically identify subgroups of children defined by profiles across nine characteristics of their home environment, including home composition, language, and food stamp status, and their primary caregiver, including mother's immigrant status, depression, age at study child's birth (teen mother), education, and employment status. Then, we examine whether treatment effects vary across these subgroups. The choice of indicators was based on the socio-demographic factors included in the initial subgroup analysis mandated by Congress in the original HS Impact Study (USDHHS-ACF, 2010). A traditional moderation framework is ideal for focusing in on differential effects of HS based on a single characteristic of the child or family (e.g., low maternal education). A latent class moderation approach can complement this understanding by enabling a comprehensive examination of types of children–defined by a broad set of child and family characteristics–that simultaneously characterize the child's developmental context. This approach adds to and broadens the perspective of prior studies.

Prior Research on the Impact of Head Start on Child Skills & Competencies

Research on HS efficacy is mixed: some studies show few to no positive effects, others show significant but small positive effects on child competencies and skills (Aughinbaugh, 2001; Barnett, 2011; Blau, 2001; Currie & Thomas, 2000; Currie, 2001; Deming, 2009; Garces, Thomas, & Currie, 2002; Gormley, Phillips, Adelstein, & Shaw, 2010; Ludwig & Miller, 2007). Some early studies examined the impact of HS on academic skills using quasi-experimental (i.e., nonrandomized) designs. They concluded that 1 year of HS had a positive effect on children's cognitive skills, but these small gains typically declined and faded out completely by 1st grade (e.g., Haskins, 1989; Lee, Brooks-Gunn, Schnur, & Liaw, 1990).

The HS Impact Study was designed to fill the need for high quality, nationally representative research on HS efficacy. Overall, the study found relatively small main effects of HS for children's school readiness skills (average effect sizes around .17, rarely exceeding .30). Specifically, children who entered HS at age 3 (age 3 cohort) and those who entered at age 4 (age 4 cohort) had improved language and pre-literacy skills, compared to control children (U.S. Department of Health & Human Services, Administration for Children and Families, 2010). The age 3 cohort also showed additional benefits in math, perceptual motor skills, pre-writing, and behavior. These gains, however, did not persist. The age 4 cohort outperformed non-HS children only in receptive vocabulary at the end of 1st grade; the age 3 cohort outperformed only in oral comprehension (USDHHS-ACF, 2010). However, traditional main effects analyses may obscure the benefit of HS for specific subgroups of children (Bloom & Michalopoulos, 2012), limiting our understanding of program efficacy for the diverse population of low-income families in HS.

Differential effects of HS on child skills and competencies

The recent emphasis on targeting interventions to those who will benefit most has led to increased interest in differential treatment analyses (i.e., subgroup analyses or moderation analyses; Supplee, Kelly, MacKinnon, & Barofsky, 2013). The Congressional mandate that initiated the HS Impact Study tasked evaluators with exploring variance across seven dimensions: (1) lowest quartile of pre-academic skills at baseline, (2) dual-language learner at baseline, (3) parent-reported special needs, (4) mother/caregiver race/ethnicity, (5) parent/caregiver depression, (6) household risk (sum of receipt of TANF or food stamps, neither parent with high school diploma or GED, neither parent employed or in school, biological mother is single parent, biological mother was a teen when child was born) and (7) urbanicity (USDHHS-ACF, 2010). Although they are highly related, each of these characteristics were examined in turn. A consistent pattern emerged for several subgroups (USDHHS-ACF, 2010). For example, the impact of HS was moderated by parental depression for both cohorts. On average, children in the age 4 cohort whose parents had mild depressive symptoms benefit more from HS than those whose parents had no symptoms. Children in the age 3 cohort whose parents had no depressive symptoms benefitted most from HS; those whose parents had mild or moderate depressive symptoms showed consistent, unfavorable impacts. Differential treatment effects for both cohorts were also evident for dual-language learners—they experienced greater benefits from HS than children who spoke only English. For the age 4 cohort, USDHHS-ACF (2010) noted significantly better effects for Black children and children with fewer skills. For the age 3 cohort, children with special needs, children exposed to greater household risk, and children in non-urban settings had stronger beneficial effects. These analyses indicate that HS does impact children differently, but the pattern is difficult to interpret.

Other studies show that children at higher risk in terms of behavior or home context benefit more from early intervention than lower-risk children (Bierman et al., 2010; Bierman, Nix, Greenberg, Blair, & Domitrovich, 2008; Conduct Problems Prevention Research Group, 2011; Lee, 2003; Richardson et al., 1999). For example, several studies found that children of immigrants and Spanish-speaking dual language learners who attend preschool may improve more in terms of academic skills (Farver, Lonigan, & Eppe, 2009; Gormley, 2008;

Magnuson, Lahaie, & Waldfogel, 2006). Research with HS children has shown that reading scores of children whose mothers had less education and lower verbal abilities benefit more from entering HS at age 3 than age 4 (Lee, 2008) and that the association between duration of HS and children's skills was moderated by the number of family risks (Lee, 2011). For example, among children with 4 or more family risks (out of 15 possible risks), those who had 2 years of HS had higher math scores than those who had 1 year; HS was not related to math skills for children with 3 or fewer family risks. Baseline academic score also moderated the effect of dose on outcome—children with lower initial scores benefit more from longer HS. Lee (2011) also uncovered a three-way interaction between duration, family risk, and baseline academic score: positive impacts of two years of HS were stronger for children with greater family risk and lower baseline academics. Overall, Lee argues that two years of HS can help compensate for lack of a high quality, nurturing, stimulating home environment.

Zhai, Brooks-Gunn, & Waldfogel (2011) examined the impact of HS on children's cognitive and social skills at age 5 in a large, diverse sample of low-income families from the Fragile Families and Child Wellbeing Study (18 low-income, urban cities). Using propensity score matching to account for potential selection effects, Zhai and colleagues compared the effectiveness of HS to multiple non-HS care arrangements. Overall, effects of HS at age 5 replicated the initial short-term HS Impact Study findings, suggesting a modest positive impact of HS on cognitive skills, attention, and social competence (no significant effect on externalizing or internalizing behaviors). Further, they found that the impact of HS varied depending on the care arrangement of the control group. In general, they found that HS children who would otherwise have been in parental or non-parental (e.g., non-center setting) care had more positive outcomes than the average of all non-HS children. However, HS did impact some non-cognitive domains in comparison to children who attended pre-kindergarten programs (on social competence) and center-based care (on social competence, attention and behavior problems). This suggests that HS may be most effective for children who would have otherwise received parental or non-parental care, which may be those families with fewer economic resources.

In summary, existing research on differential effects of preschool suggests that household and maternal characteristics, which affect level of risk for future academic and social difficulties, may be important in the prediction of response to HS. Observational research consistently demonstrates a negative association between fewer economic and social resources (e.g., single parenthood, receipt of public assistance, teen motherhood, low maternal education, maternal depression) and children's academic and social competencies (e.g., Brooks-Gunn & Duncan, 1997; Campbell, March, Pierce, Ewing, & Szumowski, 1991; Raver, Blair, & Willoughby, 2013; McLoyd et al., 2011; Yoshikawa, Aber, & Beardslee, 2012). Therefore, the present study will examine how such factors are associated with differential treatment effects of HS.

Advancing Methodological Approaches to Differential Treatment Analyses

Given the importance of differential treatment (moderation) analyses for early childhood education, it is crucial that we refine our methodological strategies for identifying and

understanding the characteristics that predict which children are likely to benefit most. In traditional differential treatment analyses, a single variable (e.g., gender, race) is incorporated into a multiple regression model as a moderator (i.e., main effects and the treatment × individual characteristic interaction are included in a regression model predicting an outcome; Aiken & West, 1991). Most of the studies described above used this approach to determine how impact of early childhood interventions vary across groups of children (e.g., Bierman et al., 2008; CPPRG, 2011; Lee, 2011). Standard moderation analyses can be particularly informative, as they permit researchers to estimate the overall differences in the effect of HS across a single dimension such as child's language or maternal education.

Latent class moderation (i.e., latent classes defined by the intersection of multiple potential moderators) has different advantages (Lanza & Rhoades, 2013; Wang & Ware, 2013) and can address some of the limitations to the traditional approach (e.g., low statistical power for examining higher-order interactions). A primary strength of latent class analysis (LCA) is its ability to empirically identify population subgroups based on the most common combinations of a particular set of variables. In the present study, it can identify subgroups comprised of like individuals (i.e., children) who share a particular combination of co-occurring social risk factors, and who may respond differently to a particular treatment (e.g., Head Start). In comparison to variable-centered approaches, like cumulative risk indices, the LCA approach provides a more comprehensive, person-centered picture of both the type and quantity of risk factors in a population (Lanza, Rhoades, Greenberg, Cox, & The Family Life Project Key Investigators, 2011; Lanza, Rhoades, Nix, Greenberg, & The Conduct Problems Prevention Research Group, 2010).

Because baseline characteristics do not operate in isolation, it can be beneficial to consider individuals' configurations on the entire set when examining differential effects of HS. Latent class moderation does just that; we introduce this approach to empirically identify the most common subgroups of HS children and to determine which subgroups benefit more from HS than others.

The Present Study

Existing literature on differential treatment effects suggests that household and maternal factors that place children at higher levels of risk are differentially impacted by early intervention. To build on this research and the original evaluation of differential effects in the HS Impact Study, we examine nine baseline characteristics[1] that may help define groups of children who benefit more or less from HS: three characteristics of the home environment (home composition, language, and food stamp status) and six characteristics of the mothers (marital status, immigrant status, depression, age at first birth, education, and employment status) at baseline. We use a latent class framework, described in detail below, to (1) empirically identify subgroups of children defined by their profiles across nine characteristics and (2) test whether the effects of HS on cognitive, and behavioral and

[1]Preliminary latent class models also included children's pre-academic skills and special needs status at baseline. However, because they did not clearly distinguish classes, they were eliminated in further analyses.

relationship skills at three time points differed across these subgroups. As mentioned earlier, a latent moderation approach allows us to characterize children by multidimensional profiles based on these nine characteristics. This approach will advance our understanding of the types of children who benefit most from HS.

Based on previous research, we hypothesize that subgroups composed of children of immigrants (particularly those who speak a language other than English at home) and subgroups composed of children with single mothers (particularly those with low education) will benefit most from HS. We expect that a subgroup characterized by both of these characteristics would particularly benefit from HS in terms of vocabulary and reading outcomes. We also hypothesize that subgroups characterized by relatively lower levels of risk (e.g., living in a married household, children of mothers higher than a high school diploma) will benefit least from HS.

Methods

Participants & Procedure

Data for the present study came from the HS Impact Study. This large-scale study included a nationally representative sample of 84 randomly selected HS grantees across 23 states in 2002, from which 383 HS centers were randomly selected for participation. A total of 4,442 children (2,449 age 3, 1,993 age 4) from these centers were randomly assigned to either the treatment group, which was given access to HS, or the control group, which was not but could enroll in other non-HS programs as desired. Control group children were most likely to receive care by parent or another relative (50%). Data were collected via direct assessments of children's cognitive skills and parent- and teacher-rated social-emotional behaviors and adult-child relationship quality from fall 2002 to spring 2006, assessing children at either age 3 or 4 in fall and in the spring of the preschool year, kindergarten and 1st grade. This design enabled us to examine the differential long-term impact of HS two years after their exposure to HS. For more details on the study sample, evaluation design, and so forth, see USDHHS-ACF (2010).

Given the stronger effect for children who enter HS early (e.g., USDHHS-ACF, 2010), the present study focused on data from the age 3 cohort (N = 2,449); 1464 children were randomly assigned to receive HS and 985 were assigned to control. As in any randomized control trial (RCT) under real-world conditions, violations of randomization occurred for a small proportion of the sample. For the age 3 cohort, 14.9% of the children assigned to receive HS were "no-shows" and did not receive any HS; about 17.3% of the control group were "crossovers" and did participate in HS during the first year of the study (USDHHS-ACF, 2010). Intent-to-treat analysis (i.e. the estimated effect of the treatment to which children were assigned) is recommended as the most appropriate method for determining an unbiased estimate of treatment efficacy (Flay et al., 2005) and therefore was used in the present study to determine differential treatment effects. Because these violations of randomization were relatively minimal and because HS and control groups were deemed comparable in terms of baseline characteristics, the only effect they should have is to possibly underestimate the program's effects (USDHHS-ACF, 2010). Also, due to the ethical concerns related to withholding services from the control group, children were not

prohibited from enrolling in HS during the second year. As a result, about 50% of the control group (compared to 63% of the HS group) attended HS during the second year (age 4) and, therefore, the present analyses represent a test of the differential treatment effect of having 1 year of HS at age 3.

The HS Impact Study provided sample weights for each assessment time so that outcome analyses relying on data from spring of preschool, spring of kindergarten, and spring of 1st grade were nationally representative in terms of the national population of entering HS children and their families for the 2002–2003 year. Thus, sample sizes for our outcome analyses vary by outcome time point (and outcome assessment) but remained representative. Attrition analyses showed no significant differences between children who remained in the study compared to those who dropped out on all nine baseline characteristics, with one exception: more children who remained in the study and had teacher-report data at 1st grade spoke a language other than English at home (21.5% of those in the study vs. 13.9% of those who dropped out; Roa-Scott = 4.19, $p = .04$). Children were evenly distributed across gender and race/ethnicity (53% female; 35% Black; 32% Hispanic, 33% White/other). At baseline, children's mean age was 3.21 years ($SE = 0.02$). Additional demographics for the sample at baseline are reported in the measures section.

Measures

Baseline characteristics for defining latent class moderator—All data related to baseline characteristics were collected via parent interview in the child's first year of preschool.

Home composition—Respondents were asked whether the biological mother and the biological father lived in the study child's household. Responses were dichotomized: 1=one or neither biological parent lives in the household (50.6%) and 2=both biological parents live in the household (49.4%).

English language learner (ELL)—At the baseline assessment (Fall 2002), the main caregiver (i.e., teacher or care provider if child was in child care; parent if the child was not in child care) was asked (1) what language the child spoke most often at home, (2) what language the child spoke most often in this care setting, and (3) what language the child preferred to speak. The child was tested in the language that corresponded to their response on two or more of these questions. For the purpose of the present study, we used the language children were tested in at baseline to represent whether that child was an English language learner (ELL) and coded it as 1=Not ELL (80.3%) and 2=ELL (19.7%). The vast majority of ELL children were tested in Spanish, but about 1% (N = 54) were tested in Creole or Mandarin.

Home food stamp status—Respondents were asked, "Do you or anyone in your household get food stamps?" Responses were coded where 1=no (54.2%) and 2=yes (45.8%).

Mother's marital status—Mothers were asked to indicate their marital status. Responses were grouped into the following three categories for the present analyses: 1=never married (41.6%), 2=married (45.5%), and 3=separated or divorced or widowed (12.9%).

Mother's immigrant status—Mothers were asked, "How many years have you lived in the United States?" Responses were dichotomized: 1=recent immigrant (foreign born and lived in US <10 years, 15.4%) and 2=not a recent immigrant (not foreign born or foreign born and lived in the US >10 years, 84.6%).

Maternal depression—Mothers completed a shortened, 12-item version of the Center for Epidemiologic Studies-Depression (CES-D) scale (Seligman, 1993). Items covered symptoms related to depressive thoughts, mood, behavior, and the impact depression has in other areas of the individual's life. Responses were used to categorize mothers into four subgroups: (1) no depressive symptoms (score of 0–4), (2) mild depressive symptoms (score of 5–9), (3) moderate depressive symptoms (score of 10–14), and (4) severe depressive symptoms (score of 15–36). For the present analyses, we created a dichotomous[2] indicator such that 1=no depressive symptoms (score of 0–4, 54.3%) and 2=any depressive symptoms (score of 5–36, 45.7%).

Teen motherhood was calculated based on dates of birth. If the mother was under age 20 when the study child was born, she was considered a teen mother (coded 2, 14.8%), otherwise she was not (coded 1, 85.2%).

Maternal education—Mothers were asked their highest level of education attained. Responses were collapsed into three categories: 1=less than high school diploma (33.1%), 2=high school diploma or GED (34.6%), and 3=beyond high school (32.3%).

Maternal employment status—Mothers were asked to report their employment status. Responses were collapsed into three categories: 1=full-time (35 hours or more per week), in military, 2=part-time, in school or job training, keeping house (i.e., stay-at-home mother), and 3=looking for work, laid off from work, in jail or prison, something else. For ease of interpretation for the remainder of the paper we will refer to these groups as (1) full-time paid work (33.4%), (2) not full-time, paid work (56.1%), and (3) unemployed (10.5%).

In addition to the above measures, pre-academic skills were measured using the Woodcock-Johnson III Brief Achievement standard score from baseline when children were approximately age 3 (M = 100, SD =15). This composite measure assessed children's pre-reading and letter and word identification skills, early math skills, and early writing and spelling skills.

[2]We also ran the 5-class latent class model using a 3-level version of maternal depression (1=no symptoms, 2=mild/moderate symptoms, 3=severe symptoms). The class structure was highly similar to the model which included the 2-level version of maternal depression.

Cognitive Outcome Measures

For cognitive skills, the present study focused on standardized measures from the Peabody Picture Vocabulary Test, third edition (PPVT-III) and two Woodcock-Johnson III Tests of Achievement, which were available at the data collection time points of interest, end of the second year of HS (preschool), kindergarten, and 1st grade.

PPVT-III (Dunn & Dunn, 1997)—The PPVT-III is an established measure of children's receptive vocabulary. For each item, children were shown four pictures and asked to identify the picture of the word spoken by the interviewer. An adapted version of the PPVT-III was used for the HS Impact Study (see USDHHS-ACF, 2010). The PPVT-III has good validity, with published reliability of .95 for the English version. Scores were standardized based on national norms (M = 100, SD = 15). Reliability for the age 3 cohort was α=.61–.78 across time points.

Woodcock-Johnson III Tests of Achievement: letter-word identification (Woodcock, McGrew, & Mather, 2001)—This subtest measured children's ability to match a rebus with the photograph of that object and then measured their ability to identify letters and words in large type in the test book. Scores were standardized based on national norms (M = 100, SD = 15). Reliability for the age 3 cohort was α=.82–.94 across time points.

Woodcock-Johnson III Tests of Achievement: applied problems (Woodcock et al., 2001)—This measured children's ability to analyze and solve practical math problems: children were asked to recognize and understand the mathematical procedure needed to solve a problem and then count and/or preform that procedure. Scores were standardized based on national norms (M = 100, SD = 15). Reliability for the age 3 cohort was α=.88–.89 across time points.

Social-Emotional Competence & Behavioral Outcomes

For social-emotional competence and behavior, we focused on three teacher-rated and three parent-rated measures that were available at all time points of interest.

Parent-rated social skills & positive approaches to learning—This was based on the scale used in the HS Family and Child Experiences Survey (FACES) and was designed to assess parents' perceptions of how their children approach learning and their children's ability to get along with and cooperate with others. Parents rated how true each of seven items was for their child on a scale from 0=*not true* to 2=*very true*. Social skill items included cooperative and empathetic behaviors like "makes friends easily" and "comforts or helps others." Approaches to learning items asked about curiosity, imagination, openness to new tasks, and attitudes toward learning new things, such as "enjoys learning." Scores were summed across items with possible scores ranging from 0 to 14. Reliability for the age 3 cohort was α=.62–.85 across time points.

Parent-rated total behavior problems—This measure is based on the Total Behavior Problem scale used in FACES and assessed parents' perceptions of their child's overall

problem behavior including aggressive, hyperactive and withdrawn behavior. Parents rated 14 items on how true each statement was for their child on a scale from 0=*not true* to 2=*very true*. Example items included "hits or fights with others" and "is very restless and fidgets a lot." We focused on the total problem behavior scale score, which was the sum of the scores across all item, with possible scores ranging from 0 to 28. Reliability for the age 3 cohort was α=.74–.96 across time points. Higher scores indicate more total behavior problems.

Parent-rated parent-child relationship—This measure assessed parents' perceptions of closeness and conflict in their relationship with their child (Pianta, 1992). Parents rated the degree to which each of 15 items represented their relationship with their child on a scale from 1=*definitely does not apply* to 5=*definitely applies*. Example items included, "If upset, this child will seek comfort from me," and "This child easily becomes angry at me." The present study used the total positive relationship scale, which was the sum of all 15 items; possible total scores ranged from 15 to 75 (items were coded so that higher scores indicate a more positive relationship). Reliability for the age 3 cohort was α=.65–.70 across time points.

Teacher-rated problem behaviors—The Adjustment Scales for Pre-School Intervention (ASPI; Lutz, Fantuzzo, & McDermott, 2002) were used to assess teachers' perceptions of students' behavioral adjustment in the classroom. This measure included 24 classroom situations and 144 descriptions of behavior. Teachers selected all descriptions that matched the specified child's behavior for the past two months in their classroom. For example, teachers were asked, "How does this child seek your help?" and selected the appropriate behavioral descriptors: *too lethargic to ask, asks for help when needed, seeks help when not needed, rarely needs help, not shy but never seeks help, or too timid to ask.* Raw scores were computed by summing all of the checked behavior descriptions that load on each factor (shy-socially reticent, aggressive, withdrawn-low energy, oppositional, and inattentive-hyperactive). Raw scores were converted to *t*-scores based on the ASPI standardization sample. The present study relied on *t*-scores of the aggressive and inattentive-hyperactive scales. Reliability was α=.61–.87 for the aggressive scale and α=.58–.84 for the inattentive-hyperactive scale across time points. Note that *N*=106 (12%) of control children were not in center-based care at the "end of preschool" time point and therefore do not have teacher ratings at this time point.

Transforming non-normal outcomes—Before conducting outcome analyses, we examined the distribution of all outcome variables at each time point for normality. The cognitive outcome measures showed relatively normal distributions and therefore were not transformed. All social-emotional competence and behavior measures were substantially skewed. For the parent-rated total problem behaviors and parent- and teacher-rated positive relationships, transformations significantly improved normality. Total problem behaviors (positively skewed) was transformed by taking the square root for the spring of preschool time and 1st grade time points and by taking the natural log for the spring of kindergarten. Parent- and teacher-rated positive relationships (negatively skewed) were transformed by squaring the outcome at all time points. For all transformed variables, we created z-scores

based on the weighted means and standard deviations of the control group at the corresponding time point.

Transformations did not sufficiently improve the distributions for parent-rated social skills and teacher-rated aggression and inattention-hyperactivity outcomes, so we dichotomized these outcomes using cut-points derived from the weighted distributions of control children at the corresponding time point. Based on the shape of these distributions, children were divided into proficiency groups that represented the top or bottom 25% vs. top or bottom 75% of the distribution. For the negative outcomes (e.g., aggression), variables were coded such that 1=poor performance (top 25%) and 0=proficient performance (bottom 75%). For positive outcomes (e.g., social skills), variables were coded such that 1=proficient performance (top 75%) and 0=poor performance (bottom 25%).

Analytic Strategy

Analysis was conducted in three steps. First, we used latent class analysis (LCA) to define the latent moderator reflecting subgroups of children with particular patterns across the nine baseline characteristics. Second, we classified the children into their most likely baseline class based on the posterior probabilities retained from the LCA model at baseline. Third, we estimated the differential impact of HS across these latent subgroups for a variety of cognitive and behavioral skills at the end of the second year of HS (preschool), kindergarten, and 1st grade.

Step 1: Use LCA to define the latent moderator—LCA is a measurement model for latent categorical variables. It posits two or more unobservable subgroups that can be indicated by multiple observed categorical indicators. All nine baseline characteristics used to indicate the latent subgroup were binary, as described in the Measures section. In LCA, the user must select the optimal number of latent classes, which in this case characterize subgroups of children in HS with common profiles on the nine baseline characteristics. We fit a series of LCA models with 1 through 7 classes to explore the number and structure of latent subgroups at baseline in the age 3 cohort. To validate the structure of the selected latent class model, we examined latent class models in the age 4 cohort (i.e., to confirm that the subgroups identified in the age 3 could be replicated in the age 4 cohort).

All subsequent analyses were then conducted with the age 3 cohort. All models accounted for the clustering of children within HS center or care-setting and baseline survey weights. Model identification was assessed using 100 sets of starting values. For model selection, we relied on information criteria (e.g., Bayesian information criterion, BIC; Akaike information criterion, AIC), which balance the tradeoff between model fit and parsimony, as well as conceptual clarity (Collins & Lanza, 2010). All latent class models were fit using SAS Proc LCA (Lanza, Dziak, Huang, Wagner, & Collins, 2013).

Step 2: Classify children according to the LCA model—In order to assign individuals to latent classes for subsequent analyses, we used an improved inclusive classify-analyze strategy that reduces or eliminates any attenuation in the association between latent class membership and a distal outcome (see Bray, Lanza, & Tan, in press). In this approach, the outcome variable is included as a covariate in the latent class model to

derive individuals' posterior probabilities of class membership. Class assignment is made based on these posterior probabilities, and then used in the outcome analysis. This approach preserves the association between class membership and the outcome, which is modeled in Step 3. We classified children into the latent class corresponding to their maximum posterior probability, then created four dummy-coded variables representing membership in one of the five latent classes at baseline; this is the moderator of treatment effects.

Step 3: Examine differential effects of HS across the latent moderator—We then included the moderator (i.e., the dummy-coded class membership variable), a 2-level categorical indicator of whether the child was randomly assigned to HS or the control group, and the interaction between the two as predictors in a regression or logistic model predicting the continuous or binary outcomes, respectively. Outcomes were measured at the end of the second year of HS (preschool), kindergarten, and 1st grade and included: child-assessed receptive vocabulary, reading skills, and math skills (all continuous); parent-rated social skills and positive approaches to learning (binary), total behavior problems (continuous), and parent-child relationships (continuous); and teacher-rated aggression (binary), inattention-hyperactivity (binary), and teacher-child relationships (continuous). This outcome model specification reflects an intent-to-treat analysis with moderation. Because the latent class moderator reflects baseline/pre-randomization characteristics, the moderation effect represented by the set of interaction terms between latent class membership and HS indicate whether the effect of HS, in terms of an ITT analysis, varies across latent classes (Kraemer, Wilson, Fairburn, & Agras, 2002). To adjust for multiple comparisons across outcomes and time points in the tests for differential treatment effects, p-values were adjusted using the Benjamini-Hochberg method (Benjamini, & Hochberg, 1995). Differential treatment effects were only interpreted if significant based on this adjusted p-value. All outcome analyses were conducted with SAS 9.3 Survey Procedures. For each outcome analysis, we used the appropriate weights for the given time point and outcome and accounted for cluster of children within HS center or care-setting. We used jackknife estimation with replicates to calculate standard errors, given the complex sampling design (see USDHHS-ACF, 2010 for more details).

Results

Step 1: Using LCA to Define the Latent Moderator

Model selection—Models with 1 through 7 classes were compared to choose a baseline latent moderator. All models were well-identified. Lower values of information criteria suggest better balance between fit and parsimony. Both the BIC (1775.3 for 5-class vs. 1843.2 for 4-class and 1790.0 for the 6-class) and the certain AIC (1839.3 for 5-class vs. 1894.2 for 4-class, 1867.0 for 6-class) suggested that the 5-class model fit the data best, but the AIC (1290.4 for 7-class vs. 1343.2 for 6-class, 1403.9 for 5-class) and adjusted BIC (1526.7 for 7-class vs. 1545.4 for 6-class, 1572.0 for 5-class) suggested that the 7-class model was superior. Given this information, we carefully reviewed the 5–7 class models for conceptual interpretability and clarity. For parsimony and because the additional classes in the 6 and 7-class models did not contribute to interpretation of the sample[3], we chose the 5-class model described below (see Table 1). To validate the structure of the selected 5-class

model, we freely estimated a 5-class model using data from the 4-year old cohort (see Table 2). All defining characteristics of the latent classes were replicated (e.g., in both samples latent class 1 was characterized by a non-teen mother who is married and speaks English at home, does not receive food stamps, and works part-time) and thus supported the validity of our final model in the 3-year-old cohort. For the purpose of the present study, we conducted the remaining analyses using the age 3 cohort.

Interpreting the latent moderator—Two sets of parameters are estimated in a latent class model: (1) the prevalence of each latent class, and (2) item-response probabilities for each response category, given membership in a particular class. Class prevalences indicate the relative size of each latent class (i.e., distribution of the latent moderator). Because the latent classes are mutually exclusive and exhaustive, the proportion of individuals in all classes sums to 1. Item-response probabilities indicate, for each latent class, the proportion of individuals endorsing a particular response on each item. These probabilities are used to label the subgroups or latent classes; probabilities closer to 1 indicate characteristics of the children in a particular latent class (bolded numbers in Table 1). Table 1 presents latent class prevalences and item-response probabilities for the 5-class model.

We identified two subgroups of children that lived with their married parents and three that lived with a single parent. The largest subgroup (37%) was characterized by living with one or neither of their biological parents, speaking English at home, and having a mother who was not a recent immigrant and who was over age 20 when the child was born. Children in this subgroup were much more likely to receive food stamps and more likely to have a mother who was experiencing depression symptoms. Thus, this subgroup was labeled Single, Food Stamps, Depression. The second largest subgroup of children with a single mother (13%) was very similar to the previous subgroup except that their mothers were likely to have more than a high school diploma and to be employed full-time. This subgroup was labeled Single, Higher Education, Full-Time. Finally, the smallest subgroup (4%) was likely to have a single mother who was divorced or separated or widowed, was a recent immigrant, and had less than a high school diploma. Children in this subgroup were also more likely to speak a language other than English at home and therefore were English language learners (ELL). This subgroup was labeled Single, ELL, Low Education. Overall, the LCA revealed that within these classes, combinations of language, food stamps, immigrant status, depression, education, and employment levels varied substantially, resulting in five distinct and diverse subgroups. Interestingly, there was one baseline characteristic, teen mother, that did not distinguish any of the subgroups, although the rate was highest among children in the Single, Food Stamps, Depressed subgroup (24.5%).

The larger subgroup with married parents (29% of the overall sample) had a high probability of living in a home with both biological parents where English was spoken and food stamps

[3]In the 6-class model, all classes from the 5-class model were replicated. The largest class (labeled Single, Food Stamps, Depression) split into two very similar classes characterized by unmarried mothers who received food stamps and were likely to report depression symptoms. The primary difference between these two classes was that one was more likely to have never been married and the other was about equally likely to have never been married as they were to be divorced/separated/widowed. In the 7-class model, all classes from the 5 and 6-class model were replicated. A new married class emerged, which was similar to the Married Lower Risk class from the 5-class model, but was characterized by a mother with less than a high school diploma.

were not received. Children in this group were also more likely to have mothers who were over 20 when the child was born; were not experiencing depressive symptoms; and were working part-time, in school or job training, or stay-at-home moms. Relative to the rest of the sample, children in this subgroup were exposed to low risk at baseline and therefore were labeled Married, Lower Risk. The second subgroup of children with married parents (17% of the sample) were very similar to the first subgroup, but were very likely to be ELL and to have mothers who were recent immigrants with lower education (no high school diploma). This subgroup was labeled Married, ELL, Low Education.

To further describe the children in these subgroups we examined gender, race, and baseline pre-academic skills. Table 3 shows the distributions of these characteristics and their distributions within each subgroup. Each subgroup was characterized by about equal numbers of males and females, and the overall chi-square showed no significant differences in the proportions across subgroups (Rao-Scott $\chi^2 = 1.84$, $p = .76$). Race (Rao-Scott $\chi^2 = 17.12$, $p = .03$) and pre-academic skills ($F = 11.71$, $p < .0001$), on the other hand, were more different across subgroups. In the two subgroups that were more likely to speak a language other than English at home, all children were Hispanic. For the Married, Lower Risk group, about half of the children were White and the remaining were about evenly split between Black and Hispanic. In contrast, about half of the children in the Single, Food Stamp, Depression and the Single, Higher Education, Full-Time subgroups were Black with the remaining split between White and Hispanic. For pre-academic skills, children in the Married, Lower Risk and Single, Higher Education, Full-Time subgroups performed only about 1/3 of a standard deviation below the national norm of $M = 100$, $SD = 15$, whereas all other subgroups performed nearly 1 or more standard deviations below the national norm for pre-academic skills at baseline.

Step 2: Classify Children according to the LCA Model

The final model used to classify individuals into latent classes (described above) was characterized by high average posterior probabilities for all latent classes and time points, suggesting low classification error. Specifically, average posterior probabilities (across all outcomes and time points) were .92 for Married, Lower Risk; .96 for Married, ELL, Low Education; .86 for Single, Food Stamps, Depression; .75 for Single, Higher Education, Full-time; and .85 for Single, ELL, Low Education. This five-level variable was then used to represent the latent moderator in the differential treatment analyses reported next.

Step 3: Differential Impact of HS Across Latent Subgroups

Table 4 shows F-statistics and p-values for the main effects for HS and latent class as well as the interaction between HS and latent class (i.e., the differential treatment effect) on children's outcomes at spring of preschool, kindergarten and first grade. The impact of HS varied significantly across the latent classes for children's vocabulary and reading scores at the end of preschool and for math at the end of kindergarten and 1st grade. Also, for teacher-rated outcomes, the HS effect varied across classes for aggression, inattentive/-hyperactive behavior, and teacher-child relationships at the end of preschool. For parent-rated behaviors, the HS effect varied across classes for children's total problem behaviors at all three time

points and for child-parent positive relationships and social skills and positive approaches to learning at the end of kindergarten and 1st grade.

For outcomes with a significant differential treatment effect (i.e. significant interaction term), we examined all pairwise comparisons among the five subgroups; the right-most column in Table 4 indicates pairs with significantly different effects of HS. Finally, for those outcomes with a significant differential treatment effect, we examined within-class means for the control and treatment groups. These means are described below and are presented in Tables 5–7.

Positive effects of HS on cognitive outcomes—Table 5 shows the estimated means and standard errors on all cognitive outcomes for the HS and control children by latent class (i.e., baseline subgroup) for outcomes assessed at spring of preschool (top panel), spring of kindergarten (middle panel), and spring of 1st grade (bottom panel). The strongest, most consistent effect of HS was for reading at the end of preschool (the end of two years of HS for those children assigned to the treatment group), when children in HS significantly out-scored children in the control group for three of the five subgroups: Married, ELL, Low Education; Single, Higher Education, Full-Time; and Single, ELL, Low Education. For these subgroups, children assigned to HS scored about ½ to 1 standard deviation above children in the control group on reading skills (Married, ELL, Low Education: M = 90.63 HS, M = 86.46 control, $p < .05$; $d = .17$; Single, Higher Education, Full-Time: M = 98.78 HS, M = 94.62 control, $p < .01$; $d = .28$; Single, ELL, Low Education*:* M = 90.64 HS, M = 78.64 control, $p < .05$; $d = .16$). For vocabulary scores at the end of preschool, only children in the Married, Lower Risk subgroup showed a significant treatment effect (M = 98.94 HS, M = 97.10 control, $p < .01$; $d = .13$). These effects, however, were not evident at 1st grade for any subgroups. Interestingly, the differential impact of HS on math skills emerges in kindergarten and persists into 1st grade. Specifically, for children in the Married, ELL, Low Education subgroup, there is a significant treatment effect on both kindergarten and 1st grade math skills—on average, HS children out-performed control children by about ½ standard deviation and scored near national norms on math (kindergarten: M = 97.31 HS, M = 91.31 control, $p < .05$; $d = .18$; 1st grade: M = 100.86 HS, M = 93.72 control, $p < .05$; $d = .28$).

Positive effects of HS on social-emotional competence & behavioral outcomes—Table 6 shows teacher-rated behaviors of HS and control children, by baseline subgroups, by time point, as estimated proportions for dichotomized outcomes and as means and standard errors for continuous outcomes. For all three outcomes, Single, Higher Education, Full-Time was the only subgroup showing a significant benefit from HS—fewer children in the HS group than in the control group were rated as aggressive (25.4% vs. 59.5%; OR = .23) or inattentive-hyperactive (22.7% vs. 47.1%; OR = .33). Also, HS children in this subgroup had a more positive relationship with their teacher at the end of HS (M = .22 HS, M = −.10 control, $p < .01$; $d = .35$). As with the cognitive outcomes, these differential effects were no longer evident in kindergarten or 1st grade.

Table 7 shows parent-rated behaviors and skills of HS and control children, by baseline subgroups, for spring of preschool (top panel), spring of kindergarten (middle panel), and spring of 1st grade (bottom panel). Estimated proportions are shown for dichotomized

outcomes; estimated means and standard errors are shown for continuous outcomes. For children in the Single, Food Stamps, Depression subgroup, HS had a significant effect on social skills and positive approaches to learning at the end of kindergarten (77.0% vs. 73.1%; OR =1.28) and a positive trend at the end of 1st grade. Children in the Single, Food Stamps, Depression subgroup who attend HS also had significantly fewer behavior problems at the end of kindergarten (M = .01 HS, M = .15 control, $p < .001$; d = .21) and 1st grade (M = .09 HS, M = .22 control, $p < .001$; d = .20). For children in the Single, Food Stamps, Depression and Single, Higher Education, Full-Time subgroups, HS had a significant effect on parent-child relationships at end of kindergarten (for Single, Higher Education, Full-Time; M = .29 HS, M = .12 control, $p < .05$; d = .26) and end of 1st grade (for both groups; M = −.10 HS, M = −.26 control, $p < .01$; d = .21; M = .25 HS, M = .14 control, $p < .05$; d = .22, respectively). Finally, HS children in the Married, ELL, Low Education subgroup had fewer total behavior problems at the end of HS (M = .03 HS, M = .20 control, $p < .01$; d = .23) and kindergarten (M = .03 HS, M = .14 control, $p < .05$; d = .20).

Mixed findings on the effects of HS—As shown above, for outcomes with significant latent moderation, the majority of pairwise comparisons were positive, with children in the HS group outperforming those in the control group. However, the analyses also revealed a few iatrogenic effects. For example, HS children in the Single, Food Stamps, Depression subgroup had significantly more behavior problems (aggressive: 36.9% vs. 16.5%; OR = 2.96; inattentive-hyperactive: 40.6% vs. 27.0%; OR = 1.85) and fewer positive relationships (M = −.18 HS, M = .01 control, $p < .01$, d = .28) at the end of preschool according to teachers. This is in contrast to the parent-rated behavior problems and social skills outcomes, which showed a positive effect of HS for this subgroup at the end of kindergarten and 1st grade. Also, for the Single, Higher Education, Full-Time subgroup, HS had a consistent positive effect on teacher-rated outcomes. It had a negative impact on parent-rated behavior problems at the end of preschool (M = −.14 HS, M = −.61 control, $p < .01$, d = .41) and social skills at the end of kindergarten (74.1% HS vs. 87.2% control, OR = .42), and a positive effect on parent-child relationships at the end of kindergarten and 1st grade.

Discussion

Overall, these findings suggest that the effectiveness of HS for children entering at age 3 varies quite substantially. Specifically, this study found differential treatment effects for six of nine outcomes at the end of preschool, four of nine at the end of Kindergarten, and four of nine at the end of 1st grade. For some children HS appears to have a significant, and in some cases, long-term, positive impact with effect sizes in the small to medium range, similar to previous research on the impact of HS (USDHHS-ACF, 2010). For other children, HS had little to no effect, and for still others, HS appears to have a mix of positive and negative effects. Similar to other HS impact studies (e.g., Lee, 2011; USDHHS-ACF, 2010), the present study found that the most consistent, positive effect of HS was on short-term cognitive outcomes (specifically reading skills). This was true for four of the five subgroups, where children in HS outperformed their non-HS peers in vocabulary or reading skills at the end of HS or preschool; however, like in other studies (e.g., Haskins, 1989; Lee et al., 1990; USDHHS-ACF, 2010), these advantages were not evident at 1st grade. For behavioral

measures, there was less consistency: three subgroups had more positive behavioral and relationship outcomes than non-HS peers.

We found mixed support of our hypotheses. When looking within subgroup and across outcomes, there appears to be a fairly consistent, positive effect for the children in the Married, ELL, Low Education subgroup, with benefits lasting into 1st grade. This finding is in line with our hypothesis that ELL children of immigrants (particularly those with low education) would benefit most from HS. Also, our hypothesis regarding children of single mothers showing more positive benefits is supported in two of the three subgroups with this characteristic. Children in the Single, Food Stamps, Depression subgroup benefitted from lasting effects of HS, but their benefit was mostly in terms of parent-rated behavioral and relationship outcomes. The Single, Higher Education, Full-Time subgroup also benefitted in behavioral and relationship outcomes; compared to control children, HS children in this group were rated higher by teachers immediately after HS and rated higher by parents at the end of kindergarten and 1st grade. Finally, as predicted, for the lowest risk subgroup, Married, Lower Risk, HS had limited effects, with a significant positive effect only on vocabulary skills at the end of preschool.

Potential Explanations for Differential Treatment Effects of HS

There are several explanations for the pattern of differential treatment effects, some of which align with previous differential treatment analyses in early childhood intervention or education more generally. First, past studies indicate that children at lower risk in terms of individual behavior/skills or family/home context benefit less from early intervention than higher-risk children (Bierman et al., 2010, 2008; CPPRG, 2011). We found this to be true in the present analyses: children in the lowest risk subgroup were mostly unaffected by HS. Second, differential treatment studies with HS children show that children with lower academic skills and children of mothers with lower education benefit more from longer exposure to HS, especially with regards to their academic scores (Lee, 2008, 2011). Also, a growing literature suggests that Spanish-speaking, English language learners and children of immigrants benefit significantly from preschool programs, particularly in academic and school readiness skills (Farver et al., 2009; Gormley, 2008; Magnuson et al., 2006). Specifically, Magnuson and colleagues (2006) examined the links between HS attendance and outcomes for children of immigrants. They found that the positive association between attending HS and English proficiency scores was strongest for children of immigrants whose mothers had less than a high school diploma. The combination of low academic skills at baseline, speaking a language other than English at home, and having a mother who is a recent immigrant with a low education is clearly reflected in the Married, ELL, Low Education subgroup. Children in this subgroup scored about 1 standard deviation below national norms in pre-academic skills at baseline and were much more likely to have mothers who were recent immigrants with no high school diploma. This combination of risk factors may help explain why they benefited more from HS than other subgroups: HS likely helped to compensate for the cognitive stimulation and practice with English that was lacking in their home environment.

It is less clear why other subgroups, like the Single, Higher Education, Full-Time subgroup, also experienced greater benefits from HS. Unlike the Married, ELL, Low Education subgroup, children from this subgroup had mothers with high levels of education and scored near national norms on pre-academic skills at baseline. It is interesting that for this group, the benefit of HS was primarily in behavioral and social skills. According to analyses by Zhai et al., (2011), HS impact was evident in non-cognitive domains when compared to children who attended pre-kindergarten programs (on social competence) and center-based care (on social competence, attention and behavior problems). Therefore, one possible explanation may be that control children in the Single, Higher Education, Full-time group were more likely to be exposed to center-based care arrangements. Future research should reveal a clearer picture of the variety of care arrangements experienced by the non-HS children.

Despite these moderated positive treatment effects of HS, it is concerning that a number of outcomes were unaffected or negatively affected by HS for certain subgroups of children. Specifically, for the Single, Food Stamp, Depression and the Single, Higher Education, Full-Time subgroups, findings were mixed. HS had a positive impact on some outcomes and a negative impact on others, within the same subgroup. When comparing the balance of positive to negative outcomes for these subgroups, it appears that children in the Single, Higher Education, Full-Time subgroup fared better. One possible explanation for this is that although this group experienced some risk (i.e., living with a single mother), the level of risk was not so significant (i.e., mothers had a higher education and a full-time job) as to interfere with the benefits of HS. On the other hand, one could argue that children in the Single, Food Stamp, Depression subgroup were exposed to a substantially higher level of risk (i.e., living with a single, depressed mother with few financial resources) and thus the risks experienced at home overwhelmed and in some cases reversed the positive effects of being in HS.

Parent involvement is a cornerstone of the HS program and is hypothesized to be an important mechanism through which HS has its effects (Zigler & Muenchow, 1994). However, research shows that parent involvement varies across and within HS sites. For those higher risk parents who are experiencing tremendous amounts of stress may encounter barriers that make the type of involvement expected from HS quite challenging (Lamb-Parker, Piotrokowski, Baker, Kessler-Sklar, Clark, & Peay, 2001). A recent study by Miller and colleagues, who also uses data from the HS Impact Study, examined if academic effects of HS were moderated by parental preacademic stimulation. They found some evidence to suggest that children receiving moderate levels of stimulation (in comparison to those with the highest or lowest levels) at home benefited most from HS in terms of early reading skills (Miller, Farkas, Vandell, & Duncan, 2014). They referred to this as the “Goldilocks” pattern suggesting that parents with neither too much risk nor too little would benefit from HS. They go on to conclude that children who do not receive sufficient support at home to complement what they are learning at HS are unlikely to reap the full benefits of the program, which may help explain our mixed findings for the Single, Food Stamp, Depression subgroup. Future research is needed to explain the different processes by which high-risk children and their families respond to early childhood interventions like HS.

The relatively limited effects of HS found in the present study may in part be due to the intent-to-treat analyses. Some of the children assigned to receive HS did not in fact receive the full dosage of the program; further, some of the children not assigned to receive HS found their way into other HS programs not involved in the study (or other similar programs). To the extent that this occurred, the effect of HS may be attenuated, making it more difficult to uncover differential treatment effects. It is also possible that, compared to the comprehensive early childhood education program implemented by HS, a more narrowly focused curriculum could produce stronger impacts on certain developmental competencies. For instance, meta-analyses of early education programs have found that the strongest cognitive impacts were found for programs with direct instruction and individualized teaching and no comprehensive social services (Camilli et al., 2010). Another explanation for the lack of effect past kindergarten is that children who are eligible for HS tend to enter low-quality, low-performing elementary schools, which may negate any earlier benefits of HS. In a study of the moderating effect of school quality following a preschool intervention, Zhai, Raver, & Jones, (2012) found that children who received the preschool intervention performed significantly better than comparison children at kindergarten, but only if they entered a high-performing school following the intervention.

Contributions of a Latent Moderator Approach to Differential Treatment Analyses

As opposed to the more traditional examination of differential treatment effects where each individual characteristic is examined as a moderator, the present study used a latent class framework to empirically identify subgroups of children defined by their comprehensive profiles across nine characteristics of the child's primary caregiver and home environment. The HS model was created based on Bronfenbrenner's ecological systems theory, which emphasizes the importance of the influence of multiple interacting systems on children's development (Bronfenbrenner & Morris, 1998). Using a latent moderation approach, we incorporated this theoretical view into our approach to uncover how multiple maternal and home environment characteristics may jointly explain which types of children benefit most from HS. This approach also captured a more nuanced picture of low-income families by illustrating how risk was manifested differently across five subgroups with various combinations of maternal and household characteristics. Although this is a relatively new approach to modeling risk, our previous research with low-income samples has uncovered a similar range of lower and higher risk families that vary according to marital status, economic resources, maternal education, and maternal depression (Lanza, et al., 2011; Lanza, et al., 2010; Rhoades, Greenberg, Lanza, & Blair, 2011).

We confirmed and extended findings of the national evaluators of the HS Impact Study. For example, USDHHS-ACF (2010) found that children who spoke Spanish at home benefited more from HS than children who spoke English. We identified two subgroups of children who were English language learners (the vast majority of whom spoke Spanish). Both had mothers who were recent immigrants and had no high school diploma, but these mothers' marital status varied. The subgroup of children whose mothers were married to and lived with their biological father (Married, ELL, Low Education) experienced more positive effects in reading and math. In contrast, for children with single mothers (Single, ELL, Low Education) HS had less consistent positive effects. It is possible that children exposed to this

combination of low education and a single mother were also exposed to higher levels of stress compared to children of married mothers. Again, in line with the Miller et al. "Goldilocks" pattern (Miller et al., 2014), it may be that the benefits typically associated with attending HS were not sufficient to overcome the stressors of having a non-English-speaking, low education, single mother. USDHHS-ACF (2010) also found that, on average, children whose parents had no depressive symptoms benefited most from HS and that those whose parents had mild or moderate depressive symptoms showed consistent, unfavorable impacts. In our analyses, we found that all but one of the subgroups had very low levels of maternal depression and that the one subgroup with higher levels of maternal depression, Single, Food Stamps, Depression, had somewhat mixed outcomes, with consistent favorable impacts on parent-rated behavioral and social skills and some negative impacts on of teacher-rated behavioral outcomes at the end of preschool.

The latent moderation approach also uncovered variability within subgroups with the same marital status, which appeared to be systematically associated with differential treatment effects of HS. Although mothers in the three single subgroups all reported being unmarried and not being teen mothers, the combination of the remaining characteristics was unique for each subgroup. Similarly, although the two married subgroups shared some characteristics, such as not receiving food stamps and showing few symptoms of depression, they varied in education level, recent immigrant status, and whether their child was ELL – and ultimately in their response to HS. For example, the children in the Married, ELL, Low Education subgroup benefitted significantly more from HS than the Married, Lower Risk subgroup suggesting that marital status alone does not predict differential treatment effects of HS. Rather it the combination of characteristics that better inform for whom HS is most beneficial.

With a latent moderation approach, we extended the one-dimensional nature of traditional moderation analyses (e.g., Aiken & West, 1991) to uncover a more complex story about the impact of HS for children who enter at age 3. Unlike previous moderation approaches that made general conclusions about high-risk children benefitting most, we identified specific profiles, allowing us to examine how specific combinations of risks moderate the impact of HS. If these findings are replicated with other samples, this could inform policies related to early education services for low-income children.

Strengths, Limitations & Future Directions

There are several important strengths to the present study. First, the HS Impact Study is the first large-scale, nationally representative RCT of HS. Given the strengths of this study design, we can be confident in making causal conclusions about the effectiveness of HS. Second, to our knowledge, other than the initial findings in USDHHS-ACF (2010), this is the only independent study to use these data to examine moderation in the impact of HS. Third, using a novel latent moderation approach, we were able to present new, generalizable findings about the differential effects of HS; our findings complement and extend the initial analyses presented by USDHHS-ACF (2010). Ultimately, by using this innovative methodological approach, we were able to paint a more comprehensive picture of the children who benefit most from HS.

Despite these contributions, it is important to recognize several limitations of this study. First, caution should be taken when interpreting differences in teacher-rated outcomes at the end of preschool, as not all children were in center-based care and therefore only a select subset of children received teacher ratings. Second, our latent moderator was defined by nine characteristics of the caregiver and home environment which were found to be important moderators of effects (USDHHS-ACF, 2010). However, there are other important potential moderators of treatment effects to examine in future analyses. These include but are not limited to characteristics of each HS program, neighborhood characteristics, and home environment. We did examine how sensitive the latent moderator (i.e., the structure of the latent classes) was to changes in how several binary indicators (e.g., maternal depression, English language learner status) and outcomes (e.g., teacher-rated aggression and inattention) were coded and found that the number and structure of subgroups and their association with the outcomes remained consistent. We also examined the structure of the latent moderator within the 4-year-old cohort and the results were highly consistent, providing additional validation for this model. An important extension of this research is to determine whether the differential treatment effects found with the 3-year-old cohort are similar for the children who enter HS at age 4. Finally, we examined differential effects of HS for nine outcomes at three time points. This resulted in a large number of statistical tests, increasing the likelihood of chance findings, although this is less likely than if we had conducted traditional moderation analyses (Aiken & West, 1991). The present study represents the first attempt to use a latent moderation approach to examine different effects of HS and therefore is exploratory, so we wanted to comprehensively present our findings. Future analyses, however, should use this approach to examine differential effects on the developmental trajectories of these outcomes, rather than the repeated measures, cross-sectional approach used here. Examining the outcomes using a longitudinal framework would enable a more direct assessment of whether HS effects become significantly smaller across time and whether this fade-out varies by subgroup.

Ultimately, our findings suggest that there is no simple answer to the question, “Does HS work?”. With a latent moderation approach, however, the present study takes a step toward determining for whom HS is most beneficial. Findings must be replicated with other samples to determine their reliability and generalizability. If replicated, they could have great policy and practice significance, particularly for decisions related to which low-income children should be given priority for early education services.

Acknowledgments

This research was supported in part by Grant P50-DA010075-17 from the National Institute on Drug Abuse (NIDA). The content is the sole responsibility of the authors and does not necessarily represent the official views of NIDA or the National Institutes of Health. The authors would like to also thank Dr. Cybele Raver who reviewed and gave feedback on an earlier version of this manuscript.

References

Aiken, LS.; West, SG. Multiple Regression: Testing and Interpreting Interactions. SAGE; 1991.

Alexander KL, Entwisle DR. Achievement in the first 2 years of school: Patterns and processes. Monographs of the Society for Research in Child Development. 1988; 53:1–157. [PubMed: 3226426]

Aughinbaugh A. Does Head Start yield long-term benefits? The Journal of Human Resources. 2001; 36:641–665.

Barnett WS. Effectiveness of early educational intervention. Science. 2011; 333:975–978. [PubMed: 21852490]

Belfield CR, Nores M, Barnett, Schweinhart WSL. The Perry pre-school 40-year follow-up cost-benefit analysis. Journal of Human Resources. 2006; 42:215–246.

Belsky J, MacKinnon C. Transition to school: Developmental trajectories and school experiences. Early Education and Development. 1994; 5:106–119.

Benjamini Y, Hochberg Y. Controlling the false discovery rate: a practical and powerful approach to multiple testing. Journal of the Royal Statistical Society. Series B (Methodological). 1995:289–300.

Bierman KL, Coie JD, Dodge KA, Greenberg MT, Lochman JE, McMahon RJ, Pinderhughes E. The effects of a multiyear universal social–emotional learning program: The role of student and school characteristics. Journal of Consulting and Clinical Psychology. 2010; 78:156–168. [PubMed: 20350027]

Bierman KL, Nix RL, Greenberg MT, Blair C, Domitrovich CE. Executive functions and school readiness intervention: Impact, moderation, and mediation in the Head Start REDI program. Development and Psychopathology. 2008; 20:821–843. [PubMed: 18606033]

Blau, DM. The child care problem: An economic analysis. New York, NY: Russell Sage Foundation; 2001.

Bloom HS, Michalopoulos C. When is the story in the subgroups? Prevention Science. 2012:1–10. [PubMed: 21932067]

Bray BC, Lanza ST, Tan X. Eliminating bias in classify-analyze approaches for latent class analysis. Structural Equation Modeling: A Multidisciplinary Journal. (in press).

Bronfenbrenner, U.; Morris, PA. The ecology of developmental processes. In: Damon, W.; Lerner, RM., editors. Handbook of child psychology. 5. Vol. 1. Hoboken, NJ: Wiley & Sons, Inc; 1998. p. 993-1028.Theoretical models of human development

Brooks-Gunn J, Duncan GJ. The effects of poverty on children. The Future of Children. 1997; 7:55–71. [PubMed: 9299837]

Burger K. How does early childhood care and education affect cognitive development? An international review of the effects of early interventions for children from different social backgrounds. Early Childhood Research Quarterly. 2010; 25:140–165.

Camilli G, Vargas S, Ryan S, Barnett WS. Meta-analysis of the effects of early education interventions on cognitive and social development. Teachers College Record. 2010; 112:579–620.

Campbell FA, Pungello EP, Miller-Johnson S, Burchinal M, Ramey CT. The development of cognitive and academic abilities: Growth curves from an early childhood educational experiment. Developmental Psychology. 2001; 37:231–42. [PubMed: 11269391]

Campbell SB, March CL, Pierce EW, Ewing LJ, Szumowski EK. Hard-to-manage preschool boys: Family context and the stability of externalizing behavior. Journal of Abnormal Child Psychology. 1991; 19:301–18. [PubMed: 1865047]

Collins, LM.; Lanza, ST. Latent class and latent transition analysis with applications in the social behavioral, and health sciences. Hoboken, NJ: Wiley; 2010.

Conduct Problems Prevention Research Group (CPPRG). The effects of the Fast Track preventive intervention on the development of conduct disorder across childhood. Child Development. 2011; 82:331–345. [PubMed: 21291445]

Currie J. Early childhood education programs. Journal of Economic Perspectives. 2001; 15:213–38.

Currie J, Thomas D. School quality and the longer-term effects of Head Start. Journal of Human Resources. 2000; 35:755–774.

Deming D. Early childhood intervention and life-cycle skill development: Evidence from Head Start. American Economic Journal: Applied Economics. 2009; 1:111–134.

Dunn, L.; Dunn, L. Peabody Picture Vocabulary Test-III (PPVT-III). Circle Pines, MN: American Guidance Service; 1997.

Farver JAM, Lonigan CJ, Eppe S. Effective early literacy skill development for young Spanish-speaking English language learners: An experimental study of two methods. Child Development. 2009; 80:703–719.10.1111/j.1467-8624.2009.01292.x [PubMed: 19489898]

Flay BR, Biglan A, Boruch RF, Castro FG, Gottfredson D, Kellam S, Ji P. Standards of Evidence: Criteria for efficacy, effectiveness and dissemination. Prevention Science. 2005; 6:151–175.10.1007/s11121-005-5553-y [PubMed: 16365954]

Garces E, Thomas D, Currie J. Longer term effects of Head Start. American Economic Review. 2002; 4:999–1012.

Gormley WT. The effects of Oklahoma's pre-K program on Hispanic children. Social Science Quarterly. 2008; 89:916–936.10.1111/j.1540-6237.2008.00591.x

Gormley WT, Phillips D, Adelstein S, Shaw C. Head Start's comparative advantage: Myth or reality? Policy Studies Journal. 2010; 38:397–418.

Haskins R. Beyond metaphor: The efficacy of early childhood education. American Psychologist. 1989; 44:274–82.

Kraemer HC, Wilson GT, Fairburn CG, Agras WS. Mediators and moderators of treatment effects in randomized clinical trials. Archives of General Psychiatry. 2002; 59:877–883. [PubMed: 12365874]

Lamb-Parker F, Piotrkowski CS, Baker AJL, Kessler-Sklar S, Clark B, Peay L. Understanding barriers to parent involvement in Head Start: a research-community partnership. Early Childhood Research Quarterly. 2001; 16:35–51.10.1016/S0885-2006(01)00084-9

Lanza, ST.; Dziak, JJ.; Huang, L.; Wagner, AT.; Collins, LM. Proc LCA & Proc LTA users' guide (Version 1.3.0). University Park: The Methodology Center, Penn State; 2013.

Lanza ST, Rhoades BL. Latent class analysis: An alternative perspective on subgroup analysis in prevention and treatment. Prevention Science. 2013; 14:157–168. [PubMed: 21318625]

Lanza ST, Rhoades BL, Nix RL, Greenberg MT. the Conduct Problems Prevention Research Group. Modeling the interplay of multilevel risk factors for future academic and behavior problems: A person-centered approach. Development and Psychopathology. 2010; 22:313–335.10.1017/S0954579410000088 [PubMed: 20423544]

Lanza ST, Rhoades BL, Greenberg MT, Cox M. Modeling multiple risks during infancy to predict quality of the caregiving environment: Contributions of a person-centered approach. Infant Behavior and Development. 2011; 34:390–406.10.1016/j.infbeh.2011.02.002 [PubMed: 21477866]

Lee K. The effects of children's Head Start enrollment age on their short-and long-term developmental outcomes. Social Service Review. 2008; 82:663–702.

Lee K. Maternal coping skills as a moderator between depression and stressful life events: Effects on children's behavioral problems in an intervention program. Journal of Child and Family Studies. 2003; 12:425–437.

Lee K. Impacts of the duration of Head Start enrollment on children's academic outcomes: Moderation effects of family risk factors and earlier outcomes. Journal of Community Psychology. 2011; 39:698–716.

Lee VE, Brooks-Gunn J, Schnur E, Liaw FR. Are Head Start effects sustained? A longitudinal follow-up comparison of disadvantaged children attending Head Start, no preschool, and other preschool programs. Child Development. 1990; 61:495–507. [PubMed: 2344785]

Ludwig J, Miller DL. Does Head Start improve children's life chances? Evidence from a regression discontinuity design. The Quarterly Journal of Economics. 2007; 122:159–208.

Lutz MN, Fantuzzo J, McDermott P. Multidimensional assessment of emotional and behavioral adjustment problems of low-income preschool children: Development and initial validation. Early Childhood Research Quarterly. 2002; 17:338–355.

Macartney, S. Child poverty in the United States 2009 and 2010: Selected race groups and Hispanic origin. American Community Survey Briefs. 2011 Nov. U.S. Census Bureau Report No. ACSBR/10-05. http://www.census.gov/prod/2011pubs/acsbr10-05.pdf

Magnuson K, Lahaie C, Waldfogel J. Preschool and school readiness of children of immigrants. Social Science Quarterly. 2006; 87:1241–1262.10.1111/j.1540-6237.2006.00426.x

McClelland MM, Acock AC, Morrison FJ. The impact of kindergarten learning-related skills on academic trajectories at the end of elementary school. Early Childhood Research Quarterly. 2006; 21:471–490.

McLoyd VC. Socioeconomic disadvantage and child development. American Psychologist. 1998; 53:185–204. [PubMed: 9491747]

McLoyd VC, Kaplan R, Purtell KM, Huston AC. Assessing the effects of a work-based antipoverty program for parents on youth's future orientation and employment experiences. Child Development. 2011; 82:113–132. [PubMed: 21291432]

Miller EB, Farkas G, Vandell D, Duncan GJ. Do the effects of Head Start vary by parental preacademic stimulation? Child Development. 2014 Advanced online publication. 10.1111/cdev.122233

National Institute of Child Health and Human Development Early Child Care Research Network. Duration and developmental timing of poverty and children's cognitive and social development from birth through third grade. Child Development. 2005; 76:795–810. [PubMed: 16026497]

Patterson GR, Forgatch MS, Yoerger KL, Stoolmiller M. Variables that initiate and maintain an early-onset trajectory for juvenile offending. Development and Psychopathology. 1998; 10:531–547. [PubMed: 9741681]

Pianta. New Directions in Child Development. Vol. 57. San Francisco: Jossey-Bass; 1992. Beyond the parent: The role of other adults in children's lives.

Ramey CT, Ramey SL. Early learning and school readiness: Can early intervention make a difference? Merrill-Palmer Quarterly. 2004; 50:471–491.

Raver C, Blair C, Willoughby M. Poverty as a predictor of 4-year-olds' executive function: New perspectives on models of differential susceptibility. Developmental Psychology. 2013; 49:292–304. [PubMed: 22563675]

Reynolds AJ, Temple JA. Extended early childhood intervention and school achievement: Age thirteen findings from the Chicago Longitudinal Study. Child Development. 1998; 69:231–246. [PubMed: 9499569]

Rhoades BL, Greenberg MT, Lanza ST, Blair C. Demographic and familial predictors of early executive function development: Contribution of a person-centered perspective. Journal of Experimental Child Psychology. 2011; 108:638–662.10.1016/j.jecp.2010.08.004 [PubMed: 20828709]

Richardson DK, Shah BL, Frantz ID, Bednarek F, Rubin LP, McCormick MC. Perinatal risk and severity of illness in newborns at 6 neonatal intensive care units. American Journal of Public Health. 1999; 89:511–516. [PubMed: 10191793]

Ryan, RM.; Fauth, RC.; Brooks-Gunn, J. Childhood poverty: Implications for school readiness and early childhood education. In: Spodek, B.; Saracho, ON., editors. Handbook of research on the education of young children. 2. Mahwah, NJ: Lawrence Erlbaum Associates Publishers; 2006. p. 323-346.

Smith, JR.; Brooks-Gunn, J.; Klebanov, PK. Consequences of living in poverty for young children's cognitive and verbal ability and early school achievement. In: Duncan, GJ.; Brooks-Gunn, J., editors. Consequences of growing up poor. New York, NY: Russell Sage Foundation; 1997. p. 132-189.

Supplee LH, Kelly BC, MacKinnon DM, Barofsky MY. Introduction to the special issue: Subgroup analysis in prevention and intervention research. Prevention Science. 2013; 14:107–110. [PubMed: 23090721]

U.S. Department of Health and Human Services, Administration for Children and Families (USDHHS-ACF). Final report. Washington, DC: 2010. Head Start impact study.

Wang R, Ware JH. Detecting moderator effects using subgroup analyses. Prevention Science. 2013; 14:111–120. [PubMed: 21562742]

Woodcock, RW.; McGrew, KS.; Mather, N. Technical manual: Woodcock-Johnson III. Itasca, IL: Riverside Publishing; 2001.

Yoshikawa H, Aber JL, Beardslee WR. The effects of poverty on the mental, emotional, and behavioral health of children and youth: Implications for prevention. American Psychologist. 2012; 67:272. [PubMed: 22583341]

Zhai F, Brooks-Gunn J, Waldfogel J. Head Start and urban children's school readiness: A birth cohort study in 18 cities. Developmental Psychology. 2011; 47:134. [PubMed: 21244155]

Zhai F, Raver CC, Jones SM. Academic performance of subsequent schools and impacts of early interventions: Evidence from a randomized controlled trial in Head Start settings. Children and Youth Services Review. 2012; 34:946–954. [PubMed: 22773872]

Zigler, E.; Muenchow, S., editors. Head Start: The Inside Story of America's Most Successful Educational Experiment. New York, NY: Basic Books; 1994.

Zigler E, Styfco SJ. Head Start: Criticisms in a constructive context. American Psychologist. 1994; 49:127–32.

Table 1

3-Year-Old Cohort: Five Latent Classes Comprising Latent Moderator at Baseline

Baseline Characteristic	Overall Proportion	Latent Baseline Subgroups (Proportions)				
		Married, Lower Risk (29%)	**Married, Span, Low Ed (17%)**	**Single, Food Stp, Dep (37%)**	**Single, Higher Ed, FT (13%)**	**Single, Span, Low Ed (4%)**
Living with Both Bio. Parents						
No	0.494	0.050	0.056	**0.865**	**0.904**	**0.979**
Yes	0.506	**0.950**	**0.945**	0.135	0.096	0.021
English Language Learner						
No	0.803	**0.988**	0.077	**0.960**	**0.999**	0.239
Yes	0.197	0.012	**0.923**	0.040	0.001	**0.761**
Receiving Food Stamps						
No	0.542	**0.656**	**0.801**	0.232	**0.781**	0.525
Yes	0.458	0.344	0.199	**0.768**	0.219	0.475
Mom Marital Status						
Single (Never Married)	0.416	0.069	0.207	**0.728**	**0.539**	0.402
Married	0.455	**0.930**	**0.772**	0.073	0.154	0.004
Divorced/Separated/Widow	0.129	0.001	0.021	0.199	0.307	**0.594**
Mom Recent Immigrant						
No	0.846	**0.922**	0.403	**0.992**	**0.990**	0.291
Yes	0.154	0.078	**0.597**	0.008	0.010	**0.709**
Mom Depression						
No symptoms	0.543	**0.651**	**0.668**	0.412	0.510	**0.594**
Some symptoms	0.457	0.349	0.332	**0.588**	0.490	0.406
Teen Mom						
No	0.852	**0.873**	**0.932**	**0.755**	**0.926**	**0.952**
Yes	0.148	0.127	0.068	0.245	0.074	0.048
Mom Education						
No HS Diploma	0.331	0.200	**0.616**	0.385	0.073	**0.535**
HS Diploma	0.346	0.409	0.232	0.377	0.299	0.138
Beyond HS Diploma	0.323	0.391	0.152	0.238	**0.628**	0.328
Mom Employment						

Baseline Characteristic	Overall Proportion	Latent Baseline Subgroups (Proportions)				
		Married, Lower Risk (29%)	Married, Span, Low Ed (17%)	Single, Food Stp, Dep (37%)	Single, Higher Ed, FT (13%)	Single, Span, Low Ed (4%)
Full-Time Paid Work	0.334	0.320	0.163	0.256	**0.901**	0.287
Not Full-Time Paid Work	0.561	**0.626**	**0.816**	**0.527**	0.082	**0.615**
Unemployed	0.105	0.054	0.022	0.217	0.017	0.098

Note. Bolded item response probabilities indicate a high probability of children in that class having that characteristic.

Table 2

4-Year-Old Cohort (Validation Sample): Five Latent Classes Comprising Latent Moderator at Baseline

Baseline Characteristic	Overall Proportion	Latent Baseline Subgroups (Proportions)				
		Married, Lower Risk (26%)	Married, Span, Low Ed (27%)	Single, Food Stp, Dep (33%)	Single, Higher Ed, FT (7%)	Single, Span, Low Ed (7%)
Living with Both Bio. Parents						
No	0.449	0.136	0.044	**0.867**	**0.991**	**0.737**
Yes	0.551	**0.864**	**0.956**	0.133	0.009	0.263
English Language Learner						
No	0.628	**0.996**	0.027	**0.994**	**0.535**	0.009
Yes	0.372	0.004	**0.973**	0.006	0.465	**0.991**
Receiving Food Stamps						
No	0.599	**0.691**	**0.795**	0.338	**0.985**	0.233
Yes	0.401	0.309	0.205	**0.663**	0.015	**0.767**
Mom Marital Status						
Single (Never Married)	0.331	0.092	0.135	**0.574**	**0.519**	**0.592**
Married	0.491	**0.904**	**0.855**	0.044	0.007	0.141
Divorced/Separated/Widow	0.178	0.004	0.010	0.382	0.474	0.267
Mom Recent Immigrant						
No	0.732	**0.899**	0.327	**0.993**	**0.658**	0.417
Yes	0.268	0.101	**0.673**	0.007	0.342	**0.583**
Mom Depression						
No symptoms	0.494	0.493	**0.691**	0.360	0.460	0.407
Some symptoms	0.506	0.507	0.309	**0.640**	**0.540**	**0.593**
Teen Mom						
No	0.854	**0.863**	**0.917**	**0.749**	**0.920**	**0.837**
Yes	0.146	0.137	0.083	0.251	0.080	0.163
Mom Education						
No HS Diploma	0.423	0.263	**0.683**	0.315	0.194	**0.849**
HS Diploma	0.325	0.348	0.203	0.396	**0.550**	0.024
Beyond HS Diploma	0.252	0.389	0.114	0.289	0.256	0.127
Mom Employment						

Baseline Characteristic	Overall Proportion	Latent Baseline Subgroups (Proportions)				
		Married, Lower Risk (26%)	Married, Span, Low Ed (27%)	Single, Food Stp, Dep (33%)	Single, Higher Ed, FT (7%)	Single, Span, Low Ed (7%)
Full-Time Paid Work	0.333	0.265	0.183	0.434	**0.878**	0.106
Not Full-Time Paid Work	0.589	**0.698**	**0.789**	0.381	0.116	**0.839**
Unemployed	0.078	0.037	0.028	0.185	0.006	0.055

Note. Bolded item response probabilities indicate a high probability of children in that class having that characteristic.

Table 3

3-Year-Old Cohort: Gender and Race Proportions and Pre-Academic Skills Means and Standard Deviations by Latent Moderator

Characteristic	Overall Sample	Latent Moderator				
		Married, Lower Risk (29%)	Married, Span, Low Ed (17%)	Single, Food Stp, Dep (37%)	Single, Higher Ed, FT (13%)	Single, Span, Low Ed (4%)
Gender						
Male	0.50	49.9	41.1	44.7	51.1	60.0
Female	0.50	50.1	58.9	55.3	48.9	40.0
Race						
White	0.31	56.7	2.0	29.1	35.0	3.4
Black	0.36	22.1	0.5	56.1	52.3	3.3
Hispanic	0.33	21.2	97.4	14.7	12.7	93.3
Pre-Academic Skills (Standard Score Mean)	90.24 (.46)	94.72 (.94)	84.35 (1.36)	88.68 (0.64)	93.80 (1.43)	82.35 (2.50)

Note. Race proportions and mean pre-academic skills are significantly different across classes ($p < .05$ for both).

Table 4

Test Statistics for Head Start, Latent Class, and Head Start by Latent Class Predicting Child Outcomes

Outcome	Spring Time-point	Predictor					
		Head Start (HS)		Latent Class (LC)		HS X LC	
		F/Chi-Sq	p-value	F/Chi-Sq	p-value	F/Chi-Sq	BH adjusted p-value[1]
Cognitive Skills							
Vocabulary	Head Start (Preschool)	3.87	.05	104.86	<.0001	3.71	0.03
	Kindergarten	0.36	ns	82.62	<.0001	1.11	ns
	1st Grade	0.09	ns	56.09	<.0001	0.99	ns
Reading	Head Start (Preschool)	13.02	<.001	10.44	<.0001	4.52	0.02
	Kindergarten	0.05	ns	13.49	<.0001	1.02	ns
	1st Grade	0.48	ns	13.63	<.0001	1.56	ns
Math	Head Start (Preschool)	3.48	ns	29.29	<.0001	1.22	ns
	Kindergarten	2.09	ns	22.39	<.0001	4.24	0.02
	1st Grade	0.08	ns	15.92	<.0001	3.09	0.04
Teacher-Rated Behaviors							
Aggressive	Head Start (Preschool)	0.09	ns	31.94	<.0001	36.58	<.001
	Kindergarten	0.01	ns	10.07	.04	8.01	ns
	1st Grade	0.31	ns	38.93	<.0001	7.76	ns
Inattentive/Hyperactive	Head Start (Preschool)	0.19	ns	9.29	.05	22.04	<.01
	Kindergarten	0.03	ns	0.03	ns	7.61	ns
	1st Grade	0.03	ns	5.59	ns	7.13	ns
Teacher-Child Positive Relationship	Head Start (Preschool)	2.84	ns	5.50	<.001	5.54	0.01
	Kindergarten	6.12	.02	6.86	<.001	2.25	ns
	1st Grade	0.13	ns	26.23	<.0001	0.47	ns
Parent-Rated Behaviors & Skills							
Total Problem Behavior	Head Start (Preschool)	0.01	ns	23.56	<.0001	4.20	0.02
	Kindergarten	0.00	ns	9.10	<.0001	3.44	0.03
	1st Grade	0.08	ns	16.12	<.0001	4.77	0.02
Social Skills & Positive Approaches to Learning	Head Start (Preschool)	2.37	ns	20.17	<.001	9.70	ns

Outcome	Spring Time-point	Predictor					
		Head Start (HS)		Latent Class (LC)		HS X LC	
		F/Chi-Sq	p-value	F/Chi-Sq	p-value	F/Chi-Sq	BH adjusted p-value[1]
Child-Parent Positive Relationship	Kindergarten	3.81	ns	15.21	<.01	12.63	0.03
	1st Grade	0.14	ns	11.80	.02	14.21	0.02
	Head Start (Preschool)	0.17	ns	14.08	<.0001	0.20	ns
	Kindergarten	0.07	ns	25.09	<.0001	3.48	0.03
	1st Grade	1.84	ns	16.69	<.001	3.92	0.02

Note. Head Start df = 1, 51; Latent Class df = 4, 51; Head Start X Latent Class df = 4, 51.

[1] Benjamini-Hochberg adjustment (Benjamini & Hochberg, 1995) applied to control for multiple comparisons.

Table 5

3-Year-Old Cohort: Estimated Means and Standard Errors on Cognitive Outcomes for Head Start and Control Children by Latent Moderator

Outcome	LC Pairwise Treatment Differences[1]	Latent Moderator									
		1. Married, Lower Risk		2. Married, ELL, Low Education		3. Single, Food Stamps, Depression		4. Single, Higher Education, Full-Time Employment		5. Single, ELL, Low Education	
		Control	HS	Control	HS	Control	HS	Control	HS	Control	HS
Vocabulary											
Preschool/HS	1–3, 3–5	97.10 (.87)	98.94** (.76)	82.17 (1.30)	82.93 (.94)	91.14 (.87)	89.87 (.87)	100.61 (1.00)	100.47 (1.13)	76.17 (1.81)	81.57 (1.95)
Kindergarten	---	98.50 (.68)	98.85 (.61)	87.37 (1.03)	88.48 (.73)	92.21 (.59)	92.59 (.68)	101.04 (.86)	101.65 (1.11)	89.04 (1.72)	84.50 (1.79)
1st Grade	---	91.95 (1.45)	92.86 (.92)	82.74 (1.00)	84.99 (.91)	87.99 (.64)	88.11 (.77)	95.14 (.78)	95.58 (.98)	83.15 (1.31)	80.40 (2.10)
Reading											
Preschool/HS	1–5, 2–3, 3–4, 3–5	94.40 (1.60)	95.62 (1.15)	86.46 (1.78)	90.63* (1.89)	92.33 (1.33)	89.94 (1.02)	94.62 (1.11)	98.78** (1.07)	78.64 (3.05)	90.64* (3.96)
Kindergarten	---	102.62 (2.61)	102.36 (1.58)	97.79 (1.70)	102.13 (1.99)	97.62 (1.21)	95.91 (1.17)	105.43 (1.50)	105.83 (1.39)	101.43 (1.79)	99.81 (3.85)
1st Grade	---	106.42 (2.09)	106.13 (1.59)	100.32 (2.34)	106.08 (1.35)	99.36 (.76)	99.06 (1.27)	109.69 (1.67)	105.38 (1.84)	100.10 (3.52)	95.88 (3.87)
Math											
Preschool/HS	---	97.62 (.90)	98.91 (.78)	90.47 (1.85)	90.51 (1.56)	93.03 (.88)	92.11 (.82)	99.59 (1.21)	99.45 (1.37)	70.58 (4.75)	87.66 (5.68)
Kindergarten	1–2, 2–3,2–4, 2–5	102.57 (1.32)	102.15 (1.12)	91.31 (2.55)	97.31* (2.55)	96.38 (.90)	93.82 (1.23)	106.99 (1.17)	103.06 (1.88)	94.40 (2.04)	86.62** (2.93)
1st Grade	2–3, 2–4	102.24 (1.15)	104.42 (1.36)	93.72 (2.29)	100.86* (1.67)	96.88 (1.06)	96.01 (1.21)	105.43 (1.39)	103.08 (2.38)	98.77 (3.54)	94.24 (3.48)
Cognitive Outcome N's											
Preschool/HS		N=254	N=401	N=146	N=199	N=236	N=398	N=125	N=222	N=15	N=44
Kindergarten		N=254	N=367	N=142	N=213	N=214	N=432	N=101	N=130	N=21	N=53
1st Grade		N=234	N=383	N=134	N=200	N=214	N=406	N=109	N=131	N=21	N=44

* $p < .05$,

** $p < .01$,

*** $p < .001$.

HS = Head Start.

[1] For each outcome where HS x LC was significant, we list the LC pairs with significantly different treatment effects ($p < .05$). 1 = Married, Lower Risk, 2 = Married, ELL, Low Education, 3 = Single, Food Stamps, Depression, 4 = Single, Higher Education, Full-Time Employment, 5 = Single, ELL, Low Education. Dashes indicate no significant HS x LC moderation.

Table 6

3-Year-Old Cohort: Estimated Means and Proportions on Teacher-Rated Outcomes for Head Start and Control Children by Latent Moderator

Outcome	LC Pairwise Treatment Differences[1]	Latent Moderator									
		1. Married, Lower Risk		2. Married, ELL, Low Education		3. Single, Food Stamps, Depression		4. Single, Higher Education, Full-Time Employment		5. Single, ELL, Low Education	
		Control	HS	Control	HS	Control	HS	Control	HS	Control	HS
Aggressive											
Preschool/HS	1–3, 1–4, 2–3, 3–4, 4–5	25.16	22.96	25.36	16.08	16.52	36.91***	59.46	25.40***	15.83	26.45
Kindergarten	---	24.22	27.45	22.77	17.16	39.08	44.31	0	29.69	27.78	33.33
1st Grade	---	17.50	19.52	16.10	11.76	40.38	37.78	29.67	13.04	21.74	21.95
Inattentive/Hyperactive											
Preschool/HS	2–3, 3–4	28.77	25.52	27.27	20.07	26.96	40.64**	47.07	22.73**	13.89	13.87
Kindergarten	---	26.92	26.86	24.27	16.76	36.90	31.60	26.53	35.24	11.76	31.03
1st Grade	---	24.05	26.27	26.89	19.65	37.04	39.16	34.44	22.81	30.00	26.32
Teacher-Child Positive Relationship											
Preschool/HS	1–4, 2–3, 3–4	0.01 (0.07)	−0.00 (0.04)	0.08 (0.06)	0.21 (0.05)	0.01 (0.04)	−0.18** (0.04)	−0.1 (0.12)	0.22** (0.06)	−0.16 (0.18)	0.03 (0.19)
Kindergarten	---	−0.00 (0.06)	−0.02 (0.04)	0.11 (0.05)	0.1 (0.05)	−0.1 (0.05)	−0.07 (0.03)	0.22 (0.09)	−0.1 (0.09)	0.18 (0.1)	−0.09 (0.1)
1st Grade	---	0.07 (0.06)	0.17 (0.04)	0.03 (0.05)	0.07 (0.04)	−0.19 (0.04)	−0.17 (0.04)	0.25 (0.05)	0.27 (0.06)	0.01 (0.14)	−0.06 (0.15)
Teacher-rated Outcome N's											
Preschool/HS		N=170	N=257	N=90	N=148	N=157	N=334	N=97	N=181	N=12	N=41
Kindergarten		N=182	N=289	N=97	N=161	N=234	N=384	N=43	N=76	N=16	N=29
1st Grade		N=171	N=269	N=121	N=160	N=176	N=354	N=108	N=122	N=23	N=40

* p < .05,

** p < .01,

*** p < .001.

HS = Head Start.

[1] For each outcome where HS x LC was significant, we list the LC pairs with significantly different treatment effects ($p < .05$). 1 = Married, Lower Risk, 2 = Married, ELL, Low Education, 3 = Single, Food Stamps, Depression, 4 = Single, Higher Education, Full-Time Employment, 5 = Single, ELL, Low Education. Dashes indicate no significant HS x LC moderation.

Table 7

3-Year-Old Cohort: Estimated Means and Proportions on Parent-Rated Outcomes for Head Start and Control Children by Latent Moderator

Outcome	LC Pairwise Treatment Differences[1]	Latent Moderator									
		1. Married, Lower Risk		2. Married, ELL, Low Education		3. Single, Food Stamps, Depression		4. Single, Higher Education, Full-Time Employment		5. Single, ELL, Low Education	
		Control	HS	Control	HS	Control	HS	Control	HS	Control	HS
Total Problem Behavior											
Preschool/HS	1–4, 2–4, 3–4, 4–5	−0.19 (0.06)	−0.21 (0.06)	0.20 (0.06)	0.03** (0.06)	0.10 (0.04)	0.00 (0.04)	−0.61 (0.11)	−0.14** (0.09)	0.58 (0.13)	0.44 (0.12)
Kindergarten	1–2, 1–3, 3–4	−0.23 (0.07)	−0.15 (0.06)	0.14 (0.04)	0.03* (0.04)	0.15 (0.03)	0.01*** (0.03)	−0.11 (0.05)	−0.03 (0.06)	0.1 (0.07)	0.19 (0.07)
1st Grade	1–2, 1–3, 3–4	0.03 (0.04)	−0.04* (0.04)	−0.14 (0.05)	−0.05 (0.03)	0.22 (0.04)	0.09*** (0.03)	−0.11 (0.05)	−0.06 (0.04)	0.15 (0.12)	0.16 (0.07)
Social Skills & Positive Approaches to Learning											
Preschool/HS	---	71.35	62.24	61.07	69.62	49.03	57.97	80.25	66.80	80.59	57.62
Kindergarten	1–4, 2–4, 3–4	71.92	78.30	81.34	80.68	73.07	77.70*	87.18	74.11**	80.95	81.63
1st Grade	1–5, 2–5, 3–5	62.23	60.06	71.74	70.44	52.61	61.84	67.31	61.07	77.27	65.22*
Parent-Child Positive Relationship											
Preschool/HS		0.08 (0.04)	0.12 (0.03)	−0.02 (0.07)	−0.01 (0.07)	−0.1 (0.04)	−0.04 (0.03)	0.28 (0.04)	0.29 (0.06)	−0.03 (0.14)	−0.06 (0.06)
Kindergarten	1–4	0.16 (0.05)	0.06 (0.04)	−0.00 (0.07)	0.06 (0.06)	−0.24 (0.04)	−0.18 (0.04)	0.12 (0.06)	0.29* (0.05)	0.26 (0.15)	−0.03 (0.12)
1st Grade	1–2, 1–3, 1–4, 2–5, 3–5	−0.02 (0.04)	0.09 (0.05)	0.14 (0.06)	0.08 (0.04)	−0.26 (0.05)	−0.1** (0.03)	0.14 (0.05)	0.25* (0.04)	0.09 (0.07)	0.01 (0.09)
Parent-rated Outcome N's											
Preschool/HS		N=202	N=343	N=153	N=205	N=318	N=496	N=80	N=179	N=15	N=39
Kindergarten		N=237	N=361	N=130	N=200	N=237	N=437	N=134	N=189	N=20	N=46
1st Grade		N=211	N=359	N=132	N=198	N=253	N=450	N=122	N=165	N=22	N=45

* $p < .05$,

** $p < .01$,

*** $p < .001$.

HS = Head Start.

[1] For each outcome where HS x LC was significant, we list the LC pairs with significantly different treatment effects ($p < .05$). 1 = Married, Lower Risk, 2 = Married, ELL, Low Education, 3 = Single, Food Stamps, Depression, 4 = Single, Higher Education, Full-Time Employment, 5 = Single, ELL, Low Education. Dashes indicate no significant HS x LC moderation.