Do the Effects of Head Start Vary by Parental Pre-academic Stimulation?

Elizabeth B. Miller
George Farkas
Deborah Lowe Vandell
Greg J. Duncan

*University of California, Irvine*

Research reported in this publication was supported by the Eunice Kennedy Shriver National Institute of Child Health & Human Development of the National Institutes of Health under Award Number P01HD065704. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health. The authors gratefully acknowledge Anamarie Auger, Marianne Bitler, and Weilin Li, who provided invaluable support with analyses for this article.
Abstract

Data from the Head Start Impact Study (N=3,185) were used to determine whether Head Start differentially benefited children from homes with high, middle, and low levels of parental pre-academic stimulation on three academic outcome domains – early literacy, receptive vocabulary, and early math – after one year. Results from residualized growth models showed positive impacts of random assignment to Head Start on all three outcomes, and positive associations between parental pre-academic stimulation and academic performance. Two moderated effects were also found. Head Start boosted early math skills the most for children receiving low parental pre-academic stimulation. Effects of Head Start on early literacy skills were largest for children receiving parental pre-academic stimulation in the middle ranges. Implications for Head Start are discussed.

Keywords: Head Start, parenting, pre-academic stimulation, differential effects
Do the Effects of Head Start Vary by Parental Pre-academic Stimulation?

Caregiving for young children in the United States is often provided by both parents and child care settings. Close to 40% of children under five years of age who have regular child care arrangements are in center-based child care settings (Laughlin, 2013). Low-income children typically enter school a half to a full standard deviation below higher-income children in academic domains such as vocabulary, cognition, and literacy skills (Duncan & Magnuson, 2013), however, those who experience high-quality early childhood programs have been found to enter kindergarten with academic skills closer to those of high-income children (e.g., Barnett, 2011; Karoly, Kilburn & Cannon, 2005; Ramey & Ramey, 2006; Schweinhart, 2006). In 2012, the federally funded Head Start program served over 825,000, or 20%, of all the children in center-based child care settings (Barnett, Carolan, Fitzgerald, & Squires, 2012).

In the 1998 reauthorization of Head Start, Congress mandated that the U.S. Department of Health and Human Services (DHHS) determine whether the program contributed to key outcomes in children’s learning and development. The resulting Head Start Impact Study (HSIS) gathered data from a large, nationally representative sample of children assigned at random to Head Start centers or a comparison group between 2002 and 2006. The Final Report of the HSIS found statistically significant cognitive gains for Head Start children over control children after one year (U.S. DHHS, Final Report, 2010).

The HSIS report also provided some evidence of differential program effects among key subgroups. On the Peabody Picture Vocabulary Test (PPVT; Dunn & Dunn, 1997), Head Start had a larger impact for Dual Language Learners compared with monolingual English speakers. Head Start also a larger impact for children of mothers with no depressive symptoms compared with children of depressed mothers. On the Woodcock-Johnson III (WJ) Letter-Word and
Applied Problems tests (Woodcock, McGrew, & Mather, 2001), Head Start had a larger impact for children of mothers with mild depressive symptoms compared with either no symptoms or moderate or severe symptoms. Head Start also a larger impact for children living with moderate household risk compared with either no risk or high risk. These findings are important because they provide support for different types of moderated effects in the HSIS.

The purpose of the current study is to extend the examination of the differential effects of Head Start by asking if the quality of parenting prior to Head Start also moderates program effects. In particular, we ask if the effects of Head Start on early academic skills are moderated by the level of pre-academic stimulation in the home.

**Parental Role in Fostering Pre-academic Skills**

Parents play a key role in early childhood by fostering school readiness through exposure to pre-academic stimulation (e.g., Bakermans-Kranenburg, van IJzendoorn, & Bradley, 2005; Leventhal, Martin, & Brooks-Gunn, 2004). Such stimulation includes reading to children, helping them write their name, helping them to recognize and pronounce letters and words, and helping them with math skills such as counting objects. Associations between these types of activities and academic success in early childhood have been well documented (e.g., Bradley & Caldwell, 1995; Melhuish et al., 2008). Studies have shown that low-income children, in general, receive less academic stimulation at home than do children in higher-income families (e.g., Bradley et al., 1989; Klebanov, Brooks-Gunn, McCarton, & McCormick, 1998), although there is wide variability within low-income households. In the current study, we ask if effects of Head Start vary for children who receive different levels of academic stimulation at home prior to the beginning Head Start.
Potential Synergistic Effects of Head Start and Parental Pre-academic Stimulation

**Theoretical Framework.** Bronfenbrenner’s bioecological model of development (1979; Bronfenbrenner & Morris, 2006) posits that multiple interacting environments such as parents, schools, communities, and even state and national policies influence a child’s development. Thus, children’s development is determined not by one single factor, but by the interplay of child, family, and environment across time. Identifying possible differential impacts of Head Start on children from varying home environments therefore can help Head Start improve its capabilities in both serving a disadvantaged population and detecting mechanisms relevant for targeted intervention.

Several competing hypotheses involve these varying effects of stimulation in the home and child care environments. We apply these general hypotheses within a framework more specific to Head Start and parental pre-academic stimulation, the subjects of our study.

The *compensatory* hypothesis posits that Head Start will have the largest impacts on children from the most disadvantaged home environments (Sameroff & Chandler, 1975), who receive very little pre-academic stimulation at home. The compensatory hypothesis is consistent with the stated goals of Head Start (Zigler & Styfco, 2010) to serve the children who are most in need of its services. This hypothesis has received empirical support in the child care research. McCartney, Dearing, Taylor, and Bub (2007), for example, found that high quality early childhood settings had larger effects for children growing up in poverty.

A contrasting *accumulated advantages* hypothesis posits that children who bring more to encounters with people and activities are likely to derive the most benefit (Coleman, 1990). Thus, children from less-risky advantage-laden home environments come to Head Start better prepared to capitalize on the benefits of the program. Child care studies have provided some
empirical support for this hypothesis as well. The Sure Start program in the U.K. was more
effective for children of non-teenage mothers and less effective for children of teenage mothers
or single parents (Belsky, Melhuish, Barnes, Leyland, & Romaniuk, 2006).

A third possibility, which we call the “Goldilocks” hypothesis, posits that Head Start will
be less effective for children at high risk from struggling, chaotic home environments, as well as
children at very low risk owing to their reduced need and the lower quality match between
program services and these children’s needs and skills. Instead, Head Start will be most effective
for children in the middle ranges of risk. Support for this relationship has been found in other
studies of child outcomes in low-income populations, particularly when investigating the match
between welfare-to-work policies and worker needs (e.g., Huston et al., 2003; Yoshikawa,
Magnuson, Bos, & Hsueh, 2003). In these studies, antipoverty employment programs
differentially benefitted children of mothers who were moderately hard-to-employ rather than
children of mothers who were the most or least likely to be employed.

Figure 1 illustrates the differences between the three hypotheses. For the compensatory
hypothesis, Head Start has the largest effect on children’s academic skills when they receive low
levels of parental pre-academic stimulation prior to Head Start. In contrast, the accumulated
advantages hypothesis predicts that Head Start has the largest effect on children’s academic
skills when they receive high levels of parental pre-academic stimulation prior to Head Start. For
the “Goldilocks” hypothesis, Head Start has the largest effect for children whose parental pre-
academic stimulation is in the middle ranges as opposed to either of the ends.

{Insert Figure 1}

**Prior Research.** Two recent studies examined the synergistic potential of child care
programs and home learning environments. Bradley, McKelvey, and Whiteside-Mansell (2011)
analyzed the Early Head Start Research and Evaluation (EHSRE) study and asked if the 14-month Home Observation for Measurement of the Environment (HOME) Learning Stimulation subscale score moderated the effects of EHS. Consistent with the compensatory hypothesis, they found larger effects of Early Head Start on the Bayley Scales of Infant Development at 36 months and the WJ Letter-Word test at 60 months for children from homes low in cognitive stimulation at 14 months. They did not test for the kinds of non-linear relations predicted by the “Goldilocks” hypothesis. Our study investigated these non-linear possibilities, finding them important for the relationship between parental stimulation and Head Start.

Watamura, Phillips, Morrisey, McCartney and Bub (2011) analyzed the NICHD Study of Early Child Care and Youth Development (SECCYD) dataset to examine children across multiple contexts of quality home and child care. They found children from low quality home environments differentially benefitted from exposure to high quality child care, providing support for the compensatory hypothesis. While these results are consistent with prior research, the data are non-experimental and thus unable to eliminate the possibility of selection bias due to omitted variables.

**Testing for interactions.** Two common approaches to testing interactions between baseline moderators and treatment assignment are to treat moderators in continuous and categorical form. If parental pre-academic stimulation is represented as a continuous variable, its interaction with treatment tests whether the relationship between assignment to Head Start and child outcomes varies systematically with each unit increase in parental pre-academic stimulation. However, this assumes that the interaction is linear. One way of addressing non-linear continuous interactions is by adding a squared parental stimulation term and its interaction with treatment to the equation. A second approach in testing for interactions has been the use of
categorical variables. The categorical specification employs dummy variables to represent levels of parental pre-academic stimulation often chosen to correspond to percentiles (e.g., 20th, 25th) of the sample distribution. This provides a flexible test of whether the relationship between assignment to treatment and child outcomes varies systematically based on the level of parental stimulation.

A variation on this use of categorical interactions employs dummy variables based on substantive, rather than distributional, divisions of the sample. In a recent paper examining maternal sensitivity and child care using the NICHD SECCYD, Burchinal, Vandell, & Belsky (in press) tested for interactions with both the continuous and categorical approaches. They found categorical interactions of high maternal sensitivity, based on substantive divisions of 5.5 to 7 on a global 1-7 scale, and child care to be the most predictive of adolescent outcomes. In this paper we make use of both potential interactive model specifications as well.

**Present Study.** The present study extends the investigation of possible synergistic relationships between Head Start and the home learning environment. This study is to our knowledge the first to address this issue using the HSIS, which randomly assigned children to Head Start or a control condition.

Our principal research question is: After one academic year in the HSIS, do the effects of Head Start on three academic domains – early math, early literacy, and receptive vocabulary – vary by the level of parental pre-academic stimulation at baseline? If so, which of the three moderating hypotheses – compensatory, accumulated advantages, or “Goldilocks” – are supported by the data? Further, what is the non-experimental association between parental pre-academic stimulation and children’s achievement? Because evidence of all three hypotheses has been found in studies with low-income samples, we explicitly test all three hypotheses for each
of the outcomes. In addition, we hypothesize that parental pre-academic stimulation will be positively associated with higher pre-academic skills, and we expect positive main effects of Head Start consistent with those reported in the Final Report (2010). Since Head Start, but not pre-academic stimulation, was randomly assigned, only the Head Start impact estimates benefit from the experimental nature of the study.

We consider three different academic domains because the interaction between Head Start and parental pre-academic stimulation may depend on the match between the skills parents impart to their children and those provided by the Head Start program. Indeed, the HSIS Final Report (2010) found differential impacts of the Head Start program among subgroups consistent with all three hypotheses for different outcomes. For receptive vocabulary as measured by the PPVT, there was evidence of a compensatory relationship with a larger Head Start effect for Dual Language Learners compared with native English speakers. There was also evidence of an accumulated advantages relationship with a larger Head Start effect for children of mothers with no depressive symptoms compared with children of mothers with such symptoms.

For early literacy as measured by the WJ Letter-Word test and early math as measured by the WJ Applied Problems test, the HSIS found evidence of a “Goldilocks” pattern of moderation in that Head Start had a larger impact for children of mothers with mild depressive symptoms compared with either no symptoms or moderate or severe symptoms. Head Start also had a larger impact for children living with moderate household risk compared with either no risk or high risk. Other prior literature, however, demonstrates that parents provide relatively little math instruction or support at home (e.g., Blevins-Knabe, Berghout, Musun-Miller, Eddy, & Jones, 2000; Young-Loveridge, 1989), consistent with a compensatory relationship for the domain of early math.
Given the evidence for all three patterns of moderation on the extent to which Head Start complements or substitutes for the home learning environment in the acquisition of later skills, we cannot clearly predict the relationship pattern between pre-academic stimulation and Head Start as all three hypotheses are plausible and empirically supported. The dearth of details on the extent to which Head Start complements or substitutes for parental support at low, medium, and high stimulation levels is the primary motivation for this paper.

Method

Participants

We analyzed data from the HSIS, a nationally representative sample of 84 Head Start grantee/delegate agencies and nearly 5,000 newly entering, eligible three and four-year-old children. Children were randomly assigned to either: (1) a Head Start group that had access to Head Start program services; or (2) a control group that was not eligible to enroll in the Head Start center to which they applied for the lottery, but could enroll in other early childhood programs or services selected by their parents, including other Head Start centers not in the study (U.S. DHHS, Final Report, 2010).

The study employed a multi-stage sampling process to select a representative group of Head Start programs and children. It began with a list of 1,715 grantee/delegate Head Start agencies that were operating in Fiscal Year (FY) 1998-99. This pool was then organized into 161 geographic clusters across 25 strata in order to ensure variation across region of the country, urban/rural location, race/ethnicity, and state pre-kindergarten and child care policies. One cluster was then randomly selected from each of the 25 strata yielding 261 grantee/delegate agencies. Agencies were eliminated that had recently closed, merged, or were serving all eligible children in their communities, and smaller agencies were grouped together. Approximately three
grantee/delegate agencies were then randomly selected from each of the 25 strata, yielding a final 84 grantee/delegate agencies.

These 84 Head Start agencies generated lists of 1,427 individual centers that were expected to be in operation for the 2002-03 school year. After individual programs were eliminated because they had recently closed, merged, or were serving all eligible children in their communities, and groups of centers were stratified along the same dimensions as the geographical agency clusters, 383 individual centers remained (U.S. DHHS, Final Report, 2010). An average of four centers were selected from each agency with a range of 1-7 centers (C. Heid, personal communication, 10 April, 2013).

Once the centers were selected, a lottery process was used to determine which children were and were not assigned a place in Head Start. The goal was to randomly select 27 children from each center – 16 to be assigned to Head Start and 11 to the control condition. In total 4,442 children were randomly selected – 2,646 for Head Start and 1,796 for the control condition. Data collection took place from fall 2002, at the time the treatment group entered Head Start, until spring 2006, at the end of first grade (U.S. DHHS, Final Report, 2010). The resulting sample was roughly split into thirds by child’s race – black, Hispanic, and white or other. Further, about half the sample was male and about a quarter of the sample was classified as a Dual Language Learner at baseline. About a third of the mothers in the sample had less than a high school education, about 15% were teenage mothers, and about 40% were classified as moderately to severely depressed by the CES-D Depression scale (Radloff, 1977).

Measures

**Parental pre-academic stimulation.** In the fall of 2002, the primary caregiver was administered a baseline interview by project staff. The primary caregiver was considered the
person living with and most responsible for the care of the child, and in the HSIS, ¾ of the respondents were the biological or adoptive mother. This parent interview was designed for the U.S. DHHS’s Head Start Children and Families Experiences Survey (FACES) and HSIS studies. It was based loosely on Caldwell & Bradley’s (1984) HOME measure, and specifically the Learning Stimulation and Language Stimulation subscales. The HOME is one of the most widely used measures designed to assess characteristics of a child’s home environment, and we included all such similar questions. Reviews of the HOME have found it to be a reliable and valid measure (Bradley, 1994; Bradley & Caldwell, 1988; Elardo & Bradley, 1981), even in low-income and diverse populations (Bradley, Corwyn, Pipes McAdoo, & García Coll, 2001; Bradley, Corwyn & Whiteside-Mansell, 1996).

During this interview, the primary caregiver was asked to report on the use of ten educational activities in the past week: the number of times they read to their child; helping their child with letters, numbers, and words; practicing writing the alphabet with their child; helping their child with songs or music; working on arts and crafts with their child; helping their child practice writing their name; practicing rhyming words with their child; counting objects with their child; talking about size with their child; and talking about the calendar with their child. Scoring was done exactly as it is in the HOME with a binary choice (yes/no) format. A variable was coded “0” if the parent interviewed indicated they had not done the activity with the child in the past week and “1” if they indicated they had. In the case of the activity “how many times the parent read to their child in the past week” which was reported on a 1-4 scale from 1 “never” to 4 “everyday”, the item was recoded to have a 0-1 scale like the other activities with 0 “never”, .33 “once or twice”, .67 “three or more times”, and 1 “everyday.” The variables were then summed to create a composite scale of parental pre-academic stimulation varying from 0 “has done none
of these activities in the past week” to 10 “has done all of these activities in the past week” ($\alpha = .71$). Factor analysis yielded only one factor from the ten items, confirming the utility of summing these items for one omnibus parental simulation measure. A complete list of descriptive statistics for the individual items as well as the composite parental pre-academic stimulation measure is reported in Table 1. Figure 2 displays a histogram of the composite parental pre-academic stimulation measure.

{Insert Table 1 & Figure 2}

**Academic achievement outcomes.** Both at baseline and after one academic year in the treatment or control condition, children were administered a battery of assessments including the Woodcock Johnson (WJ) III Applied Problems and Letter-Word tests (Woodcock, McGrew, & Mather, 2001), and the Peabody Picture Vocabulary Test, Third Edition (PPVT; Dunn & Dunn, 1997) as measures of academic achievement. The WJ Applied Problems test measures a child’s ability to analyze and solve math problems ($\alpha = .92$). The WJ Letter-Word test measures a child’s reading identification skills of letters and words ($\alpha = .91$). The PPVT measures a child’s receptive vocabulary ($\alpha = .95$). All three assessments are norm-referenced ($M = 100$, $SD = 15$) (U.S. DHHS, Technical Report, 2010). A complete list of descriptive statistics for these outcomes at baseline and one academic year later is reported in Table 1. We chose these specific outcome domains given their respective critical importance for later academic success (Duncan et al., 2007; Whitehurst & Lonigan, 2003; Yesil-Dagli, 2011).

**Covariates.** In order to sharpen the standard errors of our point estimates and to adjust for departures from randomization, several child and family covariates were included in the model to minimize the potential for omitted variable bias. We used the same set of covariates as was used in the Final Report of the HSIS (2010), which included a broad set of key child and
family demographic characteristics. Child covariates included: number of elapsed weeks from September 1, 2002 until the spring 2003 assessment; age in weeks at the spring 2003 assessment; gender; race/ethnicity; whether the child was classified as having a disability; and the language of baseline testing. Family covariates included: caregiver age in years; an indicator of caregiver depression; highest level of maternal education; whether the mother was a recent immigrant to the United States; primary language used at home; and three family structure variables including whether both biological parents lived with the child, whether the mother was married, and whether the mother was a teenage mother at the child’s birth. A lagged dependent variable was also included in the model to control for the child’s baseline level of achievement. For purposes of analyses, all the covariates were centered at their mean. Descriptive statistics for all covariates are displayed in Table 2, which also includes tests for the treatment-control group differences for each of these variables.

Non-response. As with any longitudinal dataset, there was non-response in the HSIS. In particular, spring 2003 child assessment response rates were correlated with treatment/control status as well as child gender and race. To control for this potential bias, we weighted all our analyses using the spring 2003 child assessment final child weight (CHSPR2003WTCA), which included a weight for probability of sample selection at every stage multiplied by a weight adjusted for non-response. The weight included in our analyses helped control potential non-response bias by compensating for different data collection response rates across these demographic groups of children. Weights are important in complicated multi-stage sampling studies such as HSIS because they allow us to make inferences to the relevant general

As a robustness check, we also ran multiple imputation for missing data and found that the results of our analyses did not change significantly. Hence below we report the weighted case results (N=3,185).

Results

Descriptive Analyses

Tables 1 and 2 display the descriptive statistics for the main independent variable of parental pre-academic stimulation as well as the covariates, respectively. Table 1 reports that the mean of parental pre-academic stimulation was 6.61 on a scale of 0 to 10, indicating the distribution is slightly skewed in the negative direction. Table 2 reveals that balance was achieved between treatment and control on the covariates, with no significant differences between the groups, except on two measures. Not surprisingly due to its stated commitment to serving children with disabilities and its purposeful reservation of at least 10% of eligible enrollment slots for these children (U.S. DHHS, Final Report, 2010), Head Start was more likely to serve children with disabilities as opposed to the control condition ($p < .05$). Second, caregivers of children enrolled in Head Start were significantly more likely to be diagnosed as depressed by the CES-D Depression scale than caregivers of control group children ($p < .05$).

Preliminary Analyses

Our investigation of the best model specification for our data was an iterative conceptual and empirical process. It began with both linear and quadratic specifications for the pre-academic stimulation interactions with the Head Start treatment variable. We discovered that a parabolic relation fit the data better than a linear relation, with larger effects at the middle of the sample
distribution than at the tails. We then searched systematically for cut-points in the distribution of pre-academic stimulation that corresponded to meaningful distinctions on the pre-academic stimulation scale and provided the best fit to the data. Neither terciles nor quartiles produced significant interactions. However, when we defined very low levels of pre-academic stimulation as 0-2 activities on our composite measure, which corresponded roughly to the bottom 10% of the HSIS sample, and defined very high levels of pre-academic stimulation as 8-10 activities on our composite measure, which corresponded to roughly the top 10% of the HSIS sample, we found the strongest evidence of interactions. Similar to Burchinal et al.’s (in press) work, we found interactions that were based on substantive divisions situated in our sample to be the strongest predictors of child outcomes.

Residualized growth models were used to estimate the effect of parental pre-academic stimulation on academic achievement above and beyond what can be explained by the lagged dependent variable and what can be predicted based on covariates. Using the survey commands in Stata, we ran step-wise regressions testing for main effects of pre-academic stimulation and Head Start for all three academic outcomes first, and then subsequently included the interaction terms between the variables. Thus, our principal regression model was:

$$Y_t = \beta_0 + \beta_1 Y_{t-1} + \beta_2 (HIGH \ STIMULATION)_{t-1} + \beta_3 (LOW \ STIMULATION)_{t-1} + \beta_4 TX + \beta_5 (HIGH \ STIMULATION_{t-1} \times \ TX) + \beta_6 (LOW \ STIMULATION_{t-1} \times \ TX) + \gamma COVARIATES + \epsilon_t,$$

where $Y_t$ is the outcome variable of interest after one year; $Y_{t-1}$ is the lagged outcome variable at baseline; $HIGH/LOW \ STIMULATION_{t-1}$ are the dummy variables indicating whether a child came from a home high or low in pre-academic stimulation at baseline; $TX$ is the dummy variable for assignment to Head Start; $HIGH/LOW \ STIMULATION_{t-1} \times \ TX$ are the interactions of parental stimulation at baseline and assignment to Head Start; $COVARIATES$ is the vector of child and family covariates described above that were used the model; and $\epsilon_t$ is an error term.
Regression Analyses

Table 3 displays the three sets of regression results. For each outcome, the first model estimates the main effects of high and low pre-academic stimulation (compared with the reference group of middle ranges of stimulation) and the Head Start treatment effect, while the second model adds interactions between the Head Start treatment and each of the stimulation dummy variables. Also included are a full set of covariates, which are listed in the table’s footnote.

{Insert Table 3}

Results for Model 1 (columns 1, 3, and 5) show that children from homes low in pre-academic stimulation scored close to a 1/3 of a SD lower than children receiving middle ranges of stimulation on the WJ Applied Problems ($b = -4.49, p < .001$), over a 1/5 of a SD lower on the WJ Letter Word ($b = -3.29, p < .01$), and about .07 SD lower on the PPVT ($b = -1.06, p < .01$).

In contrast children with high parental pre-academic stimulation did not differ significantly from children receiving middle ranges of pre-academic stimulation, indicating a non-linear pattern for the outcomes as parental pre-academic stimulation increased from low to middle and from middle to high levels.

Head Start impacts in Model 1 are consistent with those reported in the HSIS Final Report, with Head Start children scoring about .17 SD higher on the WJ Applied Problems ($b = 2.46, p < .001$), over 1/5 of a SD higher on the WJ Letter-Word ($b = 3.44, p < .001$), and almost .10 SD higher on the PPVT ($b = 1.40, p < .001$) compared with the control children.

Model 2 (columns 2, 4, and 6) tests for moderation. When the interaction variables are included, the main effects of parental pre-academic stimulation are those of control group children. Consequently, for each outcome we first discuss the results for control group children
and then the Head Start children. Further, to better understand these interactions and to compare them to the idealized patterns in Figure 1, Figures 3-5 show predicted scores computed from the coefficient estimates in Model 2. The Head Start children are indicated by the light gray line, while the control group children are indicated by the dark gray line. To determine the magnitude of the differences between them, we differenced the treatment and control group scores at each stimulation level and standardized them to calculate Head Start treatment effect sizes at each level of pre-academic stimulation. These effect sizes are indicated by “d”. The asterisks indicate whether the treatment-control difference at low or high parental stimulation is statistically different from the treatment-control difference at the middle ranges of stimulation.

{Insert Figures 3-5}

For the WJ Applied Problems (column 2 and Figure 3), control group children from homes with low parental pre-academic stimulation scored over .46 SD lower (b= -6.93, p < .001) compared with children in the middle ranges of parental stimulation. However, there was no parallel increase in score for control children by moving from middle to high parental stimulation, as evidenced by a slope of essentially zero.

We compare the scores for Head Start children from homes with low parental pre-academic stimulation to those for Head Start children receiving middle ranges of stimulation by adding the significant point estimates for the main and interactive effects of low stimulation. Head Start children from homes low in parental pre-academic stimulation had a more muted decrease than controls, scoring about .12 SD lower (b= -1.78) compared with children from homes receiving middle ranges of stimulation. Similar to the control children, there was no parallel increase in score between high and middle stimulation for Head Start children, so the pattern of the movement from middle to high stimulation is essentially the same for the treatment
and control groups. The Head Start effect for the WJ Applied Problems, therefore, is largest for children receiving low pre-academic stimulation at home as indicated in Figure 3, which is consistent with the compensatory effect pattern illustrated in Figure 1.

For the WJ Letter-Word (column 4 and Figure 4), control group children from homes with low parental pre-academic stimulation scored about .15 SD lower (b = -2.20, p < .10) compared with children in the middle ranges of parental stimulation while control group children from homes with high parental pre-academic stimulation scored close to .30 SD higher (b = 4.39, p < .05) compared with children in the middle stimulation ranges. Thus, there is a consistent upward trend in score for the control children as parental pre-academic stimulation increases from low to middle to high levels.

Using the same calculation method as above in adding the significant main and interactive effects, Head Start children in the low stimulation group scored close to .30 SD lower (b = -4.36) compared with the middle stimulation group. Repeating a similar calculation, Head Start children from homes with high parental pre-academic stimulation scored .02 SD lower (b = -0.32) compared with children receiving middle ranges of stimulation. Thus as indicated in Figure 4, for Head Start children the score increase in moving from low to middle stimulation is more dramatic than the score increase in moving from middle to high stimulation which is much more muted and almost flat. This matches the “Goldilocks” effect pattern shown in Figure 1. That is, for the WJ Letter-Word, Head Start has the largest treatment effect for children whose parental stimulation is not too low and not too high, but instead in the middle ranges of “just right” pre-academic stimulation.

For the PPVT (column 6 and Figure 5), as parental stimulation increases from low to middle levels, there is no statistically significant increase in score for control group children.
However, control group children from homes high in parental pre-academic stimulation scored about .14 SD higher \((b=2.16, p < .01)\), compared with children in the middle ranges.

For Head Start children, there is also no statistically significant difference in score between low and middle stimulation. However, Head Start children from homes high in parental pre-academic stimulation scored almost .02 SD lower \((b=-0.31)\) on the PPVT compared with middle stimulation children when we add the point estimates for the main and interactive effect of high stimulation. Thus, as shown in Figure 5, children assigned to Head Start experienced essentially a flat main effect of treatment, and little to no Head Start treatment effect was observed for the children from homes high in parental pre-academic stimulation.

**Robustness Checks**

Because Head Start but not parenting were randomly assigned, we sought to ensure that the stimulation interactions that we found were not picking up the effects of correlated predictors. To do this, we reran all of our models including interactions between all of our covariates and the Head Start treatment indicator. Our primary interaction variables (low stimulation \(\times\) treatment) and (high stimulation \(\times\) treatment) retained their significance in these fully-interacted models. We also interacted our primary interaction variables (low stimulation \(\times\) treatment) and (high stimulation \(\times\) treatment) with all of our covariates (a three-way interaction between treatment, high/low parental pre-academic stimulation, and the covariates one by one), and in no case was this interaction statistically significant. These robustness checks confirmed our findings that parental pre-academic stimulation moderated the Head Start treatment, and that this interaction did not further vary by any of the other covariates including maternal education.

Moreover, factor analysis of our parental pre-academic stimulation measure revealed one omnibus factor, confirming our measure’s utility in discerning the overall level of parental pre-
academic stimulation at baseline. We ran follow up analyses testing for interactions between just the math items and Head Start on the math outcome, and similarly just the literacy and vocabulary items and Head Start on the literacy and vocabulary outcomes. These analyses did not yield any significant interactions further confirming the advantage of summing all the items for one omnibus parental simulation measure.

**Discussion**

This paper used the HSIS data to test for both main effects of parental pre-academic stimulation and Head Start as well as three moderating hypotheses involving the differential effects of Head Start for children from varying home environments on three academic domains – early math (as assessed by the WJ Applied Problems test), early literacy (as assessed by the WJ Letter-Word test), and receptive vocabulary (as assessed by the PPVT).

We found Head Start treatment main effects to be consistent with those reported in the HSIS Final Report. Head Start children scored higher on all three outcomes compared with control group children.

We also found the main effects of baseline parental pre academic stimulation to be non-linear. Children from homes low in pre-academic stimulation scored lower on all three outcomes compared with children receiving middle ranges of stimulation. However, outcomes for children from homes high in parental pre-academic stimulation did not differ significantly from outcomes for children receiving middle ranges of pre-academic stimulation.

Our primary analyses concentrated on testing for three types of moderated effects. We found the differential effects of Head Start to vary by the domain of interest. For the domain of early math, we observed a moderately-sized compensatory effect of Head Start in that children from homes with very low stimulation experienced the largest Head Start treatment effect. In
contrast, the Head Start treatment effect was virtually the same for children from homes with middle and high levels of pre-academic stimulation. This more muted treatment effect pattern is consistent with previous literature on parental numeracy activities with their children (e.g., Blevins-Knabe et al., 2000; Young-Loveridge, 1989), and it appears that the intensive time spent learning early math concepts in Head Start has a particularly strong impact on children who receive little to no parental pre-academic stimulation at home.

For the domain of early literacy, using a relatively restrictive linear specification, Bradley et al. (2011) found the largest Early Head Start effect for children receiving low levels of stimulation at home, consistent with a compensatory pattern. Our more flexible model specification uncovered a non-linear relationship rendering results more consistent with a “Goldilocks” effect in that the group of children receiving the middle ranges of parental stimulation received the largest score boost from the Head Start treatment. This finding for early literacy resonates with the growing belief in social science research that interventions must take into account both the characteristics of the program and the individual needs of the child (Duncan & Vandell, 2012). Similar to previous literature on welfare-to-work policies as well as the HSIS Final Report (2010) subgroup analyses, early childhood programs like Head Start may be least effective in the early literacy domain for the most vulnerable children from chaotic, struggling home environments, as well as for the most advantaged children due to their reduced need and less appropriate match between services provided and their needs and skills. Rather, programs may be most effective for those receiving middle levels of stimulation, in the sweet spot of “just enough” risk and heightened sensitivity.

For the domain of receptive vocabulary, similar to the different patterns of moderation found in the HSIS subgroup analyses, we found evidence of moderation that was not fully
consistent with any of our three hypotheses. A modest sized Head Start main effect was observed for children receiving low and middle ranges of parental pre-academic stimulation, but none was observed for those receiving very high levels of stimulation. Although little to no Head Start effect was observed for children from homes high in parental pre-academic stimulation, we found a particularly strong gain in scores for control children when parental pre-academic stimulation increased from middle to high levels. In fact, as demonstrated by Figures 3-5, the control group children’s achievement increased as stimulation level increased for all outcomes.

Thus, higher parental pre-academic stimulation appears to foster resiliency in Head Start eligible children who do not have the program available to them, particularly in the domains of early literacy and receptive vocabulary. Consistent with Rutter’s (1987) ideas surrounding protective mechanisms, higher parental pre-academic stimulation can be a protective factor for children not assigned to Head Start. Watamura et al. (2011) had similar findings of resiliency in children from high quality home environments, regardless of the quality of the child care environment, on caregiver-reported measures.

The differential effects we observed across outcome domains are largely consistent with the types of relationship patterns reported in prior literature on different outcome domains. Differences among these three domains in the match between the skills parents impart to their children as measured by our stimulation scale and how much the Head Start program complements or substitutes for such parental stimulation may account for our diverging results. Though our measure is a reliable indicator of global baseline parental stimulation at home, closer inspection of the individual items shows there are more early literacy-related items than items for either vocabulary or math. Factor analysis revealed one omnibus factor and follow up analyses did not find specific interactions between just the math items and Head Start on the math
outcome and just the literacy and vocabulary items and Head Start on those outcomes respectively, confirming our measure’s utility in discerning the overall level of parental pre-academic stimulation at baseline. However, this overall measure may not be sufficiently sensitive to capture how the individual stimulation items affect some outcomes more than others.

Furthermore, because the HSIS did not conduct home observations, it is possible that the items relating to parental pre-academic stimulation may not reflect the actual quality of such stimulation. It may be that parents answered the interview questions on the basis of what they perceived to be socially acceptable responses, rather than based on their actual behavior. Yet such responses would have introduced measurement error that biased our coefficient estimates down toward zero. Thus, our findings are likely conservative estimates of the true magnitudes involved.

Finally, we see considerable value in our attempt to create divisions of baseline parental pre-academic stimulation based on substantively meaningful rather than purely distributional criteria. In the Head Start sample, the “very low” (scores of 0-2 on the stimulation scale) and “very high” (scores of 8-10 on the scale) groups corresponded roughly to the bottom and top deciles of the Head Start sample distribution. In a more representative sample, we would expect to see a considerably larger fraction of the sample in the “very high” category and a smaller fraction in the “very low” category. We hypothesize that interactions with early education quality treatments will be stronger with substantively-defined rather than distributionally-defined groups. We look forward to seeing if future work confirms this hypothesis.

In sum, we have found that baseline parental pre-academic stimulation and Head Start have synergistic impacts on children’s development above and beyond what can be predicted solely on the basis of maternal and child characteristics. Our results use a relatively new data set
to increase understanding of differential effects of Head Start on children from varying home environments. In particular, we found support for compensatory and “Goldilocks” patterns of Head Start effects on early achievement outcomes. We found no support for the accumulated advantages hypothesis, suggesting that Head Start should continue targeting children at highest risk as well as those at more moderate risk.

Our findings underscore the interplay between Head Start and baseline parental pre-academic stimulation as they combine to produce early math, early literacy, and vocabulary skills for low-income children. These results are helpful for future interventions surrounding the match between child care programs and parenting, and suggest that it is particularly important that Head Start be made available to those children whose parents are not able to provide high levels of pre-academic stimulation. Despite the current political climate surrounding funding for early education, it is vital that Head Start continue to serve children at highest and moderate risk as, on average, the program aids their development even when parents are unable to provide high levels of stimulation.

Our study also suggests that children’s academic achievement may benefit from programs targeted at helping parents increase their own pre-academic stimulation capabilities. An important goal of Head Start is to increase family functioning and engagement as these factors also directly impact a child’s school readiness (Zigler & Styfco, 2010). Parents may therefore change their stimulation practices over the course of the child’s participation in Head Start. While our study focused on parental stimulation at baseline, a next research step would be to investigate whether parents change their behaviors and respond differentially to Head Start themselves. Future research studies should examine this possibility and the mechanisms that may influence it.
References


Table 1

*Descriptive Statistics - Key Independent Variable and Outcomes*

<table>
<thead>
<tr>
<th></th>
<th>N</th>
<th>Mean / % of Sample</th>
<th>Standard Deviation</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Pre-academic Stimulation - Baseline (Fall 2002)</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Number of times child is read to in a week</td>
<td>3305</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Never</td>
<td></td>
<td>0.07</td>
<td></td>
</tr>
<tr>
<td>Once or twice</td>
<td></td>
<td>0.32</td>
<td></td>
</tr>
<tr>
<td>Three or more times</td>
<td></td>
<td>0.27</td>
<td></td>
</tr>
<tr>
<td>Everyday</td>
<td></td>
<td>0.35</td>
<td></td>
</tr>
<tr>
<td>Helped with letters, words, numbers</td>
<td>3305</td>
<td>0.91</td>
<td></td>
</tr>
<tr>
<td>Practiced writing the alphabet</td>
<td>3300</td>
<td>0.63</td>
<td></td>
</tr>
<tr>
<td>Helped with songs or music</td>
<td>3297</td>
<td>0.84</td>
<td></td>
</tr>
<tr>
<td>Worked on arts and crafts</td>
<td>3288</td>
<td>0.62</td>
<td></td>
</tr>
<tr>
<td>Practiced writing or spelling name</td>
<td>3295</td>
<td>0.71</td>
<td></td>
</tr>
<tr>
<td>Practiced rhyming words</td>
<td>3281</td>
<td>0.38</td>
<td></td>
</tr>
<tr>
<td>Counted things or objects</td>
<td>3298</td>
<td>0.80</td>
<td></td>
</tr>
<tr>
<td>Talked about size</td>
<td>3300</td>
<td>0.49</td>
<td></td>
</tr>
<tr>
<td>Talked about calendar</td>
<td>3304</td>
<td>0.58</td>
<td></td>
</tr>
<tr>
<td>Total of pre-academic stimulation activities</td>
<td>3239</td>
<td>6.61</td>
<td>2.25</td>
</tr>
<tr>
<td><strong>Academic Achievement - Baseline (Fall 2002)</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>WJ Letter Word</td>
<td>3577</td>
<td>91.46</td>
<td>16.82</td>
</tr>
<tr>
<td>WJ Applied Problems</td>
<td>3577</td>
<td>89.82</td>
<td>16.16</td>
</tr>
<tr>
<td>PPVT</td>
<td>3577</td>
<td>91.55</td>
<td>8.78</td>
</tr>
<tr>
<td><strong>Academic Achievement - End of First Year Treatment (Spring 2003)</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>WJ Letter Word</td>
<td>3633</td>
<td>92.75</td>
<td>17.32</td>
</tr>
<tr>
<td>WJ Applied Problems</td>
<td>3633</td>
<td>89.20</td>
<td>17.99</td>
</tr>
<tr>
<td>PPVT</td>
<td>3633</td>
<td>92.28</td>
<td>9.57</td>
</tr>
</tbody>
</table>

*Note.* WJ = Woodcock-Johnson. PPVT = Peabody Picture Vocabulary Test. Weight used = CHSPR2003WTCA.
Table 2

_Descriptive Statistics - Background and Demographic Controls, Full Sample_

<table>
<thead>
<tr>
<th>Child Characteristics - Baseline (Fall 2002)</th>
<th>Treatment</th>
<th></th>
<th></th>
<th>Control</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>N</td>
<td>2292</td>
<td>1341</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mean / % of Sample</td>
<td>236.02</td>
<td>236.13</td>
<td></td>
<td>31.38</td>
<td>27.06</td>
<td></td>
</tr>
<tr>
<td>Standard Deviation</td>
<td>31.38</td>
<td>0.55</td>
<td></td>
<td>27.06</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Age at Spring 2003 assessment in weeks</td>
<td>2292</td>
<td></td>
<td></td>
<td>1341</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Age-3 cohort</td>
<td>0.54</td>
<td>0.55</td>
<td></td>
<td>0.55</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Gender - Male</td>
<td>0.49</td>
<td>0.49</td>
<td></td>
<td>0.49</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Race</td>
<td>2292</td>
<td></td>
<td></td>
<td>1341</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Black</td>
<td>0.32</td>
<td>0.29</td>
<td></td>
<td>0.36</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Hispanic</td>
<td>0.36</td>
<td>0.36</td>
<td></td>
<td>0.36</td>
<td></td>
<td></td>
</tr>
<tr>
<td>White &amp; Other</td>
<td>0.33</td>
<td>0.35</td>
<td></td>
<td>0.35</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Spanish as baseline testing language</td>
<td>2292</td>
<td></td>
<td></td>
<td>1341</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Disability</td>
<td>0.24</td>
<td>0.23</td>
<td></td>
<td>0.11</td>
<td></td>
<td>*</td>
</tr>
<tr>
<td>Disability</td>
<td>2292</td>
<td></td>
<td></td>
<td>1341</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

_Family Characteristics - Baseline (Fall 2002)_

| Caregiver age                                          | 2292      |       |       | 1341    |       |       |
| Maternal Education                                    | 2292      |       |       | 1341    |       |       |
| Less than high school                                 | 0.36      | 0.39  |       | 0.39    |       |       |
| High school diploma / GED                             | 0.34      | 0.33  |       | 0.33    |       |       |
| Beyond high school                                    | 0.31      | 0.28  |       | 0.28    |       |       |
| Married Mother                                         | 2292      |       |       | 1341    |       |       |
| Teenage Mother                                        | 2292      |       |       | 1341    |       |       |
| Parents lived together                                | 2292      |       |       | 1341    |       |       |
| Immigrant mother                                      | 2292      |       |       | 1341    |       |       |
| Home language Spanish                                  | 2292      |       |       | 1341    |       |       |
| Maternal Depression                                   | 2292      |       |       | 1341    |       |       |

*Note. p level of treatment/control difference: * p < .05. ** p < .01. *** p < .001.

T-tests for differences in means were conducted for continuous variables, chi-square tests for categorical variables.

\( a \) = Composited from CES-D Depression scale. Weight used = CHSPR2003WTCA.
Figure 1. Idealized graphical models of the three hypotheses surrounding Head Start and parental pre-academic stimulation.
Table 3

*Relationship between Parental Pre-academic Stimulation, Assignment to Head Start, and their Interaction*

<table>
<thead>
<tr>
<th>Outcomes</th>
<th>WJ Applied Problems</th>
<th>WJ Letter-Word</th>
<th>PPVT</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Model 1</td>
<td>Model 2</td>
<td>Model 1</td>
</tr>
<tr>
<td>Low Pre-academic Stimulation (Dummy Bottom 10%)</td>
<td>-4.49***</td>
<td>-6.93***</td>
<td>-3.29**</td>
</tr>
<tr>
<td></td>
<td>(1.21)</td>
<td>(1.74)</td>
<td>(1.00)</td>
</tr>
<tr>
<td>Middle Pre-academic Stimulation (Dummy 10-90% - REFERENCE)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>High Pre-academic Stimulation (Dummy Top 10%)</td>
<td>-0.97</td>
<td>-0.65</td>
<td>1.34</td>
</tr>
<tr>
<td></td>
<td>(1.04)</td>
<td>(1.75)</td>
<td>(0.99)</td>
</tr>
<tr>
<td>Treatment (Dummy Assignment to H.S.)</td>
<td>2.46***</td>
<td>1.93*</td>
<td>3.44***</td>
</tr>
<tr>
<td></td>
<td>(0.68)</td>
<td>(0.76)</td>
<td>(0.63)</td>
</tr>
<tr>
<td>Low Pre-academic Stimulation x Treatment</td>
<td>5.15*</td>
<td>-2.16~</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(2.32)</td>
<td>(1.25)</td>
<td></td>
</tr>
<tr>
<td>High Pre-academic Stimulation x Treatment</td>
<td>-0.41</td>
<td>-4.71*</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(2.14)</td>
<td>(2.09)</td>
<td></td>
</tr>
<tr>
<td>Intercept</td>
<td>81.08</td>
<td>80.02</td>
<td>88.80</td>
</tr>
<tr>
<td>N</td>
<td>3185</td>
<td>3185</td>
<td>3185</td>
</tr>
<tr>
<td>R²</td>
<td>0.32</td>
<td>0.33</td>
<td>0.32</td>
</tr>
</tbody>
</table>

*Note.* Standard errors in parentheses calculated using jackknife replicate weights.  
~ p < 0.10.  * p < 0.05.  ** p < 0.01.  *** p < 0.001.  
WJ = Woodcock-Johnson. PPVT= Peabody Picture Vocabulary Test. Covariates (centered at mean): lagged DV, child cohort, child race, child gender, child disability status, maternal education, maternal marital status, caregiver depression, teenage mother status, caregiver age, maternal immigration status, home language, child age at spring 2003 assessment, number of weeks elapsed between 09/01/02 and spring 2003 child assessment, and if baseline child assessment in Spanish. Weight used = CHSPR2003WTCA.
Figure 2. Histogram of composite parental pre-academic stimulation measure with normal overlay.

Figure 3. Predicted Head Start treatment effect for the WJ Applied Problems, consistent with a compensatory pattern. Head Start effect sizes calculated at each level of parental pre-academic stimulation, indicated by “d”. Asterisks indicate the treatment-control difference at low parental stimulation to differ significantly from the treatment-control difference at middle parental stimulation. (\( p < 0.10 \). * \( p < 0.05 \). ** \( p < 0.01 \). *** \( p < 0.001 \).)
Figure 4. Predicted Head Start treatment effect for the WJ Letter-Word, consistent with a “Goldilocks” pattern. Head Start effect sizes calculated at each level of parental pre-academic stimulation, indicated by “d”. Asterisks indicate the treatment-control difference at low and high parental stimulation to differ significantly from the treatment-control difference at middle parental stimulation. (~ p < 0.10.  * p < 0.05.  ** p < 0.01.  *** p < 0.001.)
Figure 5. Predicted Head Start treatment effect for the PPVT, not fully consistent with any hypothesis pattern. Head Start effect sizes calculated at each level of parental pre-academic stimulation, indicated by “d”. Asterisks indicate the treatment-control difference at high parental stimulation to differ significantly from the treatment-control difference at middle parental stimulation. (~ p < 0.10. * p < 0.05. ** p < 0.01. *** p < 0.001.)