



**Third Grade Follow-up to the  
Head Start Impact Study**

**Executive Summary**

**OPRE Report 2012-45b**

**October 2012**

# THIRD GRADE FOLLOW-UP TO THE HEAD START IMPACT STUDY

## FINAL REPORT

### OPRE Report 2012-45b

**October 2012**

Michael Puma, Chesapeake Research Associates, Stephen Bell, Abt Associates, Ronna Cook, Ronna Cook Associates, Camilla Heid, Pam Broene, and Frank Jenkins, Westat, Andrew Mashburn, Portland State University, and Jason Downer, University of Virginia

Submitted to:

Jennifer Brooks, Project Officer  
Office of Planning, Research and Evaluation  
Administration for Children and Families  
U.S. Department of Health and Human Services

Contract Number: HHSP23320062929YC

Project Director: Camilla Heid  
Westat  
1600 Research Boulevard  
Rockville, MD 20850

This report is in the public domain. Permission to reproduce is not necessary. Suggested citation: Mike Puma, Stephen Bell, Ronna Cook, Camilla Heid, Pam Broene, Frank Jenkins, Andrew Mashburn, and Jason Downer (2012). *Third Grade Follow-up to the Head Start Impact Study Final Report, Executive Summary*. OPRE Report # 2012-45b. Washington, DC: Office of Planning, Research and Evaluation, Administration for Children and Families, U.S. Department of Health and Human Services.

#### Disclaimer

The views expressed in this publication do not necessarily reflect the views or policies of the Office of Planning, Research and Evaluation, the Administration for Children and Families, or the U.S. Department of Health and Human Services.

This report and other reports sponsored by the Office of Planning, Research and Evaluation are available at <http://www.acf.hhs.gov/programs/opre>.



# ***Third Grade Follow-up to the Head Start Impact Study***

## ***Executive Summary***

### ***Introduction***

Since its beginning in 1965 as a part of the War on Poverty, Head Start’s goal has been to boost the school readiness of low-income children. Based on a “whole child” model, the program provides comprehensive services that include preschool education; medical, dental, and mental health care; nutrition services; and efforts to help parents foster their child’s development. Head Start services are designed to be responsive to each child’s and family’s ethnic, cultural, and linguistic heritage.

In the 1998 reauthorization of Head Start, Congress mandated that the US Department of Health and Human Services (DHHS) determine, on a national level, the impact of Head Start on the children it serves. As noted by the Advisory Committee on Head Start Research, this legislative mandate required that the impact study address two main research questions:<sup>1</sup>

#### **Study Goals**

- 1) Determine the impact of Head Start on:
  - Children’s school readiness, and
  - Parental practices that support children’s development.
- 2) Determine under what circumstances Head Start achieves its greatest impact and for which children.

- “What difference does Head Start make to key outcomes of development and learning (and in particular, the multiple domains of school readiness) for low-income children? What difference does Head Start make to parental practices that contribute to children’s school readiness?”
- “Under what circumstances does Head Start achieve the greatest impact? What works for which children? What Head Start services are most related to impact?”

The *Head Start Impact Study Final Report*<sup>2</sup> addressed these questions by reporting on the impacts of Head Start on children and families during the children’s preschool, kindergarten, and 1<sup>st</sup> grade years. This *Third Grade Follow-up to the Head Start Impact Study Final Report* addresses these same questions by looking at longer-term effects through the end of 3<sup>rd</sup> grade.

### ***Background for the Head Start Impact Study***

The Head Start Impact Study (HSIS) was conducted with a nationally representative sample of 84 grantee/delegate agencies and included nearly 5,000 newly entering, eligible 3- and

---

<sup>1</sup> U.S. Department of Health and Human Services, Administration for Children, Youth, and Families. (January, 1999). *Evaluating Head Start: A recommended framework for studying the impact of the Head Start program*. Washington, DC: Author.

<sup>2</sup> U.S. Department of Health and Human Services, Administration for Children and Families (January, 2010). *Head Start Impact Study: Final Report*. Washington, DC. [http://www.acf.hhs.gov/programs/opre/hs/impact\\_study/reports/impact\\_study/hs\\_impact\\_study\\_final.pdf](http://www.acf.hhs.gov/programs/opre/hs/impact_study/reports/impact_study/hs_impact_study_final.pdf)

4-year-old children who were randomly assigned to either: (1) a Head Start group that had access to Head Start program services or (2) a control group that did not have access to Head Start, but could enroll in other early childhood programs or non-Head Start services selected by their parents. Data collection began in fall 2002 and continued through 2008, following children from program application through the spring of their 3<sup>rd</sup> grade year.

The study was designed to separately examine two cohorts of children, newly entering 3-and 4-year-olds. This design reflects the hypothesis that different program impacts may be associated with different age of entry into Head Start. Differential impacts are of particular interest in light of a trend of increased enrollment of the 3-year-olds in some grantee/delegate agencies presumably due to the growing availability of preschool options for 4-year-olds. Consequently, the study included two separate samples: a newly entering 3-year-old group<sup>3</sup> (to be studied through two years of possible Head Start participation, kindergarten 1<sup>st</sup> grade, and 3<sup>rd</sup> grade), and a newly entering 4-year-old group (to be studied through one year of Head Start participation, kindergarten, 1<sup>st</sup> grade, and 3<sup>rd</sup> grade).

The study showed that the two age cohorts varied in demographic characteristics. The racial/ethnic characteristics of newly entering children in the 3-year-old cohort were substantially different from the characteristics of children in the newly entering 4-year-old cohort. While the newly entering 3-year-olds were relatively evenly distributed between Black children and Hispanic children about half of newly entering 4-year-olds were Hispanic children (see Exhibit 1). The ethnic difference was also reflected in the age-group differences in child and parent language.

**Exhibit 1. Percentage of Children by Racial/Ethnic Characteristics and By Age Cohort**

<b>Child Race/Ethnicity</b>	<b>3-Year-Old Cohort</b>	<b>4-Year-Old Cohort</b>
<b>Hispanic</b>	37.4%	51.6%
<b>Black</b>	32.8%	17.5%
<b>White/Other</b>	29.8%	30.8%

This study is unique in its design and differs from prior evaluations of early childhood programs:

- **Randomized Control.** The Congressional mandate for this study had a clearly stated goal of producing causal findings, i.e., the purpose was to determine if access to Head Start caused better developmental and parenting outcomes for participating children and families. To do this, the study randomly assigned Head Start applicants either to a Head Start group that was allowed to enroll, or to a “control” group that could not. This procedure ensured comparability between the two groups at program entry, so that later differences can be causally attributed to Head Start.
- **Representative Sample of Programs and Children.** Most random assignment studies are conducted in small demonstration programs or in a small number of operating sites, usually those that volunteer to be included in the research. In

---

<sup>3</sup> The study design allowed 3-year-old cohort control group children to reapply to Head Start after the first year.

contrast, the Head Start Impact Study is based on a nationally representative sample of Head Start programs and children, with a few exceptions for programs serving particular populations. This makes the study results generalizable to the vast majority of programs nationwide at the time the study was fielded in 2002, not just the selected study sample. Unlike most studies, it examines the average impact of programs that represent the full range of intensity and quality and adherence to the established Head Start program standards (i.e., the best, the worst, and those in the middle of a fully implemented program).

- **Examination of a Comprehensive Set of Outcomes Over Time.** The study quantifies the overall impact of Head Start separately for 3- and 4-year-old children in four key program domains—cognitive development, social-emotional development, health status and services, and parenting practices—following them through early elementary school. These impacts are measured by examining the difference in outcomes between children assigned to the Head Start group and those assigned to the control group.

Other study features that must be considered in interpreting the study findings include:

- **Control Group Children Did Not All Stay at Home.** Children who were placed in the control or comparison group were allowed to enroll in other non-parental care or non-Head Start child care or programs selected by their parents. They could remain at home in parent care, or enroll in a child care or preschool program. Consequently, the impact of Head Start was determined by a comparison to a mixture of alternative care settings rather than against a situation in which children were artificially prevented from obtaining child care or early education programs outside of their home. Approximately 60 percent of the control group children participated in child care or early education programs during the first year of the study, with 13.8 percent of the 4-year-olds in the control group and 17.8 percent of the 3-year-olds in the control group finding their way into Head Start during this year. Preventing families from seeking out alternative care or programs for their children is both infeasible and unethical. The design used here answers the policy question, how well does Head Start do when compared against the other types of services or care that low-income children could receive in fall 2002.
- **Impacts Represent the Effects of One Year of Head Start.** For children in the 4-year-old cohort, the study provides the impact of Head Start for a single year, i.e., the year before they are eligible to enter kindergarten. The impacts for the 3-year-old cohort reflect the benefits of being provided an earlier year of Head Start (as compared to the control group, which received access to Head Start at age 4.) At the end of one year of Head Start participation, the 3-year-old cohort—but not the 4-year-old cohort—had another year to go before they started kindergarten. It was not feasible or desirable for this study to prevent 3-year-olds from participating in Head Start for two years. Thus, the study could not directly assess the receipt of one year versus two years of Head Start. Rather, it addresses the receipt of an earlier year—whether having Head Start available at age three is helpful to children brought to the program at that age, or whether those children would be just as well off, if the program did not enroll them until age four. This is not only important to individual families; it also answers an important policy question. To answer this question, the

best approach is to preclude program entry at age three while allowing it at age four and contrast outcomes after that point with statistically equivalent children never excluded from the program. By design, the study did not attempt to control children's experiences after their first Head Start year.

The Head Start Impact Study is a comprehensive, carefully designed study of a large-scale early childhood program that has existed for more than 40 years. It is designed to address the overall average impact of the Head Start program as it existed in 2002. The findings cannot be directly compared to more narrowly focused studies of other early childhood programs. The Advisory Committee on Head Start Research and Evaluation, which developed the blueprint for this study, recommended that “the research and findings should be used in combination with the rest of the Head Start research effort to improve the effectiveness of Head Start programs for children and families” (Advisory Committee on Head Start Research and Evaluation, 1999, p. 44). The Third Grade Follow-up to the Head Start Impact Study builds upon the existing randomized control design in the HSIS in order to determine the longer term impact of the Head Start program on the well-being of children and families through the end of 3<sup>rd</sup> grade.

## ***Key Findings***

Looking across the full study period, from the beginning of Head Start through 3<sup>rd</sup> grade, the evidence is clear that access to Head Start improved children's preschool outcomes across developmental domains, but had few impacts on children in kindergarten through 3<sup>rd</sup> grade. Providing access to Head Start was found to have a positive impact on the types and quality of preschool programs that children attended, with the study finding statistically significant differences between the Head Start group and the control group on every measure of children's preschool experiences in the first year of the study. In contrast, there was little evidence of systematic differences in children's elementary school experiences through 3<sup>rd</sup> grade, between children provided access to Head Start and their counterparts in the control group.

In terms of children's well-being, there is also clear evidence that access to Head Start had an impact on children's language and literacy development while children were in Head Start. These effects, albeit modest in magnitude, were found for both age cohorts during their first year of admission to the Head Start program. However, these early effects rapidly dissipated in elementary school, with only a single impact remaining at the end of 3<sup>rd</sup> grade for children in each age cohort.

With regard to children's social-emotional development, the results differed by age cohort and by the person describing the child's behavior. For children in the 4-year-old cohort, there were no observed impacts through the end of kindergarten but favorable impacts reported by parents and unfavorable impacts reported by teachers emerged at the end of 1<sup>st</sup> and 3<sup>rd</sup> grades. One unfavorable impact on the children's self-report emerged at the end of 3<sup>rd</sup> grade. In contrast to the 4-year-old cohort, for the 3-year-old cohort there were favorable impacts on parent-reported social emotional outcomes in the early years of the study that continued into early elementary school. However, there were no impacts on teacher-reported measures of social-emotional development for the 3-year-old cohort at any data collection point or on the children's self-reports in 3<sup>rd</sup> grade.

In the health domain, early favorable impacts were noted for both age cohorts, but by the end of 3<sup>rd</sup> grade, there were no remaining impacts for either age cohort. Finally, with regard to parenting practices, the impacts were concentrated in the younger cohort. For the 4-year-old cohort, there was one favorable impact across the years while there were several favorable impacts on parenting approaches and parent-child activities and interactions (all reported by parents) across the years for the 3-year-old cohort.

In summary, there were initial positive impacts from having access to Head Start, but by the end of 3<sup>rd</sup> grade there were very few impacts found for either cohort in any of the four domains of cognitive, social-emotional, health and parenting practices. The few impacts that were found did not show a clear pattern of favorable or unfavorable impacts for children.

In addition to looking at Head Start's average impact across the diverse set of children and families who participated in the program, the study also examined how impacts varied among different types of participants. There is evidence that for some outcomes, Head Start had a differential impact for some subgroups of children over others. At the end of 3<sup>rd</sup> grade for the 3-year-old cohort, the most striking sustained subgroup findings were found in the cognitive domain for children from high risk households as well as for children of parents who reported no depressive symptoms. Among the 4-year-olds, sustained benefits were experienced by children of parents who reported mild depressive symptoms, severe depressive symptoms, and Black children.

## ***Overview of Study Methods***

### **Random Assignment**

Newly entering 3- and 4-year-old Head Start applicants were randomly assigned either to a Head Start group that for one year had access to Head Start services, or to a control group that could receive any other non-Head Start services chosen by their parents.

To reliably answer the research questions outlined by Congress, a nationally representative sample of Head Start programs and newly entering 3- and 4-year-old children was selected, and children were randomly assigned either to a Head Start group that had access to Head Start services in the initial year of the study or to a control group that could receive any other non-Head Start services available in the community, chosen by their parents. In fact, approximately 60 percent of control group parents enrolled their children in some other type of preschool program in the first year. In addition, all children in the 3-year-old cohort could receive Head Start services in the second year. Under this randomized design, a simple comparison of outcomes for the two groups yields an unbiased estimate of the impact of access to Head Start in the initial year on children's school readiness. This research design ensured that the Head Start and control groups did not differ in any systematic or unmeasured way except through their access to Head Start services. It is important to note that, because the control group in the 3-year-old cohort was given access to Head Start in the second year, the findings for this age group reflect the added benefit of providing access to Head Start at age 3 vs. at age 4, *not* the total benefit of having access to Head Start for two years.

In addition to random assignment, this study is set apart from most program evaluations because it includes a nationally representative sample of programs, making results generalizable to the Head Start program as a whole, not just to the selected samples of programs

and children. However, the study does not represent Head Start programs serving special populations, such as tribal Head Start programs, programs serving migrant and seasonal farm workers and their families, or Early Head Start. Further, the study does not represent the 15 percent of Head Start programs in which the pool of applicants for Head Start slots was too small to allow for an adequate control group.

### **Study Sample**

The nationally representative study sample, spread over 23 different states, consisted of a total of 84 randomly selected grantees/delegate agencies, 383 randomly selected Head Start centers, and a total of 4,667 newly entering children; 2,559 3-year-olds and 2,108 4-year-olds.

Selected Head Start grantees and centers had to have a sufficient number of applicants for the 2002-2003 program year to allow for the creation of a control group without requiring Head Start slots to go unfilled. As a consequence, the study was conducted in communities that had more children eligible for Head Start than could be served with the existing number of funded slots.

At each of the selected Head Start centers, program staff provided information about the study to parents at the time enrollment applications were distributed. Parents were told that enrollment procedures would be different for the 2002-2003 Head Start year and that some decisions regarding enrollment would be made using a lottery-like process. Local agency staff implemented their typical process of reviewing enrollment applications and screening children for admission to Head Start based on criteria approved by their respective Policy Councils. No changes were made to these locally established ranking criteria.

Information was collected on all children determined to be eligible for enrollment in fall 2002, and an average sample of 27 children per center was selected from this pool: 16 who were assigned to the Head Start group and 11 who were assigned to the control group. Random assignment was done separately for two study samples—newly entering 3-year-olds (to be studied through two years of potential Head Start participation, kindergarten, 1<sup>st</sup> grade, and 3<sup>rd</sup> grade) and newly entering 4-year-olds (to be studied through one year of Head Start participation, kindergarten, 1<sup>st</sup> grade, and 3<sup>rd</sup> grade).

The total sample, spread over 23 different states, consisted of 84 randomly selected Head Start grantees/delegate agencies, 383 randomly selected Head Start centers, and a total of 4,667 newly entering children, including 2,559 in the 3-year-old group and 2,108 in the 4-year-old group.<sup>4</sup>

Data collection began in the fall of 2002 and continued through the spring of 2008, following children from entry into Head Start through the end of 3<sup>rd</sup> grade. Comparable data were collected for both Head Start and control group children, including interviews with parents, direct child assessments, surveys of Head Start, other early childhood, and elementary school teachers, interviews with center directors and other care providers at the preschool level, direct observations of the quality of various preschool care settings, and teacher or care provider assessments of children. For the Third Grade Follow-up, principal surveys and teacher ratings by the principal were added to the data collection. Response rates were consistently quite high, approximately 80 percent for parents and children throughout the study. Teacher response rates were higher at the preschool level (about 80 percent) and gradually decreased as the child

---

<sup>4</sup> The sample of 3-year-olds is slightly larger than the sample of 4-year-olds to ensure that an adequate sample size was maintained, given the possibility of higher study attrition resulting from an additional year of longitudinal data collection for the younger children.

progressed through school (slightly above 60 percent by the end of 3<sup>rd</sup> grade). Principal data were collected only during 3<sup>rd</sup> grade and the response rate was about the same as for 3<sup>rd</sup> grade teachers.

Although every effort was made to ensure compliance with random assignment, some children accepted into Head Start did not participate in the program (about 15 percent for the 3-year-old cohort and 20 percent for the 4-year-old cohort), and some children assigned to the non-Head Start group nevertheless entered the program in the first year (about 17 percent for 3-year-olds and 14 percent for 4-year-olds), typically at centers that were not in the study sample. These families are referred to as “no shows” and “crossovers.” Statistical procedures for dealing with these events are discussed in the report. Thus, the findings in this report provide estimates of both the impact of *access* to Head Start using the sample of all randomly assigned children (referred to as Intention to Treat, or ITT) and the impact of *actual* Head Start participation (adjusting for the no shows and crossovers, referred to as Impacts on the Treated or IOT).

## ***Findings: Head Start Through 3<sup>rd</sup> Grade***

### ***Impact on Children’s Experiences in Preschool and Early Elementary School***

There are clear impacts on the types and quality of children’s child care, early education, and school experiences at the preschool level but not in the early elementary grades. At the preschool level, the story is far clearer, as providing access to Head Start was found to have a positive impact on children’s experiences across many measures of early childhood experience. There were statistically significant differences between the Head Start group and the control group on every measure of children’s preschool experiences measured in this study. These effects were found both for the 4-year-old cohort and for the 3-year-old cohort during the year in which they were admitted to Head Start. The measures that were examined included, but were not limited to, teacher qualifications, including their training and education; classroom literacy and math instructional activities; classroom teacher-child ratios; the nature of teacher-child interactions; and global measures of the care environment as measured by ECERS-R/FDCRS scores. The differences in magnitude were quite large, driven in part by the large proportion of children in the control group who were in parent care (i.e., nearly four out of ten children remained at home with their parents when Head Start was unavailable to them).<sup>5</sup> Yet, analyses excluding those children, and thus comparing only children in the Head Start and control groups who were in non-parental care, largely showed the same pattern of differences, albeit somewhat smaller.

The preschool experiences of children in the 3-year-old cohort were very different in the second year of the study, when most were 4 years old. The majority of the children (both Head Start and control group) were in some type of center-based care by the this year, and with three small exceptions, the observed treatment and control differences disappeared in the age 4 year. That is, once the control group had access to Head Start, the earlier differences on the measures of their early childhood care environments all but vanished.

---

<sup>5</sup> For these analyses, children in parent care were included and given a score of zero.

Providing access to Head Start did not have much impact on the types of schools children attended from kindergarten through 3<sup>rd</sup> grade. By the end of 3<sup>rd</sup> grade, the study sample had dispersed into nearly 1,600 individual elementary schools. On average, both Head Start and control group children attended public schools, with the percentage enrolled in public school increasing from kindergarten to 3<sup>rd</sup> grade for children in the study sample as a whole. For the 4-year-old cohort, approximately 80 percent of the children were enrolled in public school at the end of kindergarten. By the end of 3<sup>rd</sup> grade, 98 percent of the children were enrolled in public school. For the 3-year-old cohort, approximately 85 percent were enrolled in public schools at the end of kindergarten and 96 percent by the end of the 3<sup>rd</sup> grade. The schools' percentages of students at or above proficient on state assessments in math and reading were in the middle of the respective state averages (55 to 67 percent depending on the subject and year), indicating that on average the schools attended by the study children were not among the worst or best schools in their respective states. In the 3-year-old cohort's kindergarten year, a significant difference was found in the school-wide average math proficiency scores for Head Start children and control group children, with the difference favoring the control group. For the 3-year-old cohort in the 3<sup>rd</sup> grade, a significant difference was found between average reading/language arts proficiency scores at the schools attended by the Head Start and control group children, this time favoring the Head Start group.

Not surprisingly, the study children attended schools with much higher levels of poverty than schools nationwide (as indicated by proportions of students eligible for free- and reduced-price lunch—66-67 percent) and were in schools with higher proportions of minority students (approximately 60 percent of students). With only a few exceptions, teacher and classroom characteristics did not differ significantly between children in the Head Start group and those in the control group. The school experiences measures were limited in kindergarten and 1<sup>st</sup> grade, while a wider range of school, classroom, and teacher measures assessed the children's 3<sup>rd</sup> grade year. The few differences that were found varied and most were found at the end of 3<sup>rd</sup> grade, sometimes favoring the control group and sometimes favoring the Head Start group.

For children in the 4-year-old cohort at the end of 3<sup>rd</sup> grade year, there is evidence that the Head Start children were in schools that, according to principals, had greater access to computers (compared to the non-Head Start children), and were more likely to have summer school programs. Head Start children were more likely to have a teacher with a standard state teaching certificate, and their schools were more likely to use standardized tests to a great extent to compare subgroups of students. On the other hand, the schools attended by the Head Start children were more likely to have higher levels of student mobility.

At the end of 3<sup>rd</sup> grade for children in the 3-year-old cohort, principal reports showed that Head Start children were in schools that had more adequate school facilities, lower staff turnover, and a higher percentage of 3<sup>rd</sup> grade students scoring at the proficient or higher level on the state reading/language arts assessment. On the other hand, the schools attended by Head Start children had higher percentages of students with disabilities, and according to principals, required more attention to deal with student discipline problems. In terms of classroom and teacher measures, the Head Start children were more likely to have access to an instructional aide in their classroom, to have a lower percentage of children reading below grade level in their reading/language arts class, and to have a teacher who majored in education as an undergraduate.

## ***Impacts on Children's Cognitive Development***

The cognitive domain consisted of: (1) direct assessments of language and literacy skills, pre-writing skills (in Head Start years only), and math skills; (2) teacher reports of children's school performance; and (3) parent reports of child literacy skills and grade promotion.

There is clear evidence that Head Start had a statistically significant impact on children's language and literacy development while children were in Head Start. These effects, albeit modest in magnitude, were found for both age cohorts during their first year of admission to the Head Start program. However, these early effects dissipated in elementary school, with only a single impact remaining at the end of 3<sup>rd</sup> grade for children in each age cohort: a favorable impact for the 4-year-old cohort (ECLS-K Reading) and an unfavorable impact for the 3-year-old cohort (grade promotion).

Impacts aside, these children remain disadvantaged compared to their same-age peers; the scores of both the Head Start and the control group children remained lower than the norm for the population. At the end of 3<sup>rd</sup> grade, HSIS children (both Head Start and control group children) in the 4-year-old cohort, on average, scored about eight points (approximately one-half of a standard deviation) lower than a national sample of third graders on the ECLS-K Reading Assessment and the promotion rate<sup>6</sup> for the 3-year old cohort was two to three percent lower than the predicted national promotion rate for children at the end of 3<sup>rd</sup> grade.

For mathematics, impacts were found only on a single outcome measure (Woodcock Johnson III Applied Problems) and only for the 3-year-old cohort at the end of their Head Start year.

The findings from the cognitive domain are summarized by age cohort below.<sup>7</sup> Exhibits 2a and 2b present all statistically significant cognitive impacts and their effect sizes<sup>8</sup> from the Intent to Treat (ITT) analysis.

### **4-Year-Old Cohort**

- At the end of the Head Start year, there was strong evidence that the Head Start group demonstrated better skills on the following six child outcomes related to children's language and literacy development: (1) Peabody Picture Vocabulary Test (PPVT) (vocabulary); (2) Woodcock-Johnson III (WJIII) Letter-Word Identification; (3) WJIII Spelling; (4) WJIII Pre-Academic Skills; (5) Color Identification; and (6) Letter Naming.

---

<sup>6</sup> Warren and Saliba (2012) generated a predicted 3<sup>rd</sup> grade national retention rate using an age-grade delay model as a proxy for retention. See Chapter 4 for additional information.

<sup>7</sup> Three levels of evidence are considered in this report: (1) strong evidence is used for impacts statistically significant at  $p \leq 0.05$ , and the result holds up under the Benjamini-Hochberg test for multiple comparisons; (2) moderate evidence signifies a particular impact is statistically significant at  $p \leq 0.05$ , but this result does not hold up under the test for multiple comparisons; and (3) suggestive evidence signifies a particular impact is statistically significant under a relaxed standard  $p \leq 0.10$ , and the result may or may not hold up under the test for multiple comparisons.

<sup>8</sup> The effect size is simply the impact estimate divided by the standard deviation of the outcome measure in the population. The effect size provides an indication of the magnitude of each impact that is independent of the particular instrument or measure used. More discussion of the interpretation of effect sizes is provided in Chapter 2.

- Parents of children in the Head Start group reported that their children had greater emerging literacy skills at the end of Head Start than did parents of children in the control group.
- There were no impacts for 4-year-olds in the cognitive domain at the end of kindergarten.
- At the end of 1<sup>st</sup> grade, there was suggestive evidence of a positive impact of access to Head Start on PPVT (vocabulary) scores.
- At the end of 3<sup>rd</sup> grade, there was suggestive evidence of a positive impact of access to Head Start on the ECLS-K Reading Assessment.
- No significant impacts were found for math skills, pre-writing, children's promotion, or teacher report of children's school accomplishments or abilities in any year.

### **3-Year-Old Cohort**

- At the end of their Head Start year, there was strong evidence of better skills for the Head Start group on the following five outcomes related to children's language and literacy development: (1) PPVT (vocabulary), (2) WJIII Letter-Word, (3) Preschool Comprehensive Test of Phonological and Print Processing (CTOPPP) Elision, (4) Letter Naming, and (5) WJIII Pre-Academic Skills. There was also a statistically significant impact on the measure of children's pre-writing skills. Children in the Head Start group were found to have more advanced math skills than their counterparts at the end of the Head Start year on the WJIII test of Applied Problems.
- Favorable impacts of Head Start were also found on parental reports of children's emergent literacy skills at the end of the Head Start year.
- At the end of the age 4 year, few statistically significant impacts were found. However, two impacts persisted related to children's literacy skills. Children in the Head Start group scored higher than children in the control group on CTOPPP Elision as well as on parents' reports of their literacy skills.
- As with the 4-year-old cohort, there was no strong evidence of impacts on children's language, literacy, or math measures at the end of kindergarten or at the end of 1st grade. However, there was suggestive evidence of an impact on Oral Comprehension at the end of 1st grade.
- At the end of 3rd grade, there was suggestive evidence of an unfavorable impact—the parents of the Head Start group children reported a significantly lower child grade promotion rate than the parents of the non-Head Start group children.
- No statistically significant impacts were found for teacher reports of children's school performance, with the exception of a lower teacher assessment in kindergarten of Head Start children's math ability. This was not supported by children's scores on the three direct math assessments, which showed no evidence of math differences. However, the schools attended by the control group children in the 3-year-old cohort during their kindergarten year reported a significantly higher percentage of students at or above the proficient level in math than the schools attended by the Head Start group children.

**Exhibit 2a. Summary of ITT Cognitive Impacts for 4-Year-Olds by Year**

Measure	Age 4 (Head Start Year)	K	1 <sup>st</sup> Grade	3 <sup>rd</sup> Grade
<b>Language, Literacy, and Pre-Writing</b>				
Color Identification	0.16	NA	NA	NA
Pre-Writing (McCarthy Draw a Design)		NA	NA	NA
Emergent Literacy Scale (parent report)	0.31	NA	NA	NA
Letter Naming	0.25		NA	NA
Test of Phonological Processing (CTOPPP Elision)			NA	NA
Receptive Vocabulary (PPVT)	0.09		0.09	
Letter-Word Identification (WJIII)	0.22			
Spelling (WJIII)	0.15			NA
Oral Comprehension (WJIII)				NA
Pre-Academic Skills (WJIII)	0.19			NA
Phonetic Skills/ Word Attack (WJIII)	NA			NA
Basic Reading (WJIII)	NA			NA
Academic Applications (WJIII)	NA	NA		NA
Academic Skills (WJIII)	NA	NA		NA
Passage Comprehension (WJIII)	NA	NA		NA
ECLS-K Reading	NA	NA		0.11
Writing Sample (WJIII)	NA	NA		NA
<b>Spanish Language</b>				
Receptive Vocabulary (TVIP)				NA
Batería WM Identificación de letras y palabras				
<b>Math</b>				
One-to-One Counting (Counting Bears)		NA	NA	NA
Applied Problems (WJIII)				
Quantitative Concepts (WJIII)	NA			NA
Math Reasoning (WJIII)	NA			NA
Calculation (WJIII)	NA	NA		
<b>School Performance</b>				
School Accomplishments	NA			NA
Promotion (parent report)	NA			
Language and Literacy Ability	NA			
Math Ability	NA			
Math Skills	NA	NA	NA	
Reading/Language Arts Skills	NA	NA	NA	
Social Studies and Science Ability	NA			NA

**KEY:**

-  Dark gray cell indicates a significant favorable impact ( $p \leq 0.10$ ).
-  Diagonal mark cell indicates a significant unfavorable impact ( $p \leq 0.10$ ).
-  Blank cell indicates no significant impact.

An \* indicates that the interpretation of the outcome is unclear.  
 NA indicates that no data were collected for this outcome at this data collection point.

NOTE: Intent to Treat (ITT) effect sizes are presented only for statistically significant differences ( $p \leq 0.10$ ). The effect size is simply the impact estimate divided by the standard deviation of the outcome measure in the population. The effect size provides an indication of the magnitude of each impact that is independent of the particular instrument or measure used. More discussion of the interpretation of effect sizes is provided in Chapter 2.

## Exhibit 2b. Summary of ITT Cognitive Impacts for 3-Year-Olds by Year

Measure	Age 3 (Head Start Year)	Age 4	K	1 <sup>st</sup> Grade	3 <sup>rd</sup> Grade
<b>Language, Literacy, and Pre-Writing</b>					
Color Identification			NA	NA	NA
Pre-Writing (McCarthy Draw a Design)	<b>0.14</b>		NA	NA	NA
Emergent Literacy Scale (parent report)	<b>0.35</b>	<b>0.16</b>	NA	NA	NA
Letter Naming	<b>0.24</b>			NA	NA
Test of Phonological Processing (CTOPPP Elision)	<b>0.10</b>	<b>0.15</b>		NA	NA
Receptive Vocabulary (PPVT)	<b>0.18</b>				
Letter-Word Identification (WJIII)	<b>0.26</b>				
Spelling (WJIII)					NA
Oral Comprehension (WJIII)				<b>0.08</b>	NA
Pre-Academic Skills (WJIII)	<b>0.22</b>				NA
Phonetic Skills/Word Attack (WJIII)	NA	NA			NA
Basic Reading (WJIII)	NA	NA			NA
Academic Applications (WJIII)	NA	NA	NA		NA
Academic Skills (WJIII)	NA	NA	NA		NA
Passage Comprehension (WJIII)	NA	NA	NA		NA
ECLS-K Reading	NA	NA	NA		
Writing Sample (WJIII)	NA	NA	NA		NA
<b>Spanish Language</b>					
Receptive Vocabulary (TVIP)					NA
Batería WM Identificación de letras y palabras			<b>0.26</b>		
<b>Math</b>					
One-to-One Counting/Counting Bears			NA	NA	NA
Applied Problems (WJIII)	<b>0.15</b>				
Quantitative Concepts (WJIII)	NA	NA			NA
Math Reasoning (WJIII)	NA	NA			NA
Calculation (WJIII)	NA	NA	NA		
<b>School Performance</b>					
School Accomplishments	NA	NA			NA
Promotion (parent report)	NA	NA			<b>-0.11</b>
Language and Literacy Ability	NA	NA			
Math Ability	NA	NA	<b>-0.19</b>		
Math Skills	NA	NA	NA	NA	
Reading/Language Arts Skills	NA	NA	NA	NA	
Social Studies and Science Ability	NA	NA			NA

### KEY:

-  Dark gray cell indicates a significant favorable impact ( $p \leq 0.10$ ).
-  Diagonal mark cell indicates a significant unfavorable impact ( $p \leq 0.10$ ).
-  Blank cell indicates no significant impact.

An \* indicates that the interpretation of the outcome is unclear.

NA indicates that no data were collected for this outcome at this data collection point.

NOTE: Intent to Treat (ITT) effect sizes are presented only for statistically significant differences ( $p \leq 0.10$ ). The effect size is simply the impact estimate divided by the standard deviation of the outcome measure in the population. The effect size provides an indication of the magnitude of each impact that is independent of the particular instrument or measure used. More discussion of the interpretation of effect sizes is provided in Chapter 2.

## ***Impacts on Children's Social-Emotional Development***

The social-emotional domain consisted of parent-reported measures during the Head Start years, reports by both parents and teachers in all elementary school years, with child self-reports added at the end of 3<sup>rd</sup> grade. Measures of children's behavior, social skills and approaches to learning, parent-child relationships, teacher child relationships, school adjustment, peer relationships and school experiences were assessed.

With regard to children's social-emotional development, the results differed by age cohort and by the source of the information on the child's behavior. For children in the 4-year-old cohort, there were no observed impacts through the end of kindergarten and then favorable impacts reported by parents and unfavorable impacts reported by teachers at the end of 1<sup>st</sup> and 3<sup>rd</sup> grades and children at the end of 3<sup>rd</sup> grade.

In contrast, the early favorable social emotional impacts reported by parents for the 3-year-old cohort continued into early elementary school. There were favorable impacts at all data collection points through the end of 3<sup>rd</sup> grade on parent-reported measures of children's social-emotional development. However, there were no impacts on teacher-reported measures of social-emotional development for the 3-year-old cohort at any data collection point or on the children's self-reports in 3<sup>rd</sup> grade.

The findings from the social-emotional domain are summarized by age cohort below. Exhibits 3a and 3b provide all statistically significant social-emotional impacts and their effect sizes from the ITT analysis.

### **4-Year-Old Cohort**

- There were no significant differences between the Head Start group and the control group on any measures of social-emotional development during the Head Start year or during kindergarten.
- At the end of 1<sup>st</sup> grade, impacts on social-emotional development were few and mixed.
  - There were two unfavorable findings based on teacher reports of children's behavior: (1) children in the Head Start group demonstrated moderate evidence of more socially reticent behavior (i.e., shy and hesitant behavior) as reported by teachers, and (2) there was suggestive evidence of more problematic student-teacher interactions.
  - In contrast, there was suggestive evidence of less withdrawn behavior for children in the Head Start group as reported by their parents.
- At the end of 3<sup>rd</sup> grade, parents reported less aggressive and total problem behaviors for the Head Start group children. However, teachers reported unfavorable impacts with a higher incidence of children's emotional symptoms, less closeness, and a less positive relationship with the Head Start children. Finally, Head Start children in the 4-year-old cohort reported less positive peer relations at school compared to the control group.

### **3-Year-Old Cohort**

- At the end of the Head Start year, children in the Head Start group showed strong evidence of less hyperactive behavior and fewer overall problem behaviors as reported by their parents.
- At the end of the age 4 year and the end of kindergarten, children in the Head Start group demonstrated suggestive evidence of better social skills and positive approaches to learning as reported by their parents. Further, children in the Head Start group also continued to show moderate evidence of less hyperactive behavior at the end of kindergarten.
- By the end of 1st grade, parents of Head Start group children reported moderate evidence of a closer relationship with their child than parents of control group children. At the same time, parents of Head Start group children reported (suggestive evidence) a more positive overall relationship with their child than parents of children in the control group.
- There were no impacts on teacher-reported measures of social-emotional development for the 3-year-old cohort in either the kindergarten or 1st grade year.
- For this age cohort, there was only a single statistically significant social-emotional impact at the end of 3rd grade. Children in the Head Start group demonstrated better social skills and positive approaches to learning as reported by their parents, compared with the non-Head Start group.

**Exhibit 3a. Summary of ITT Social-Emotional Impacts for 4-Year-Olds by Year**

Measure	Age 4 (Head Start Year)	K	1 <sup>st</sup> Grade	3 <sup>rd</sup> Grade
<b>Parent-Reported Measures</b>				
Aggressive Behavior				-0.13
Hyperactive Behavior				
Withdrawn Behavior			-0.13	
Total Problem Behavior				-0.12
Social Competencies				NA
Social Skills and Positive Approaches To Learning				
Closeness with Parent				NA
Conflict with Parent				NA
Positive Parent-Child Relationships				NA
<b>Teacher-Reported Measures</b>				
Aggressive (ASPI)	NA			NA
Interactive/Hyperactive (ASPI)	NA			
Withdrawn/Low Energy (ASPI)	NA			NA
Oppositional (ASPI)	NA			NA
Problems with Peer Interaction (ASPI)	NA			NA
Shy/Socially Reticent (ASPI)	NA		-0.19	NA
Problems with Structured Learning (ASPI)	NA			NA
Problems with Teacher Interaction (ASPI)	NA		-0.13	NA
Closeness with Teacher	NA			-0.13
Conflict with Teacher	NA			
Positive Teacher-Child Relationships	NA			-0.14
Conduct Problems-% in Normal Category	NA	NA	NA	
Emotional Symptoms-% in Normal Category	NA	NA	NA	-0.24
Hyperactivity-% in Normal Category	NA	NA	NA	
Peer Problems-% in Normal Category	NA	NA	NA	
Pro-social Behavior-% in Normal Category	NA	NA	NA	
Total Difficulties-% in Normal Category	NA	NA	NA	
Social Competency	NA	NA	NA	
<b>Child-Reported Measures</b>				
Externalizing	NA	NA	NA	
Internalizing	NA	NA	NA	
Peer Relations	NA	NA	NA	-0.14
School	NA	NA	NA	

**KEY:**

 Dark gray cell indicates a significant favorable impact ( $p \leq 0.10$ ).

 Diagonal mark cell indicates a significant unfavorable impact ( $p \leq 0.10$ ).

 Blank cell indicates no significant impact.

An \* indicates that the interpretation of the outcome is unclear.

NA indicates that no data were collected for this outcome at this data collection point.

NOTE: Intent to Treat (ITT) effect sizes are presented only for statistically significant differences ( $p \leq 0.10$ ). The effect size is simply the impact estimate divided by the standard deviation of the outcome measure in the population. The effect size provides an indication of the magnitude of each impact that is independent of the particular instrument or measure used. More discussion of the interpretation of effect sizes is provided in Chapter 2.

**Exhibit 3b. Summary of ITT Social-Emotional Impacts for 3-Year-Olds by Year**

Measure	Age 3 (Head Start Year)	Age 4	K	1 <sup>st</sup> Grade	3 <sup>rd</sup> Grade
<b>Parent-Reported Measures</b>					
Aggressive Behavior					
Hyperactive Behavior	-0.21		-0.12		
Withdrawn Behavior					
Total Problem Behavior	-0.14				
Social Competencies					NA
Social Skills and Positive Approaches To Learning		0.11	0.14		0.12
Closeness with Parent				0.10	NA
Conflict with Parent					NA
Positive Parent-Child Relationships				0.10	NA
<b>Teacher-Reported Measures</b>					
Aggressive (ASPI)	NA	NA			NA
Interactive/Hyperactive (ASPI)	NA	NA			
Withdrawn/Low Energy (ASPI)	NA	NA			NA
Oppositional (ASPI)	NA	NA			NA
Problems with Peer Interaction (ASPI)	NA	NA			NA
Shy/Socially Reticent (ASPI)	NA	NA			NA
Problems with Structured Learning (ASPI)	NA	NA			NA
Problems with Teacher Interaction (ASPI)	NA	NA			NA
Closeness with Teacher	NA	NA			
Conflict with Teacher	NA	NA			
Positive Teacher-Child Relationships	NA	NA			
Conduct Problems-% in Normal Category	NA	NA	NA	NA	
Emotional Symptoms-% in Normal Category	NA	NA	NA	NA	
Hyperactivity-% in Normal Category	NA	NA	NA	NA	
Peer Problems-% in Normal Category	NA	NA	NA	NA	
Pro-social Behavior-% in Normal Category	NA	NA	NA	NA	
Total Difficulties-% in Normal Category	NA	NA	NA	NA	
Social Competency	NA	NA	NA	NA	
<b>Child-Reported Measures</b>					
Externalizing	NA	NA	NA	NA	
Internalizing	NA	NA	NA	NA	
Peer Relations	NA	NA	NA	NA	
School	NA	NA	NA	NA	

**KEY:**

-  Dark gray cell indicates a significant favorable impact ( $p \leq 0.10$ ).
-  Diagonal mark cell indicates a significant unfavorable impact ( $p \leq 0.10$ ).
-  Blank cell indicates no significant impact.

An \* indicates that the interpretation of the outcome is unclear.

NA indicates that no data were collected for this outcome at this data collection point.

NOTE: Intent to Treat (ITT) effect sizes are presented only for statistically significant differences ( $p \leq 0.10$ ). The effect size is simply the impact estimate divided by the standard deviation of the outcome measure in the population. The effect size provides an indication of the magnitude of each impact that is independent of the particular instrument or measure used. More discussion of the interpretation of effect sizes is provided in Chapter 2.

## ***Impact on Health Status and Access to Health Services***

The health domain consisted of two categories: (1) children’s receipt of health care services and (2) their current health status. Early favorable impacts in the health domain were noted for both age cohorts but by the end of 3<sup>rd</sup> grade, there were no remaining impacts for either age cohort.

The findings from the health domain are summarized by age cohort below, while Exhibits 4a and 4b present all statistically significant health impacts and their effect sizes from the ITT analysis.

### **4-Year-Old Cohort**

- At the end of the Head Start year, there was strong evidence that access to Head Start increased children’s receipt of dental care—a difference of 15 percentage points.
- In kindergarten, there was suggestive evidence of an improvement in children’s health status and an increase in health insurance coverage (differences of five and four percentage points, respectively).
- By the end of 1<sup>st</sup> grade, there was still moderate evidence of increased health insurance coverage among the Head Start group—a difference of four percentage points.
- There were no significant impacts at the end of 3<sup>rd</sup> grade.

### **3-Year-Old Cohort**

- At the end of the Head Start year and again at the end of the age 4 year, there was strong evidence that access to Head Start increased children’s receipt of dental care—differences of 17 and 10 percentage points, respectively.
- There was moderate evidence of improvements on children’s reported overall health status at the end of the Head Start year and moderate evidence of an impact on health insurance coverage at the end of kindergarten.
- There was evidence of a significant impact on care for injuries<sup>9</sup> at the end of the age 4 year, although the interpretation of this impact is unclear.
- There were no significant impacts at the end of 1<sup>st</sup> or 3<sup>rd</sup> grades.

---

<sup>9</sup> The interpretation of child had care for injury in the last month is unclear. The change may reflect an increase in injuries, an increase in care-seeking, or both.

### Exhibit 4a. Summary of ITT Health Impacts for 4-Year-Olds by Year

Measure	Age 4 (Head Start Year)	K	1 <sup>st</sup> Grade	3 <sup>rd</sup> Grade
<b>Parent-Reported Measures</b>				
Child Received Dental Care	0.31			
Child Has Health Insurance Coverage		0.11	0.11	
Child's Overall Health Status is Excellent/ Good		0.13		
Child Needs Ongoing Health Care				
Child Had Care for Injury in Last Month				

**KEY:**

-  Dark gray cell indicates a significant favorable impact ( $p \leq 0.10$ ).
-  Diagonal mark cell indicates a significant unfavorable impact ( $p \leq 0.10$ ).
-  Blank cell indicates no significant impact.

An \* indicates that the interpretation of the outcome is unclear.

NA indicates that no data were collected for this outcome at this data collection point.

NOTE: Intent to Treat (ITT) effect sizes are presented only for statistically significant differences ( $p \leq 0.10$ ). The effect size is simply the impact estimate divided by the standard deviation of the outcome measure in the population. The effect size provides an indication of the magnitude of each impact that is independent of the particular instrument or measure used. More discussion of the interpretation of effect sizes is provided in Chapter 2.

### Exhibit 4b. Summary of ITT Health Impacts for 3-Year-Olds by Year

Measure	Age 3 (Head Start Year)	Age 4	K	1 <sup>st</sup> Grade	3 <sup>rd</sup> Grade
<b>Parent-Reported Measures</b>					
Child Received Dental Care	0.33	0.20			
Child Has Health Insurance Coverage			0.14		
Child's Overall Health Status is Excellent/ Good	0.11				
Child Needs Ongoing Health Care					
Child Had Care for Injury in Last Month <sup>10</sup>		0.10*			

**KEY:**

-  Dark gray cell indicates a significant favorable impact ( $p \leq 0.10$ ).
-  Diagonal mark cell indicates a significant unfavorable impact ( $p \leq 0.10$ ).
-  Blank cell indicates no significant impact.

An \* indicates that the interpretation of the outcome is unclear.

NA indicates that no data were collected for this outcome at this data collection point.

NOTE: Intent to Treat (ITT) effect sizes are presented only for statistically significant differences ( $p \leq 0.10$ ). The effect size is simply the impact estimate divided by the standard deviation of the outcome measure in the population. The effect size provides an indication of the magnitude of each impact that is independent of the particular instrument or measure used. More discussion of the interpretation of effect sizes is provided in Chapter 2.

<sup>10</sup> See footnote 9.

## ***Impact on Parenting Practices***

This domain consisted of six categories of outcomes: (1) disciplinary practices, (2) educational supports, (3) safety practices, (4) parenting styles, (5) parent participation in and communication with school and (6) parent and child time together. With regard to parenting practices, the impacts were concentrated in the younger cohort, which showed favorable parent-reported impacts across all years of the study. For the 4-year-old cohort, in contrast, there were few impacts.

The findings from the parenting practices domain are summarized by age cohort below, and Exhibits 5a and 5b provide the statistically significant parenting practices impacts and their effect sizes from the ITT analysis.

### **4-Year-Old Cohort**

- There were minimal impacts for the 4-year-old cohort in this domain, with two exceptions: at the end of the Head Start year, parents in the Head Start group were less likely to use time out<sup>11</sup> as a disciplinary practice than were parents in the control group and at the end of 3<sup>rd</sup> grade, there was strong evidence of a large favorable impact on parental reports of the amount of time they spent with their child.

### **3-Year-Old Cohort**

- In the Head Start year, there were three impacts on parenting practices, of which two impacts (spanking and cultural enrichment) were supported by strong evidence:
  - Parents of children in the Head Start group were less likely to have spanked their children than parents in the control group (a difference of seven percentage points).
  - Parents of children in the Head Start group were more likely to have read to their child in the last week than parents in the control group.
  - Parents of children in the Head Start group involved their child in cultural enrichment activities more than parents of children in the control group.
- At the end of the age 4 year, there was a favorable impact on parenting, with parents of children in the Head Start group less likely to use an authoritarian parenting style (characterized by high control and low warmth) than parents of children in the control group.
- Evidence of impacts on parenting continued in kindergarten, 1<sup>st</sup> and 3<sup>rd</sup> grades.
  - At the end of kindergarten, there was suggestive evidence that parents of children in the Head Start group were less likely to spank their children and moderate evidence that these parents were less likely to use time out.<sup>12</sup>

---

<sup>11</sup> The interpretation of time out is unclear. The change may reflect favorable changes in the children's behavior, changes in the parents' reactions (whether to less or more desirable forms of discipline), or both.

<sup>12</sup> See footnote 11.

- At the end of 1<sup>st</sup> grade, there was also suggestive evidence that parents of children in the Head Start group were less likely to use time out and moderate evidence that these parents were less likely to use an authoritarian parenting style.
- At the end of 3<sup>rd</sup> grade, there was a favorable impact on the use of the preferred authoritative parenting style (characterized by high warmth and high control).

**Exhibit 5a. Summary of ITT Parenting Practices Impacts for 4-Year-Olds by Year**

Measure	Age 4 (Head Start Year)	K	1 <sup>st</sup> Grade	3 <sup>rd</sup> Grade
<b>Parent-Reported Measures</b>				
Parent Spanked Child in Last Week				
Parent Used Time Out in Last Week <sup>13</sup>	-0.17*			NA
Parent Read to Child in Last Week				NA
Parental Safety Practices Scale				NA
Family Cultural Enrichment Scale				NA
Parenting Style: Authoritarian	NA			
Parenting Style: Authoritative	NA			
Parenting Style: Neglectful	NA			
Parenting Style: Permissive	NA			
Supportive School Environment	NA	NA	NA	
Effect of Parenting on Parent's Life	NA	NA	NA	
Doing Things Together	NA	NA	NA	
Time Spent with Child	NA	NA	NA	0.27
Parent Perception of School Services	NA	NA	NA	
<b>Teacher-Reported Measures</b>				
School Contact and Communication	NA			
Parent Participation	NA			

**KEY:**

-  Dark gray cell indicates a significant favorable impact ( $p \leq 0.10$ ).
-  Diagonal mark cell indicates a significant unfavorable impact ( $p \leq 0.10$ ).
-  Blank cell indicates no significant impact.

An \* indicates that the interpretation of the outcome is unclear.

NA indicates that no data were collected for this outcome at this data collection point.

NOTE: Intent to Treat (ITT) effect sizes are presented only for statistically significant differences ( $p \leq 0.10$ ). The effect size is simply the impact estimate divided by the standard deviation of the outcome measure in the population. The effect size provides an indication of the magnitude of each impact that is independent of the particular instrument or measure used. More discussion of the interpretation of effect sizes is provided in Chapter 2.

<sup>13</sup> See footnote 11.

**Exhibit 5b. Summary of ITT Parenting Practices Impacts for 3-Year-Olds by Year**

Measure	Age 3 (Head Start Year)	Age 4	K	1 <sup>st</sup> Grade	3 <sup>rd</sup> Grade
<b>Parent-Reported Measures</b>					
Parent Spanked Child in Last Week	-0.14		-0.09		
Parent Used Time Out in Last Week <sup>14</sup>			-0.13*	-0.11*	NA
Parent Read to Child in Last Week	0.15				NA
Parental Safety Practices Scale					NA
Family Cultural Enrichment Scale	0.18				NA
Parenting Style: Authoritarian	NA	-0.14		-0.11	
Parenting Style: Authoritative	NA				0.16
Parenting Style: Neglectful	NA				
Parenting Style: Permissive	NA				
Supportive School Environment	NA	NA	NA	NA	
Effect of Parenting on Parent's Life	NA	NA	NA	NA	
Doing Things Together	NA	NA	NA	NA	
Time Spent with Child	NA	NA	NA	NA	
Parent Perception of School Services	NA	NA	NA	NA	
<b>Teacher-Reported Measures</b>					
School Contact and Communication	NA	NA			
Parent Participation	NA	NA			

**KEY:**

-  Dark gray cell indicates a significant favorable impact ( $p \leq 0.10$ ).
-  Diagonal mark cell indicates a significant unfavorable impact ( $p \leq 0.10$ ).
-  Blank cell indicates no significant impact.

An \* indicates that the interpretation of the outcome is unclear.

NA indicates that no data were collected for this outcome at this data collection point.

NOTE: Intent to Treat (ITT) effect sizes are presented only for statistically significant differences ( $p \leq 0.10$ ). The effect size is simply the impact estimate divided by the standard deviation of the outcome measure in the population. The effect size provides an indication of the magnitude of each impact that is independent of the particular instrument or measure used. More discussion of the interpretation of effect sizes is provided in Chapter 2.

## Variation in Impact

This report examines differences in impact among different groups of children and parents. Seven dimensions were used to define subgroups: (1) whether a child had low pre-academic skills, (2) whether the child was a Dual Language Learner, (3) whether the child had special needs (as reported by the parent), (4) mother's race/ethnicity, (5) reported level of depressive symptoms for the child's parent/primary caregiver, (6) a composite index of household risks, and (7) urban location. All categorizations were based on data collected at the time of random assignment. Sample sizes by subgroup, age cohort, and random assignment status are presented in Chapter 5.

<sup>14</sup> See footnote 11.



The approach to analyzing subgroups was to highlight patterns in the findings. There is no scientific consensus for what constitutes a pattern of impacts. Yet, given the large number of comparisons tested (over 13,000, taking into consideration the study's two cohorts, five time points for measuring outcomes, and multiple outcomes across many subgroups), it was important to find an approach that balances the risk of reporting on chance findings with that of ignoring important findings. To this end, the subgroup findings concentrate on differential impacts, that is, impacts where there was a statistically significant difference in Head Start's effects for one subgroup compared to another. Accordingly, the discussion primarily focuses on results where there was both a statistically significant difference in impacts between subgroups and a statistically significant impact for at least one subgroup in the comparison.

Particular attention was paid to end of 3<sup>rd</sup> grade results that showed a pattern across domains and how those results related to prior time points.<sup>15</sup> The subgroup findings should be viewed as secondary and exploratory as compared to the main impact findings that are considered primary as well as confirmatory.

At the end of 3<sup>rd</sup> grade, the most striking sustained subgroup finding was related to children from high risk households. For this subgroup, children in the 3-year old cohort demonstrated sustained cognitive impacts across all the years from pre-K through 3<sup>rd</sup> grade. At the end of 3<sup>rd</sup> grade, the Head Start children from high risk households showed favorable impacts on the ECLS-K Reading Assessment, the WJIII Letter-Word Identification, and the teacher-reported reading/language arts skills. This was in contrast to the impacts for children in lower and moderate risk households, for whom there were no impacts. Those children who started out with more familial stressors than their peers were found to have multiple positive impacts on the direct student assessments over time. Also, among the 3-year-old cohort, children of parents with no reported depressive symptoms experienced sustained benefits of Head Start in the cognitive domain through the end of 3<sup>rd</sup> grade and in the social-emotional and parenting practices domain through the end of 1<sup>st</sup> grade.

Among the 4-year-olds, the subgroups that demonstrated sustained benefits are children of parents who reported mild depressive symptoms, severe depressive symptoms, and Black children. Head Start children of parents reporting mild depressive symptoms demonstrated favorable cognitive impacts through the end of 3<sup>rd</sup> grade. This was in contrast to those with no, moderate, or severe depressive symptoms. However, favorable impacts were reported only at the end of the Head Start year for parents with severe depressive symptoms. In the parenting and social-emotional domains, predominantly favorable parent-reported impacts were sustained for children of parents with severe depressive symptoms. Black children experienced favorable impacts in the social-emotional domain at the end of kindergarten through 3<sup>rd</sup> grade as reported by teachers, parents, and the child self-report.

Finally, several subgroups experienced solely-or primarily-unfavorable impacts of Head Start that were sustained through 3<sup>rd</sup> grade. For the 4-year-old cohort, this included White children, who had unfavorable impacts in the social-emotional domain, and for the 3-year-old cohort, children of parents with mild depressive symptoms, who also had unfavorable impacts in the social-emotional domain. Many subgroups in both age cohorts experienced a mixture of favorable and unfavorable impacts, particularly in the social-emotional domain.

---

<sup>15</sup> The Benjamini-Hochberg test of multiple comparisons was also applied to the subgroup analysis, and the results are included in the Chapter 8 tables of this report.

Some of the subgroup impacts from earlier years were not sustained through 3<sup>rd</sup> grade. For example, the favorable social-emotional impacts for children in the 4-year-old cohort of parents with mild depressive symptoms and the favorable cognitive impacts found at the end of 1<sup>st</sup> grade for children in the 3-year-old cohort from non-urban settings were not sustained through 3<sup>rd</sup> grade.

## ***Final Thoughts***

Head Start has the ambitious mandate of improving educational and developmental outcomes for children from economically disadvantaged families. Head Start's mandate requires that it meet the needs of the whole child, including the cognitive, social-emotional, and health needs of children, and positively influence the parenting practices of their parents. This study examines the impacts of Head Start on these four domains and whether earlier impacts were sustained into 3<sup>rd</sup> grade.

The lasting effects of Head Start and early childhood education in general on children's outcomes have been the focus of much study. Considering only outcomes through early elementary school and middle childhood, results for the HSIS cognitive outcomes are in line with other experimental and non-experimental early education studies. Non-experimental Head Start studies showed initial positive impacts of a roughly similar magnitude to those found in the HSIS that dissipated as the children entered early elementary school (Currie & Thomas, 1995; Garces, et al., 2002; Ludwig & Phillips 2008; Deming 2009). Moreover, recent longitudinal data from the experimental evaluation of Early Head Start (Vogel, et al., 2010) showed a similar pattern of early positive impacts that were not sustained into elementary school. Experimental results from the HighScope Preschool Curriculum Comparison study found negligible differences between study groups in cognitive and academic outcomes in the first decade of study (Schweinhart & Weikart, 1997). Similar conclusions about the size and lack of persistence of early impacts were reported in a recent broader meta-analysis of early childhood interventions (Leak et al., 2010). However, as we discuss later, some studies, including those that did not show differences in elementary school, reported finding positive effects later in adulthood. Although the underlying cause of the rapid attenuation of early impacts is an area of frequent speculation, we don't have a good understanding of this observed pattern. All we can say is after the initially realized cognitive benefits for the Head Start children, these gains were quickly made up by children in the non-Head Start group.

We do not yet know if there will be positive outcomes for HSIS participants later in life, however, research suggests that positive outcomes later in life are possible. Despite a growing body of research about relatively rapid dissipation of early cognitive impacts, there is some evidence suggesting that positive effects of Head Start may have an impact on participants' later life such as later school success and early adulthood outcomes (Garces, et al., 2002; Ludwig & Miller, 2007; Deming, 2009). Garces, Thomas, and Currie (2000) conducted a non-experimental study that reported evidence of long-term improvement for Head Start participants on outcomes such as school attainment, earnings and crime reduction, for some race and gender combinations. Ludwig and Miller (2007), using a regression discontinuity design, reported that increases in Head Start funding were associated with a decline in mortality rates for children ages five to nine from causes of death that could be affected by the program, an increase in high school completion, and an increase in the likelihood of attending some college. Both of these

findings were based on Head Start programs that operated in the 1960's through the 1980's. More recently, Gelber and Isen (2011), using the HSIS data, reported that parents of children assigned to Head Start were more involved with them in a variety of activities both during Head Start enrollment and the early elementary years. The authors suggested that increases in parent involvement may mediate long-term impacts on child outcomes. According to a recent paper by Gibbs, Ludwig, & Miller (2011) such delayed or "sleeper" effects may occur because of the Head Start benefits in the area of children's social and emotional development, i.e., improved socialization and emotional strength may have later school-related payoffs.

Research from non-Head Start samples with similar populations also suggests that "sleeper" effects may present years after exposure to early education. Using data from the randomized study, Project Star (1985-89 Tennessee K-3 Class Size Study), Chetty et al. (2010) reported that improvements in kindergarten test scores translate into higher lifetime earnings, more likely college attendance, retirement savings, home ownership, and residence in a better neighborhood. Children from the HighScope group completed more years of school, had less self-reported misconduct at age 15, fewer felony arrests, and fewer property crime arrests than those who received direct instruction. Initially, no early academic differences were found but the long term impacts suggested benefits from quality early childhood education in early adulthood outcomes. Although Project Star and the HighScope Preschool Curriculum Comparison study were not focused on Head Start, like the other Head Start studies, they point to the importance of early education for improving children's long-term outcomes.

In addition to considering the possible long-term impacts, there are a few other things to consider in interpreting the findings of this study. First, this was not a comparison of Head Start to parental care. This study evaluated the Head Start program as it existed in 2002 against a mixture of alternative care settings rather than against a "no services" condition. About 40 percent of the control group did not receive formal preschool education and, for those who did, quality was generally lower than in Head Start. Nevertheless, many of the control group children received services. Further, among those who participated in non-parental care, the control group children were actually in non-parental care for more hours than the Head Start group—on average, children in the control group attended some type of non-parental care about four to five hours more per week in the Head Start year, compared with children who had access to Head Start. Consequently, to achieve measurable impacts, Head Start (as noted above) had to outperform what control group children received.

Additionally, to date the findings do not differentiate impacts for children who received services of differing quality in Head Start. Although the quality was high on average, Head Start programs varied in terms of academic instruction in the key areas measured as part of this study, i.e., early development of language and literacy and mathematics skills. This is not to say that all Head Start programs were not trying their best to improve children's development in these areas, but rather on average the program may not have been potent enough in this particular domain to provide the level of overall learning gains needed to move children into a different, and more rapid, growth trajectory. The pattern for the HSIS data showed initial accelerated gains for the Head Start children, then these gains were quickly made up by the control group children, followed by continued gains at the same pace for both groups. The variation in quality may have contributed to the lack of statistically significant differences in the cognitive domain in the early elementary grades.

A separate report will explore how variation in Head Start quality is related to children's impacts as well as how children's later experiences in the school and community affect their outcomes at 1<sup>st</sup> and 3<sup>rd</sup> grade, including whether some later experiences help to sustain impacts through the early elementary grades.

The study also reflects on the impact of Head Start as it existed in 2002, and does not necessarily represent either the impact of Head Start between the time it was initiated and 2002 or the impact of any changes made to Head Start since 2002. As most evaluations, this study is designed to ask a set of questions about a program at a particular point in time. To the extent that the program has changed since the time that study participants were given access, those changes will not be reflected in the study's findings.

Finally, this study leaves many important questions about Head Start unanswered. These questions include, but are certainly not limited to: Is there a benefit to having two years of Head Start rather than one year? What accounts for the subgroup patterns observed in this report? The Head Start Impact Study is an excellent data base for methodological and child development research due to its size, longitudinal data, and multiple variables. Hopefully, researchers will take advantage of the data from this study, which will be made available through a data archive,<sup>16</sup> to further the understanding of the role Head Start plays in the well-being of children and families.

---

<sup>16</sup> The data is archived at the Child Care & Early Education Research Connections Project. <http://www.childcareresearch.org/childcare/welcome>