Head Start Impact Study Final Report

January 2010

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

Head Start Impact Study

Final Report

January 2010

Prepared for:

Office of Planning, Research and Evaluation
Administration for Children and Families
U.S. Department of Health and Human Services
Washington, D.C.
under contract 282-00-0022, Head Start Impact Study

Prepared by:

Westat
1600 Research Blvd.
Rockville, MD 20850

Chesapeake Research Associates
708 Riverview Terrace
Annapolis, MD 21401

Abt Associates
4550 Montgomery Avenue
Bethesda, MD 20814

Ronna Cook Associates
5912 Rossmore Drive
Bethesda, MD 20814

The Urban Institute
2100 M Street, N.W.
Washington, DC 20037

American Institutes for Research
1000 Thomas Jefferson Street, N.W.
Washington, DC 20007

Decision Information Resources, Inc.
2600 Southwest Freeway, Suite 900
Houston, TX 77098

Head Start Impact Study
Final Report

Authors

Prepared by:

Michael Puma
Stephen Bell
Ronna Cook
Camilla Heid

Contributing Authors:

Gary Shapiro
Pam Broene
Frank Jenkins
Philip Fletcher
Liz Quinn
Janet Friedman
Janet Ciarico
Monica Rohacek
Gina Adams
Elizabeth Spier

Suggested Citation:

U.S. Department of Health and Human Services,
Administration for Children and Families (January 2010).
Head Start Impact Study. Final Report.
Washington, DC.

Disclaimer

The Office of Planning, Research and Evaluation, Administration for Children and Families at the U.S. Department of Health and Human Services contracted with Westat to conduct the Head Start Impact Study. The views expressed in this report are those of the authors and they do not necessarily represent the opinions and positions of the Office of Planning, Research and Evaluation, Administration for Children and Families or the U.S. Department of Health and Human Services.

Acknowledgements

This final report of the Head Start Impact Study is the result of several years of design, data collection, and analysis. We gratefully acknowledge the contributions and dedication of individuals and organizations in the preparation and production of this report. A special thanks to Dr. Jennifer Brooks, the Federal Project Officer, for her expert leadership and vision.

There were those who were worried that random assignment and subsequent data collection efforts would be difficult, if not impossible to implement. Study staff have done a tremendous job in meeting these challenges to ensure the success of the study. Moreover, the partnership and support from the National Head Start Association, Head Start Grantees and Delegate Agencies and their center staff, as well as the study children's elementary schools and their staff were instrumental in the successful implementation of this study. The ongoing backing of the Head Start Bureau and Regional Office staff was critical to the recruitment process. A special thank you is extended to all the families and their children who participated in the study. Their continued contributions of time and information during the data collection years have been exceptional and greatly appreciated.

We also want to thank the many external experts who helped us along the way, particularly the members of the Advisory Committee on Head Start Research and Evaluation. Your wisdom about sample design, measures, program, policy, and analytic challenges has helped formulate the design and analysis presented in the report.

Finally, we gratefully acknowledge the staff from Westat, Chesapeake Research Associates, Abt Associates, Ronna Cook Associates, Urban Institute, and American Institutes for Research for their hard work, professionalism and dedication to the project. We also wish to thank Decision Information Resources, Inc for their assistance in the data collection.

Table of Contents

Contents (continued)

Contents (continued)

Contents (continued)

Contents (continued)

Contents (continued)

Exhibits (continued) Page

Contents (continued)

Contents (continued)

Contents (continued)

Contents (continued)

Exhibits (continued) Page

Contents (continued)

Contents (continued)

Head Start Impact Study

Executive Summary

Overall 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:[1]

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?"

This report addresses these questions by reporting on the impacts of Head Start on children and families during the children's preschool, kindergarten, and 1st grade years.

The Head Start Impact Study was conducted with a nationally representative sample of 84 grantee/delegate agencies and included nearly 5,000 newly entering, eligible 3- and 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 2006, following children from program application through the spring of their 1st grade year.[2]

---

[1] Advisory Committee on Head Start Research and Evaluation (1999). *Evaluating Head Start: A Recommended Framework for Studying the Impact of the Head Start Program.* Washington, DC: US Department of Health and Human Services.

[2] The study design allowed 3-year-old cohort control group children to reapply to Head Start after the first 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 (to be studied through two years of Head Start participation i.e., Head Start year and age 4 year, kindergarten and 1st grade), and a newly entering 4-year-old group (to be studied through one year of Head Start participation, kindergarten and 1st grade).

The study showed that the two age cohorts varied in demographic characteristics, making it even more appropriate to examine them separately. 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 (Black children 32.8%, Hispanic children 37.4%, and White/other children 29.8%), about half of newly entering 4-year-olds were Hispanic children (Black children 17.5%, Hispanic children 51.6%, and White/other children 30.8%). The ethnic difference is also reflected in the age-group differences in child and parent language.

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 contrast, the Head Start Impact Study is based on a nationally representative sample of Head Start programs and children. This makes the study results generalizable to the full national program 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 (i.e., the best, the worst, and those in the middle of a fully implemented program) and adherence to the established Head Start program standards.
- **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. These impacts are quantified 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:

- **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, i.e., how well does Head Start do when compared against what else low income children could receive in the absence of the program 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. 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 reasonable 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.[3] 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. Therefore, the research design for the 3-year-old cohort only varied the first year of Head Start participation. Hence, impacts for the 3-year-old cohort reflect the benefits of being provided an earlier year of Head Start, rather than the effects of being provided two years of Head Start. By design, the study did not attempt to control children's experiences after they received this Head Start year.

The Head Start Impact Study is a comprehensive, well-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 program. The findings cannot be directly compared to more narrowly focused studies of other early childhood programs. The Advisory Committee on Head

[3] It was not feasible or reasonable to prevent 3-year-olds from participating in Head Start for two years. Thus, the study does not directly assess the receipt of one-year versus two years of Head Start. It addresses the receipt of an earlier year.

Start Research and Evaluation, who 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).

Key Findings

The key findings are presented below. First, we present findings related to the primary questions about the average effect of Head Start as a whole. Next we present findings about subgroups of children. As described later in this summary, 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.

Confirmatory Impact Findings

- Providing access to Head Start has a positive impact on children's preschool experiences. There are statistically significant differences between the Head Start group and the control group on every measure of children's preschool experiences measured in this study.
- Access to Head Start has positive impacts on several aspects of children's school readiness during their time in the program.
  - For the 4-year-old group, benefits at the end of the Head Start year were concentrated in language and literacy elements of the cognitive domain, including impacts on vocabulary (PPVT), letter-word identification, spelling, pre-academic skills, color identification, letter naming, and parent-reported emergent literacy. There was also an impact on access to dental care in the health domain.
  - For the 3-year-old group, benefits were found in all four domains examined at the end of the Head Start and age 4 years, including impacts on vocabulary (PPVT), letter-word identification, pre-academic skills, letter naming, elision (phonological processing), parent-reported emergent literacy, McCarthy Draw-a-Design (perceptual motor skills and pre-writing), applied problems (math), hyperactive behavior, withdrawn behavior, dental care, health status, parent spanking, parent reading to child, and family cultural enrichment activities.
- However, the advantages children gained during their Head Start and age 4 years yielded only a few statistically significant differences in outcomes at the end of 1st grade for the sample as a whole. Impacts at the end of kindergarten were scattered and are mentioned below only when they appear to be related to the 1st grade impacts.
  - Cognitive Outcomes. By the end of 1st grade, only a single cognitive impact was found for each cohort. Head Start group children did significantly better on the PPVT (a vocabulary measure) for 4-year-olds and on the Woodcock-Johnson III test of Oral Comprehension for the 3-year-olds.
  - Social-Emotional Outcomes. By the end of 1st grade, there was some evidence that the 3-year-old cohort had closer and more positive relationships with their parents. These impacts were preceded by other social-emotional impacts (improvements in behavior-hyperactive behavior and total problem behavior, and social skills and positive approaches to learning) in the earlier years. The findings

for the 4-year-old cohort are inconsistent with teachers reporting that children in the Head Start group are more shy and socially reticent and have more problems with student and teacher interactions than control group children while their parents are reporting that they are less withdrawn.

- Health Outcomes. For the 4-year-old cohort, there was an impact on child health insurance coverage at the end of kindergarten and 1st grade, and an impact on child health status in kindergarten. For the 3-year-olds, there was an impact on child health insurance coverage in kindergarten only.
- Parenting Outcomes. For the 3-year-old cohort, there were positive favorable impacts on use of time-out and authoritarian parenting at the end of 1st grade and on spanking and time out in kindergarten. These favorable impacts for authoritarian parenting and spanking were also demonstrated in the earlier years. For the 4-year-old cohort, there were no significant parenting practices impacts in kindergarten or 1st grade.

Exploratory Subgroup Findings

- Selected subgroups of children showed patterns of favorable impacts, including favorable impacts through 1st grade in the cognitive, social-emotional, or health domains.
  - Among the 4-year-old cohort, these subgroups include children of parents with mild depressive symptoms, children who were Dual Language Learners, and children with lower cognitive skills. Additionally, Black children experienced favorable impacts in the social-emotional domain at the end of kindergarten.
  - Among the 3-year-old cohort, the subgroups showing favorable impacts include children with special needs, children of parents with no depressive symptoms, children from higher risk households, and children in non-urban settings. In the 3-year-old cohort, there were also several groups with more favorable impacts during the earlier years of the study: these groups included children with lower cognitive skills upon entering Head Start and Dual Language Learners.
- There were also a few subgroups of children that showed patterns of unfavorable impacts. The group that showed the most widespread unfavorable impacts was 3-year-olds whose parents reported moderate depressive symptoms. These children experienced negative impacts across the cognitive, social-emotional, and health domains.

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 in the initial 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

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, if properly implemented, would ensure that the two 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 three, *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 shortage of 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 2002-03 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 2002-03 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 Head Start participation i.e., Head Start year and age 4 year, kindergarten, and 1st grade) and newly entering 4-year-olds (to be studied through one year of Head Start participation, kindergarten, and 1st 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.[4]

Data collection began in the fall of 2002 and continued through the spring of 2006, following children from entry into Head Start through the end of 1st grade. Comparable data were collected for both Head Start and control group children, including interviews with parents, direct child assessments, surveys of Head Start and non-Head Start teachers, interviews with center directors and other care providers, direct observations of the quality of various care settings, and care provider assessments of children. Response rates were consistently quite high, approximately 80 percent for parents and children throughout the study.

Although every effort was made to ensure complete 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 and the impact of actual Head Start participation (adjusting for the no shows and crossovers).

Key Findings

Impact on Children's Experiences

Head Start Experiences

Providing access to Head Start increases the likelihood that low-income children will be enrolled in a center-based early childhood program (including center-based Head Start, preschool, and child care). Specifically, Head Start group children[5] were twice as likely as control group children to use a center-based program in spring 2003.

Conversely, control group children were substantially more likely than Head Start group children to be exclusively in parent care[6] in spring 2003. Among children in the 3-year-old cohort, 38.4 percent of control group children were in parent care as compared to only 6.7 percent of children in the Head Start group; among children in the 4-year-old cohort, the figures were 39.7 and 9.1 percent, respectively (see Exhibit 1).

---

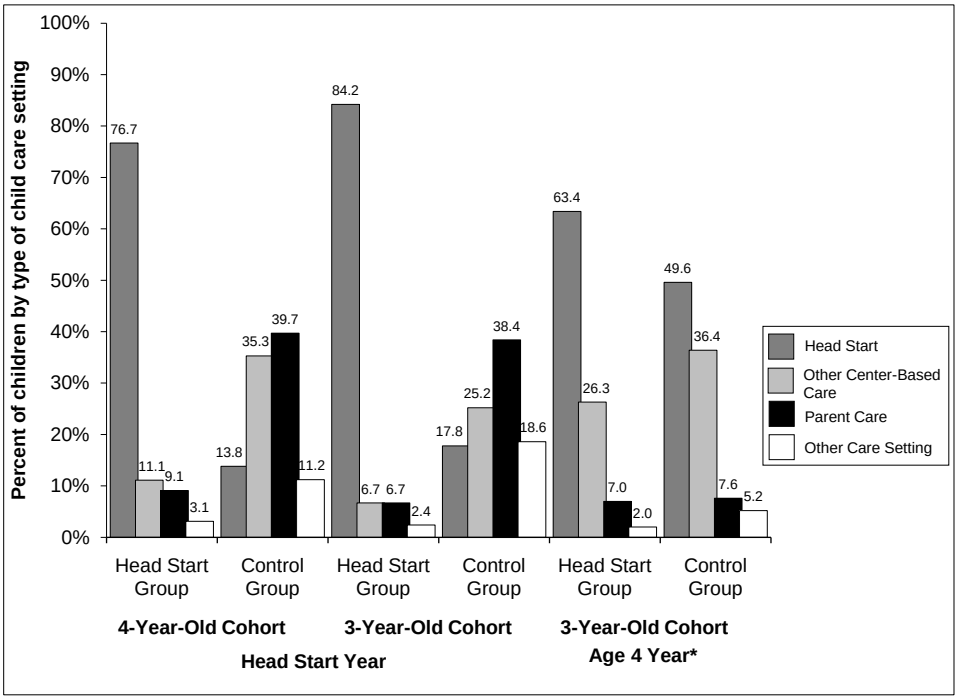
[4] The sample of 3-year-olds is slightly larger than the sample of 4-year-olds to ensure that an adequate sample size was maintained, given the possibility of higher study attrition resulting from an additional year of longitudinal data collection for the younger children.

[5] The Head Start group refers to children who were randomly assigned to have access to Head Start.

[6] Exclusively in parent care is defined as not being in any other non-parental setting for at least five hours per week.

During the second year of the study, the control group was given access to Head Start; thus the pattern of the 3-year-old cohort's second year experiences was very different from that in their first year. At the end of the second year, about 90 percent of the Head Start group was in a center-based early childhood program (primarily Head Start, 63 percent). At this point, a comparable percentage of the control group was also in a center-based program, with about 50 percent of those children in Head Start (see Exhibit 1).

Exhibit 1: Child Care Settings Used by Head Start and Control Groups During the Head Start Year, Spring 2003, and 3-Year-Old Cohort's Age 4 Year, Spring 2004



*For approximately 1 percent of respondents, setting type unknown.

There also was variation in the number of hours per week spent in Head Start as compared to other non-parental care for Head Start group and control group children. For those attending Head Start, the average number of hours spent per week was between 24 and 28 hours, with some variation by cohort and year. As discussed earlier, some control group children did receive Head Start services. Those control group children who found their way into Head Start experienced the same number of hours of Head Start as their program group counterparts. Non-parental care settings include Head Start, other center-based care, and home-based care. When averaging across all three types of non-parental care settings, control group children tended to be in non-parental care settings more hours per week than Head Start group children (Exhibit 2). For both the 3- and 4-year-old cohorts' Head Start year, control group children spent four to five more hours per week in their primary non-parental care setting than Head Start group children. The number decreases to only two more hours for the 3-year-old cohort's age 4 year.

Exhibit 2: Average Hours Per Week that the 3- and 4-Year-Old Cohorts Spent in Non-Parental Care Settings

	Hours Per Week					
	4-Year-Olds		3-Year-Olds Head Start Year		3-Year-Olds Age 4 Year	
	Head Start Group	Control Group	Head Start Group	Control Group	Head Start Group	Control Group
Overall Average Across All Non-Parental Care Settings	25	29	28	33	27	29

Providing access to Head Start had a positive impact on children's exposure to high-quality early care and education environments. Specifically, there are statistically significant differences between the Head Start group and the control group on every aspect of children's early care 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 spring of the first year of study. The measures that were examined included, among others, 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 the Early Childhood Environment Rating Scale-Revised (ECERS-R) and the Family Day Care Rating Scale (FDCRS) scores.[7]

The preschool experiences of children in the 3-year-old cohort were very different in the second preschool year. As discussed above, most of the children (both Head Start and control group) were in some type of center-based care by the second year. There were only three statistically significant differences across all the measures examined: (1) children in the Head Start group were less likely to be in a center that was affiliated with a school; (2) children in the Head Start group were more likely to have a teacher with a Child Development Associate (CDA) degree; and (3) children in the Head Start group were more likely to have hearing and vision screening referral services.

Of those 3-year-olds that attended a first year of Head Start, about 72 percent returned to Head Start for a second year. Characteristics related to an increased likelihood of returning for a second year included less competition from other early childhood programs in the area, centers with only full-day classrooms, parental satisfaction with how the center supported and respected their family's culture and background, coming from a household in which the home language was Spanish, or having a mother who was a recent immigrant.

While on average having access to Head Start resulted in more positive experiences for children, not all children in the Head Start group had the same quality of experience. The experiences of children and the services they received varied. The majority (70 percent) of children in the Head Start group in both cohorts were in centers with overall average ECERS-R scores of at least a five on a seven-point scale, indicating a good or better quality environment. Most children were also in classrooms that emphasized language and literacy and math

[7] These analyses compare the treatment and control groups, regardless of children's preschool placements. Chapter 3 provides an in depth description of the types of programs children were exposed to as well as participated in prior to kindergarten.

activities—approximately 60 percent were in classrooms that provided an emphasis on language and literacy and math activities (in which teachers reported providing more than half of a list of 12 language and literacy activities and eight math activities at least three times per week).

Conversely, the remaining Head Start group children did not experience centers with such high ECERS-R scores or as strong an emphasis on language and literacy or math activities. There is also diversity in the training and qualifications of the teachers serving the Head Start group children. Approximately 30 percent of the Head Start group children had teachers with a BA degree, and 30 percent had teachers with at least an Associate's degree, leaving approximately 40 percent of the children with teachers without a postsecondary degree. Slightly more than one-third of the 3-year-old cohort, and 40 percent of the 4-year-old cohort assigned to the Head Start group had teachers who had received 25 hours or more of training in the last year.

Experiences in Kindergarten and 1st Grade

This study collected data on children's elementary schools from secondary sources, teacher report, and parents. A subsequent examination of children through the end of 3rd grade will shed greater light on school experiences. However, based on the information that was collected, the results show that providing access to Head Start did not appear to have an impact on the types of schools children attended through 1st grade.

Few significant differences were found between the teachers of the Head Start and control group children for any of the teacher qualification measures (certification, educational attainment, educational preparation, and tenure) or on measures of job satisfaction in either kindergarten or 1st grade. There were also few significant differences on measures of teacher beliefs on how children ought to be taught or on any other measures of classroom activities.

For the 4-year-old cohort, in fact, there were no significant differences on any measures. Those differences that did emerge for the 3-year-old Head Start group suggest they had kindergarten teachers who had completed more coursework in teaching reading and 1st grade teachers with more coursework in reading and in early education than the 3-year-old control group. Likewise, the 3-year-old Head Start group was more likely to be in classrooms that conducted more math activities in the 1st grade. Yet, the vast majority of measures of school quality showed no significant difference for either cohort.

Not surprisingly, the study children – regardless of Head Start status—attended schools with much higher levels of poverty than schools nationwide (as indicated by proportions of students eligible for free- and reduced-price meals) and were in schools with higher proportions of minority students.

Most children in both the Head Start and control groups attended public schools of middle quality as measured by student proficiency on state assessments in math and reading. There was however, one significant difference in these test scores: for the 3-year-old cohort, there was a significant difference in the schools the Head Start group and control group attended for kindergarten. Math proficiency scores were higher in the schools attended by the control group than in those attended by the Head Start group.

While there were very few statistically significant differences in experiences for the Head Start and control group children, the overall findings for both groups can contribute to an understanding of the school environment experienced by both groups of children. For example,

nearly 50 percent of the 4-year-old cohort and 40 percent of the 3-year-old cohort were in classrooms where the teachers reported well-behaved students, with slightly smaller percentages in classrooms with teachers reporting occasional misbehavior and much smaller percentages in classrooms with teachers reporting frequent misbehavior. Teachers were asked about the reading, language and math activities that were provided in their classrooms on a daily basis. On average, kindergarten children in both cohorts and across both the Head Start and control group were exposed to about one-half of the reading, language, and math activities on a daily basis. In 1st grade, this dropped to about one-third of the activities.

Impacts on Children's Cognitive Development

The cognitive domain consists 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. The findings are summarized below.[8] All measures are described in Chapter 2 of this report. Exhibits 3a and 3b present all statistically significant cognitive impacts and their effect sizes.[9]

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 (WJ III) Letter-Word Identification; (3) WJ III Spelling; (4) WJ III Pre-Academic Skills; (5) Color Identification; and (6) Letter Naming.
- Parents of children in the Head Start group reported that their children had more emerging literacy skills at the end of Head Start than did parents of children in the control group. (This measure was not collected when the children were in school.)
- There were no impacts for 4-year-olds in the cognitive domain at the end of kindergarten.
- At the end of 1st grade, there is suggestive evidence of a positive impact of access to Head Start on PPVT (vocabulary) scores.
- 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.

---

[8] Three levels of evidence are considered in this report: (1) Strong evidence is used for impacts statistically significant at the $p \leq 0.05$, 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.

[9] 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 3a. Summary of Cognitive Impacts for 4-Year-Olds by Year

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

KEY:

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

Gray cell indicates the outcome is not applicable for that year.

Blank cell indicates a nonsignificant impact.

Note: Intent to Treat effect sizes are presented only for statistically significant differences ($p<.10$). 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.

3-Year-Old Cohort

- At the end of their Head Start year, there was strong evidence that the Head Start group demonstrated better skills on the following five child outcomes related to children's language and literacy development: (1) PPVT (vocabulary), (2) WJ III Letter-Word, (3) Preschool Comprehensive Test of Phonological and Print Processing (CTOPPP) Elision, (4) Letter Naming, and (5) WJ III 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 WJ III 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 some suggestive evidence of an impact on Oral Comprehension at the end of 1st grade.
- No statistically significant impacts were found for teacher reports of children's school performance in kindergarten and 1st grade with the exception of a lower teacher assessment in kindergarten of Head Start children's math ability. This was the only unfavorable cognitive impact found for either cohort as a whole in any year and was not supported by children's scores on the three direct math assessments where there was no evidence of math differences. Additionally, the schools attended by the control group children in the 3-year-old cohort during their kindergarten year report a significantly higher percent of students at or above the proficient level in math than the schools attended by the Head Start group children.

To provide context, we can compare the skill levels of children in the Head Start Impact Study with those of the general population of 3- and 4-year-olds in the United States (including those who were not from low-income families). The average 2003 PPVT score for a child in the 4-year-old control group was at the 27th percentile among children in the general population. Head Start group children's scores were four percentile points higher, at the 31st percentile. For the 3-year-olds, average 2003 PPVT scores were at the 29th percentile for the control group and the 32nd percentile for the Head Start group.

The study children also lag behind other children in the nation on letter identification. Fifty-five percent of the 4-year-old Head Start group and 65 percent of the 3-year-old Head Start group can recognize all their letters by the end of their kindergarten year. For the control group, 58 percent of the 4-year-olds and 64 percent of the 3-year-olds recognize all their letters by the end of kindergarten. Comparing these numbers to a nationally representative sample of children

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

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

KEY:

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

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

Gray cell indicates the outcome is not applicable for that year.

Blank cell indicates a nonsignificant outcome.

Note: Intent to Treat effect sizes are presented only for statistically significant differences ($p<.10$). 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.

from the Early Childhood Longitudinal Study-Kindergarten Cohort (ECLS-K) (2002),[10] 95 percent of children know all of the letters of the alphabet by the end of their kindergarten year.

Impacts on Children's Social-Emotional Development

The social-emotional domain consists of parent-reported measures during the Head Start years and reports by both parents and teachers in the early elementary school years. Measures of children's behavior, social skills and approaches to learning, parent-child relationships, teacher child relationships, and school adjustment were assessed. The findings in this domain are summarized below, and Exhibits 4a and 4b provide all statistically significant impacts for both cohorts and their effect sizes.

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 1st 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 there is suggestive evidence of more problematic student-teacher interactions.
  - In contrast, there is suggestive evidence of less withdrawn behavior for children in the Head Start group as reported by their parents.

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 three-year-old cohort in either the kindergarten or 1st grade year.

[10] U.S. Department of Education, National Center for Education Statistics. (2002). *Children's Reading and Mathematics Achievement in Kindergarten and First Grade.* Washington, DC: Author.

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

	Measure	Age 4 (Head Start Year)	K	1st Grade
Parent-Reported Measures	Aggressive Behavior			
	Hyperactive Behavior			
	Withdrawn Behavior			-0.13
	Total Problem Behavior			
	Social Competencies			
	Social Skills and Positive Approaches To Learning			
	Closeness			
	Conflict			
	Positive Relationships			
Teacher-Reported Measures	Aggressive (ASPI)			
	Interactive/Hyperactive (ASPI)			
	Withdrawn/Low Energy (ASPI)			
	Oppositional (ASPI)			
	Problems with Peer Interaction (ASPI)			
	Shy/Socially Reticent (ASPI)			0.19
	Problems with Structured Learning (ASPI)			
	Problems with Teacher Interaction (ASPI)			0.13
	Closeness			
	Conflict			
	Positive Relationships			

KEY:

Blue cell indicates a significant favorable impact ($p \leq 0.10$**).**

Yellow cell indicates a significant unfavorable impact ($p \leq 0.10$**).**

Gray cell indicates the outcome is not applicable for that year.

Blank cell indicates a nonsignificant outcome.

Note: Intent to Treat effect sizes are presented only for statistically significant differences ($p<.10$). 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. Chapter 5 provides an explanation for the directionality of outcomes.

To provide context for the social-emotional findings, a t-score of 60 or higher for any Adjustment Scales for Pre-school Intervention (ASPI) component empirically confirms a problem with that component. The percent of empirically confirmed problems for the study children at the end of 1st grade ranges from a low of five to six percent on the shy/socially reticent component to a high of 25 to 27 percent on the problems with peer interaction component.

Exhibit 4b. Summary of Social-Emotional Impacts for 3-Year-Olds by Year

	Measure	Age 3 (Head Start Year)	Age 4	K	1st Grade
Parent-Reported Measures	Aggressive Behavior				
	Hyperactive Behavior	-0.21		-0.12	
	Withdrawn Behavior				
	Total Problem Behavior	-0.14			
	Social Competencies				
	Social Skills and Positive Approaches To Learning		0.11	0.14	
	Closeness				0.10
	Conflict				
	Positive Relationships				0.10
Teacher-Reported Measures	Aggressive (ASPI)				
	Interactive/Hyperactive (ASPI)				
	Withdrawn/Low Energy (ASPI)				
	Oppositional (ASPI)				
	Problems with Peer Interaction (ASPI)				
	Shy/Socially Reticent (ASPI)				
	Problems with Structured Learning (ASPI)				
	Problems with Teacher Interaction (ASPI)				
	Closeness				
	Conflict				
	Positive Relationships				

KEY:

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

Gray cell indicates the outcome is not applicable for that year.

Blank cell indicates a nonsignificant outcome.

Note: Intent to Treat effect sizes are presented only for statistically significant differences ($p<.10$). 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. Chapter 5 provides an explanation for the directionality of outcomes.

Impact on Health Status and Access to Health Services

The findings in this domain comprise two categories: (1) children's receipt of health care services and (2) their current health status. The findings in this domain are summarized below, while Exhibits 5a and 5b present all statistically significant findings and their effect sizes for both cohorts of children.

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 1st grade, there was still moderate evidence of increased health insurance coverage among the Head Start group —a difference of four percentage points.

Exhibit 5a. Summary of Health Impacts for 4-Year-Olds by Year

	Measure	Age 4 (Head Start Year)	K	1st Grade
Parent-Reported Measures	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:

Blue cell indicates a significant favorable impact ($p \leq 0.10$**).**

Blank cell indicates a nonsignificant outcome.

Note: Intent to Treat effect sizes are presented only for statistically significant differences ($p<.10$). 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. Chapter 6 provides an explanation for the directionality of outcomes.

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 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 1st grade.

Exhibit 5b. Summary of Health Impacts for 3-Year-Olds by Year

	Measure	Age 3 (Head Start Year)	Age 4	K	1st Grade
Parent-Reported Measures	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		0.10		

- **KEY:**

Blue cell indicates a significant favorable impact ($p \leq 0.10$**).**

Blank cell indicates a nonsignificant outcome.

- Note: Intent to Treat effect sizes are presented only for statistically significant differences ($p<.10$). 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. Chapter 6 provides an explanation for the directionality of outcomes.

Comparing the health status of the children in the Head Start Impact Study with children in the general population demonstrates that Head Start children are about equal to other children on general health status, as measured here, and receipt of dental care. The target for the Healthy People 2010 initiative (U.S. Department of Health and Human Services, 2000)[11] is to increase the proportion of low-income children and adolescents who receive any preventive dental services to 66 percent. For the 4-year-olds, 73 percent of children in the Head Start group had seen a dentist since September in their Head Start year, compared with 56 percent of children in the control group. For the 3-year-olds, 68 percent of children in the Head Start group had seen a dentist since September of their Head Start year, compared with 52 percent of children in the control group and 74 percent of the Head Start group had seen a dentist since September of their age 4 year, compared with 65 percent of the control group. By the end of 1st grade, the

[11] U.S. Department of Health and Human Services. (2000). *Healthy People 2010: 21 – Oral Health*. Retrieved from:www.healthypeople.gov/data/midcourse/html/focusareas/FA21Objectives.htm.

percentage who had seen a dentist since the start of the 1st grade year in the Head Start group was about 68 percent for the 4-year-old cohort and 74 percent for the 3-year-old cohort, with no statistically significant difference from the control group.

In the ECLS-K study (2000),[12] 83 percent of parents reported that their child's health status was good or excellent at the beginning of kindergarten. In this study, at the end of kindergarten, 82 percent of the Head Start group parents for the 4-year-old cohort reported that their child's health was good or excellent, compared with 76 percent of the control group parents. Eighty-one percent of the Head Start group parents for the 3-year-old cohort reported that their child's health was good or excellent, with no statistically significant difference from the control group. These numbers compare favorably, suggesting that the low-income parents in this study do not see their children as more or less likely to be in good health than parents nationally. However, we have only very limited measures of health in this study, and cannot make definitive statements about how specific health conditions of the study children (e.g., asthma, obesity) compare to other children nationally.

Impact on Parenting Practices

This domain consists of four categories of outcomes: (1) disciplinary practices, (2) educational supports, (3) safety practices, and (4) parenting styles. The findings for the 3-year-olds in the parenting practices domain are consistent with favorable impacts on children's behavior through kindergarten and on parent-child relationships and closeness through 1st grade as reported in the social-emotional domain. The findings in this domain are summarized below, and Exhibits 6a and 6b provide the statistically significant findings and effect sizes.

4-Year-Old Cohort

- There were minimal impacts for the 4-year-old cohort in this domain, with one exception: at the end of the Head Start year, parents in the Head Start group were less likely to use time out as a disciplinary practice than were parents in the control group. In the absence of any pattern of impacts on social-emotional development or other parenting practices for this cohort, it is difficult to interpret this finding, which might reflect changes in either children's behavior or parents' reactions to it.

3-Year-Old Cohort

- In the Head Start year, there were several impacts on parenting practices, and most 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 also 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.

[12] U.S. Department of Education, National Center for Education Statistics. (2000). *America's Kindergartners.* Washington, DC: Author.

- At the end of the age 4 year, parents of children in the Head Start group were 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 continues in kindergarten and 1st grade.
  - At the end of kindergarten, there is 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.
  - At the end of 1st grade, there is 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 more likely to use an authoritarian parenting style.

Exhibit 6a. Summary of Parenting Impacts for 4-Year-Olds by Year

	Measure	Age 4 (Head Start Year)	K	1st Grade
Parent-Reported Measures	Parent Spanked Child in Last Week			
	Parent Used Time Out in Last Week	-0.17		
	Parent Read to Child in Last Week			
	Parental Safety Practices Scale			
	Family Cultural Enrichment Scale			
	Parenting Style: Authoritarian			
	Parenting Style: Authoritative			
	Parenting Style: Neglectful			
	Parenting Style: Permissive			
Teacher-Reported Measures	School Contact and Communication			
	Parent Participation			

- **KEY:**
- **Blue cell indicates a significant favorable impact (**$p \leq 0.10$**).**
- **Gray cell indicates the outcome is not applicable for that year.**
- Note: Intent to Treat effect sizes are presented only for statistically significant differences. 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. Chapter 7 provides an explanation for the directionality of outcomes.

Exhibit 6b. Summary of Parenting Impacts for 3-Year-Olds by Year

Measure		Age 3 (Head Start Year)	Age 4	K	1st Grade
Parent-Reported Measures	Parent Spanked Child in Last Week	-0.14		-0.09	
	Parent Used Time Out in Last Week			-0.13	-0.11
	Parent Read to Child in Last Week	0.15			
	Parental Safety Practices Scale				
	Family Cultural Enrichment Scale	0.18			
	Parenting Style: Authoritarian		-0.14		-0.11
	Parenting Style: Authoritative				
	Parenting Style: Neglectful				
	Parenting Style: Permissive				
Teacher-Reported Measures	School Contact and Communication				
	Parent Participation				

KEY:

Blue cell indicates a significant favorable impact ($p \leq 0.10$**).**

Gray cell indicates the outcome is not applicable for that year.

Note: Intent to Treat effect sizes are presented only for statistically significant differences. 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. Chapter 7 provides an explanation for the directionality of outcomes.

Research has demonstrated that reading to children has a positive effect on their literacy outcomes (Denton, Reaney & West, 2001; Snow, Burns, & Griffin, 1998). In the ECLS-K study (2000)[13], about 45 percent of parents reported reading to their kindergarten children every day and 80 percent of parents reported reading to their children at least three times per week. For the Head Start Impact Study at the end of kindergarten, 35 percent of Head Start group parents of children in the 4-year-old cohort reported reading to their children every day while 69 percent of these parents reported reading to their children at least three times per week, with no statistically significant difference from the control group. For the 3-year-old cohort, 34 percent of the Head Start group parents reported reading to their children every day while 65 percent of these parents reported reading to their children at least three times per week, with no statistically significant difference from the control group.

Variation in Impact

This report examines differences in impact among different types of children and parents. Seven dimensions were used to define subgroups: (1) whether a child had low pre-academic skills at the start of Head Start (lowest quartile), (2) whether the child was a Dual Language

[13] U.S. Department of Education, National Center for Education Statistics. (2000). *America's Kindergartners.* Washington, DC: Author.

Learner at the start of Head Start, (3) whether the child had special needs (as reported by the parent at the start of Head Start), (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 8.

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 (almost 10,000, taking into consideration the study's two cohorts, four time points for measuring outcomes, and multiple outcomes), 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. Particular attention was paid to multiple impacts that occur across domains or outcomes, or that persist into kindergarten and 1st grade.[14] 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. The major findings are presented below.

4-Year-Old Cohort

- **Children of parents with mild depressive symptoms** experienced favorable cognitive impacts through the end of 1st grade. At the end of the Head Start year, this subgroup experienced several benefits of Head Start in language and literacy compared to children of parents with other levels of reported depressive symptoms. No cognitive impacts were found in kindergarten but the language and literacy impacts re-appeared at the end of 1st grade.
- **Dual Language Learners** experienced health benefits from Head Start at the end of kindergarten and 1st grade, although the early impacts were mixed. Head Start had a favorable impact on the health insurance coverage of Dual Language Learners at the end of the Head Start and kindergarten years and on the receipt of dental care at the end of 1st grade.
- **Children in the lowest academic quartile at baseline** showed benefits of Head Start in the social-emotional domain through 1st grade. At the end of the Head Start year, there were favorable impacts for the lowest quartile children compared to their non-lowest quartile counterparts on parent's report of their relationship with the child. In the school years, teacher reports showed more favorable impacts for lowest quartile children on oppositional behavior, problems with peer interaction, conflict, and positive relationships with the teacher than non-lowest quartile children.

[14] 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.

- In addition, **Black children** experienced favorable impacts in the social-emotional domain at the end of kindergarten as reported by teachers. Black children in the Head Start group were reported to have reduced inattentiveness, fewer problems with structured learning, peer interactions, or teacher interactions, and better relationships with teachers.

3-Year-Old Cohort

- **Children with special needs** benefited from Head Start in the math and social-emotional areas at the end of the 1st grade. As a result of Head Start, special needs children showed a reduction in inattention/hyperactivity, in problems with structured learning, and in conflict with teachers, as well as an increase in positive teacher relationships.
- **Children of parents with no depressive symptoms** experienced sustained benefits of Head Start in the cognitive, social-emotional, and parenting domains through the end of 1st grade. In the cognitive domain, children of parents with no reported depressive symptoms benefited from Head Start on many direct assessments of language, literacy, and math skills in all years, and especially at the end of 1st grade.
- **Children from high-risk households** showed sustained favorable cognitive impacts through the end of 1st grade. Children from high-risk households experienced benefits in five direct assessments of academic skills at the end of 1st grade.
- **Children in non-urban settings** showed sustained cognitive benefits from Head Start through the end of 1st grade and some benefits in the social-emotional domain during the Head Start years. Children in non-urban settings demonstrated favorable cognitive impacts at the end of their Head Start year on three measures of language and literacy and one pre-writing measure. Additionally, favorable math impacts were demonstrated at the end of the age 4 year and favorable spelling impacts at the end of kindergarten. At the end of 1st grade, there were favorable impacts on six language and literacy measures and one math measure.
- There were also several groups with favorable impacts in the earlier years of followup that faded by the 1st grade, including children in the lowest quartile at baseline and Dual Language Learners.

There were also a few subgroups that experienced a pattern of mixed or unfavorable impacts by 1st grade. For example, White children in the 4-year-old cohort experienced unfavorable impacts on several teacher-reported social-emotional measures in the 1st grade year and one unfavorable impact on parenting in the kindergarten year. Within the 3-year-old cohort, children of parents with moderate depressive symptoms experienced sustained negative impacts of Head Start in the cognitive, social-emotional, and health domains and mixed impacts in the parenting domain through 1st grade. These children were less likely to be promoted to the next grade, as reported by their parents. This group is of particular concern because the unfavorable impacts were found across domains and methods of assessment.

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 examined the impacts of Head Start on these four domains.

The study shows that providing access to Head Start led to improvements in the quality of the early childhood settings and programs children experienced. On nearly every measure of quality traditionally used in early childhood research, the Head Start group had more positive experiences than those in the control group.

These impacts on children's experiences translated into favorable impacts at the end of one year in the domains of children's cognitive development and health, as well as in parenting practices. There were more significant findings across the measures within these domains for 3-year-olds in that first year (and only the 3-year-old cohort experienced improvements in the social-emotional domain.) Yet, by the end of 1st grade, there were few significant differences between the Head Start group as a whole and the control group as a whole for either cohort.

The few differences at the end of 1st grade included a favorable impact on the receipt of health insurance for the four-year-old cohort, consistent with earlier impacts on health insurance for both cohorts. Further, there are longer-term favorable impacts related to children's social-emotional development and relationships with their parents for the 3-year-old cohort. According to parent reports, this younger group experienced favorable impacts on behavioral and social-emotional outcomes during the early years of the program and into kindergarten. By 1st grade, these impacts were limited to outcomes related to parent-child relationships and parenting practices. It is possible that these benefits in the parent-child relationship are both related to earlier improvements in behavior and may lead to longer term benefits for children. However, this is only one hypothesis, and the issue requires further analysis.

This study evaluated the Head Start program against a mixture of alternative care settings rather than against a condition of "no services" or parental care only condition. About 40 percent of the non-Head Start 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 did receive child care or early childhood education. Further, among those that 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 who participated in some type of non-parental care attended care about four to five hours more per week for the Head Start year. Consequently, to achieve measurable impacts, Head Start (as noted above) has to outperform what control group children received. Improved childcare and pre-K standards across the nation may have reduced the differences found between the Head Start and control group children.

Although the quality is high on average, Head Start programs vary in terms of instruction in the key areas measured as part of this study, i.e., early development of language and literacy and mathematics skills. The inclusion of programs across the full spectrum in this study's nationally representative sample may help to explain why impacts in the cognitive domain are not stronger.

Head Start has always had a particular emphasis on young children with special needs, and indeed, though exploratory, we see impacts through 1st grade in the two subgroups of relevance: children whose parents have been told their child has special needs or disabilities (for 3-year-olds) and children with the lowest cognitive skills upon entering Head Start. Head Start has benefits for these groups of children that last into early elementary school.

Similarly, the Head Start performance standards emphasize the importance of respecting children and individualizing services as needed based on their cultural and linguistic backgrounds. Indeed, among the 3-year-old treatment group, parents were more likely to enroll their children for a second year of Head Start if they found the program supportive of their culture, if they were immigrants, and if English was not the primary language used in the home. Likewise, the findings from this study demonstrate that Black children (in the 4-year-old cohort) and Dual Language Learners are among the groups that benefited more than other groups from access to Head Start. However, most of these impacts only lasted through the end of kindergarten.

The subgroup findings do not present a consistent picture of favorable impacts for groups that have traditionally been emphasized as higher risk. While the children from higher risk households benefited in the 3-year-old group, there were no differences in impacts by household risk for the 4-year-olds. Further, it was the children of caregivers with less severe depressive symptoms that experienced favorable impacts through 1st grade in both cohorts. There also appears to be a pattern for the 3-year-old cohort where the children from families in the middle of the risk categories (neither highest nor lowest) actually experienced some negative impacts. This pattern is particularly strong, and concerning, for children of caregivers with moderate levels of depressive symptoms. The explanation for these patterns is unclear and warrants more attention.

This study also found that, in the 3-year-old cohort, Head Start had benefits through 1st grade for children from non-urban communities. It is possible that this finding represents the difficulties that children and families in non-urban communities have in getting comprehensive services and in finding quality early care and education for their children, absent Head Start. Indeed, children were more likely to participate in a second year of Head Start if there was less competition from other preschools in the area. These are questions that should be pursued in future research.

In sum, this report finds that providing access to Head Start has benefits for both 3-year-olds and 4-year-olds in the cognitive, health, and parenting domains, and for 3-year-olds in the social-emotional domain. However, the benefits of access to Head Start at age four are largely absent by 1st grade for the program population as a whole. For 3-year-olds, there are few sustained benefits, although access to the program may lead to improved parent-child relationships through 1st grade, a potentially important finding for children's longer term development. Moreover, several subgroups of children in this study experience benefits of Head Start into 1st grade. It will be important in future research to examine whether the positive parent-child relationships for the 3-year-old cohort translate into improved outcomes as children get older, as well as whether the findings for subgroups of children persist over the longer term. To that end, the study children have been followed through 3rd grade. The 3rd grade report will examine the extent to which impacts of Head Start on initial school readiness are altered or maintained as children enter pre-adolescence. Further, that report will provide a greater focus on

how children's later experiences in the school and community affect their outcomes at 1st and 3rd grades.

Finally, this study leaves many important questions about Head Start unanswered. These questions include, but are certainly not limited to, questions such as: Is there a benefit to having two years of Head Start rather than one year? What types of programs, centers, classrooms, and other experiences relate to more positive impacts for children and families? What accounts for the subgroup patterns observed in this report? Are there some later experiences that help to sustain impacts through the early elementary grades? Hopefully, researchers will take advantage of the data from this study, which will be made available through a data archive, to further the understanding of the role Head Start plays in the well-being of children and families.

Chapter 1: Study Goals and Purposes

Introduction

The Head Start Impact Study (HSIS) responds to the 1998 congressional mandate to determine, on a national level, the impact of Head Start on the children it serves, both at the end of their Head Start experience and during their early elementary school years. Increased focus on outcomes and accountability for Federal resources, combined with concern about the lack of a rigorous experimental design to evaluate the effectiveness of the Head Start program nationally, resulted in this mandate (see Appendix A). The evaluation was designed to determine whether Head Start has impacts on participating children and their parents and whether any impacts vary among different types of children, families, and communities.

The study was conducted across 84 nationally representative[15] grantee/delegate agencies. Approximately 5,000 newly entering 3- and 4-year-old children applying for fall 2002 admission to Head Start were randomly assigned to either a Head Start group that had access to Head Start program services or to a control group that could enroll in available community services other than Head Start, selected by their parents. Local agency staff implemented their typical process of determining eligibility, and no changes were made to locally established ranking criteria. 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. This constraint ensured that the study's need for a control group did not require slots to go unfilled. Data collection for this report began in fall 2002 and continued through spring 2006, following children through the spring of their 1st grade year.

The study quantifies the impact of Head Start separately for children who entered the program as 3-year-olds and as 4-year-olds across child cognitive, social-emotional, and health domains as well as on parenting practices. Under the study's randomized design, a simple comparison of outcomes between the Head Start group and the control group yields an unbiased

[15] It should be noted that, as mandated by Congress, several groups of programs serving special populations are not included in this study. These are the Migrant and Seasonal Head Start programs, Early Head Start (which serves pregnant women and children from birth to age three), as well as the American Indian/Alaska Native tribal programs. Programs in Puerto Rico were included in the study, but findings for these programs are reported separately. Unless otherwise stated, findings in this report are not pertinent or necessarily generalizable to these programs.

estimate of the impact of Head Start on children's school readiness. This focus on impacts distinguishes this study from many others that seek primarily to examine relationships among participant outcomes or between participant outcomes and individual or program characteristics.

In addition, this study is set apart from most program evaluations because children were selected at random from those applying for entry into Head Start in a nationally representative sample of programs, making results generalizable to the Head Start program, not just to the selected sample of programs and children.

To establish the context for the findings from the HSIS, this chapter describes the Head Start program and relevant research conducted prior to the completion of the HSIS. Information describing the characteristics of Head Start programs nationwide at the time this study was fielded (in 2002) is also provided to set the context for the findings discussed in later chapters. Finally, the chapter describes the purpose and study goals for the Head Start Impact Study, before presenting a road map of the contents of the report.

What Is the Head Start Program?

The Head Start program, created in 1965 as part of the War on Poverty, is intended to boost the school readiness of low-income children. Head Start has grown from its early days of originally offering six-week summer sessions for 4-year-olds, to providing typically nine-month and sometimes year-long programs serving children from three to five years of age. The program is dedicated to promoting school readiness and providing comprehensive child development services to low-income children, their families, and communities, with an underlying premise that low-income children and families need extra support to prepare them for the transition to school. In general, during the study, to be eligible for Head Start, a child had to be living in a family whose income was below the Federal poverty line. Programs were permitted, however, to fill ten percent of their enrollment with children from families that are over this income level. More recently, since the 2007 reauthorization of Head Start, programs may serve up to 35 percent of their enrollment from children whose families' incomes are below 130 percent of the poverty line. Programs were, and still are, required to make at least ten percent of the total number of enrollment opportunities during each enrollment year available to children with disabilities.

Based on a "whole child" model, the Head Start program focuses on ". . . helping preschoolers develop the early reading and math skills they need to be successful in school. . . . Head Start programs promote school readiness by enhancing the social-emotional and cognitive development of children through the provision of educational, health, nutritional, social and other services to enrolled children and families. They engage parents in their children's learning and help them in making progress toward their educational, literacy and employment goals. Significant emphasis is placed on the involvement of parents in the administration of local Head Start programs."[16] Head Start is administered by local grantees and public and private non-profit and for-profit agencies that must adhere to national program guidelines—the Head Start Performance Standards—to ensure that programs provide a wide array of comprehensive services for families and children. Local Head Start programs conduct a needs assessment of each child's and family's needs and strengths at the beginning of the program and use this to tailor and guide services. Likewise, ongoing evaluations of children's development and progress are used to individualize services at the classroom level. According to a Government Accounting Office report (U.S. GAO, 2003), most Head Start programs meet the requirements of the Head Start Performance Standards.[17]

The Head Start program offers services to children and families through a variety of program options. The most common of these, and concomitantly the highest proportion of the study sample, is a center-based program option in which children are enrolled in classroom settings and parents participate in at least two home visits annually. Three other options represented in the sample are: (1) a home-based program option in which staff work directly with children and parents primarily in the home on a weekly basis and also in at least twice monthly group socialization activities, (2) a family child care option, in which services to children and families are provided in a family child care setting, and (3) the combination program option that allows for a variety of combinations of center-based class sessions with home visits. Grantees may propose to offer any or all of these options, or may design a different option subject to approval from the Office of Head Start.

[16] http://www.acf.hhs.gov/opa/fact_sheets/headstart_printable.html

[17] http://eclkc.ohs.acf.hhs.gov/hslc/Program%20Design%20and%20Management/Head%20Start%20Requirements/Head%20Start%20Requirements.

Each program conducts a community needs assessment to determine which options and services best fit the strengths and needs of families in the community. Head Start programs work in partnership with other service providers, adjust schedules to meet the needs of the populations served, vary in length from school-year to full-year, and for those that provide services to children in out-of-home settings, can vary in intensity from part-day to full-day. All of these variations are represented in the sample for this study.

Since 1965, the context in which the program delivers services has changed dramatically. Most notably, greater cultural diversity of the population and increasing prevalence of Dual Language Learners have combined to increase the challenges in Head Start's responsibility to be responsive to each child's and each family's unique needs and ethnic, cultural, and linguistic heritage. Increased immigration from Latin America, the Caribbean, Asia, and the Middle East has created a more diverse population of American children (Cappella & Larner, 1999), and resulted in Head Start serving a wider variety of ethnic and racial minority groups. Since the inception of the program, family structure also has changed with the decline of what was once considered the traditional family consisting of children living with a father in the labor force and an at-home mother. For example, births to unmarried mothers in the U.S. have risen dramatically, to 1,641,946 babies in 2006, with the proportion of all births to unmarried women having reached 38.5 percent of all U.S. births that year (Martin, Hamilton, Sutton et al., 2009). In addition, there has been an increase in the number of children involved in a divorce (Johnson and O'Brien-Strain, 2000) and an increase in the number of mothers entering the workforce and seeking child care.

According to Head Start data reported by programs to the Federal government, in FY 2003, enrollment for the entire Head Start program (including Early Head Start, Migrant and Seasonal and American Indian/Alaska Native tribal programs) was 909,608 children (U.S. Department of Health and Human Services, 2004). Most of the children who received Head Start services were between three and five years old (92 percent). Eight percent were infants and toddlers (birth to three years). As of FY 2003, Head Start had served a total of over 22 million preschool-age children, infants, and toddlers since its inception in 1965. At that time, the program included 1,670 Head Start grantees, 47,000 classrooms, and 19,200 centers. Head Start programs employed 206,000 staff, who, with the assistance of 1,372,000 volunteers, worked to

provide comprehensive services to meet the early childhood development, educational, health, and family needs of children.

What Does Research Say About Early Childhood Education Programs?

There is a large body of research on the effects of early childhood education in general, but only a relatively small set of approaches has been evaluated using scientifically rigorous methods. Even fewer rigorous evaluations have concentrated on Head Start, and only some of these assessed the program's impact on a national level.

The early research on Head Start and other early childhood education programs often focused on the effects of these programs on children's IQ scores and the sustainability of these effects over time. Some of these early studies, begun in the 1960s and 1970s, have continued into the present to study children over time, providing insight into both the short- and long-term effects of participation in early childhood and preschool programs.

The first major evaluation of Head Start itself was conducted in the late 1960s (Cicirelli, 1969). The Westinghouse Study, as it is called, concluded that Head Start boosted children's intelligence, but that the impact faded after a few years in elementary school (Ramey & Ramey, 2004). However, this early study was conducted when Head Start was a very new Federal program, and most local programs offered services only in the summer. Furthermore, the study has been widely criticized because the program and comparison groups were not equivalent (Campbell & Erlebacher, 1970), and because it focused only on children's cognitive outcomes as opposed to the comprehensive goals and services of the program. Nevertheless, the findings continue to be remembered as casting doubt on the long-term effects of Head Start.

Other evaluation studies of Head Start suggest positive child outcomes in various domains including cognitive, health, and socio-emotional development. For example, preliminary results from a longitudinal study of more than 600 Head Start graduates in San Bernardino County, California (Meier, 2003), showed that final kindergarten report card grades/ratings of Head Start graduates were higher in numeracy, language, literacy, social conduct, and physical development as compared to their non-Head Start peers (including two comparison sub-samples matched by school district and kindergarten class). Hubbs-Tait and

colleagues (2002) report an association between Head Start attendance for high-risk children and increased receptive vocabulary, and Barnett (2002) reports favorable long-term effects on grade repetition, special education, and high school graduation rates for Head Start children.

Although studies have found no evidence for an increased direct benefit to children attending Head Start for two years over those with one year of attendance (Kreisman, 2003; Ritblatt, Brassert, Johnson and Gomez, 2001), there is evidence that two years of Head Start participation has positive effects on the home environment (Ritblatt, Brassert, Johnson & Gomez, 2001). Additionally, a study by Abbott-Shim, Lambert, & McCarty (2003) using a randomized wait-list design in three Head Start centers, found that 4-year-olds participating in Head Start performed better in receptive language and phonemic awareness than 4-year-olds of similar backgrounds who were wait-listed for the Head Start program. Head Start children also were significantly more likely to have had medical check-ups, routine health screenings, immunizations, and dental check-ups.

Several other evaluations of early childhood education programs also have shown impacts for children, with some even showing benefits into adulthood. The High/Scope Perry Preschool Project, which began in 1962, has been the subject of a well-known longitudinal study. A total of 123 African American 3- and 4-year-old children were randomly assigned to a program group that participated in an active learning preschool program and a control group that did not receive a preschool program. The program group significantly outperformed the control group on various intellectual and language tests from their preschool years up to age seven and later on school achievement and literacy tests, with findings of program effects through age 40 spanning the domains of education, economic performance, crime prevention, family relationships, and health (Schweinhart, 2005). The children in this program appear to have benefited into adulthood, with benefits in level of schooling completed, attitudes toward school, and at age 40, rates of employment, income, and family harmony as well as fewer arrests (Schweinhart, 2005). However, study results have been questioned because the treatment and control groups were adjusted after randomization, when about one-fifth of the sample children were moved in order to keep siblings together, and to place all children of working mothers in the control group.

Findings from another very early and well-known randomized, longitudinal study of an early childhood intervention, the Abecedarian Project, include some important, long-term benefits. The Abecedarian project, which began in Chapel Hill, North Carolina, during the 1970s, randomly assigned four cohorts of infants to either an early educational intervention group or a control group. The program focused on cognitive, language, and social/emotional areas of development and provided family support services with referrals to community social and health service providers. The program had as its goals to improve school readiness and later school performance and was conceived as an intensive preventive intervention. Compared to the control group, program participants performed better on cognitive assessments from the toddler years to age 21, had higher reading and math achievement from the primary grades through young adulthood, and completed more years of education (Ramey & Ramey, 2004).

While the evaluations of the Perry Preschool Project and the Abecedarian Project suggest that early childhood education programs can have long-lasting effects, both studies were conducted over three decades ago, when the control group largely stayed at home with their parents. Further, both studies were small and not representative of the nation's children. Nevertheless, these studies have been instrumental in helping to demonstrate the potential of early childhood education programs and to shape the evaluation described in this report.

More recently, a longitudinal study was begun in 1986 to investigate the short- and long-term effects of the Chicago Child-Parent Center and Expansion Program that has been operating in Chicago Public Schools since 1967. The program provides school-based educational enrichment and comprehensive family services from preschool to 3rd grade. Using a quasi-experimental design, researchers reported strong effect sizes for test score impacts at school entry, and while these decreased over time, statistically significant test score differences did remain between the treatment and the comparison groups at grade six (Reynolds & Temple, 1995). Moreover, long-term benefits were found for program participants in key areas of societal importance such as high school completion, educational attainment, felony arrest rates, depressive symptoms, full-time employment, and disability rates (Reynolds, Temple, et al., 2007).

Evaluations of state pre-K programs also provide evidence that early childhood education can be effective in improving children's development, particularly in the cognitive domain. A

study of five state-funded pre-K programs for 4-year-olds in Michigan, New Jersey, Oklahoma, South Carolina, and West Virginia (Barnett, Lamy, & Jung, 2005) used a regression discontinuity approach, defined by the age cutoff for kindergarten eligibility, to compare two groups of children of similar ages. Both Oklahoma and West Virginia have universal pre-K programs while Michigan, New Jersey, and South Carolina target the at-risk population. In the fall, 5,071 children were tested to measure receptive vocabulary, early literacy, and math skills. Substantial gains were found in all five states with statistically significant impacts on children's early language, literacy, and math development.

The studies of pre-K add to the research on early childhood development and school readiness, however, caution should be used in directly comparing results of pre-K studies with Head Start. The population, study design, and outcomes measured often differ across these studies (Wong, Cook, Barnett, & Jung, 2008; IPR News, Fall 2006).

Finally, studies of both Head Start and other early childhood programs have sought to determine the role of various aspects of classroom "quality" in affecting child outcomes. Some research has focused on observing the behavior of teachers and the responses of children in child care settings to measure quality, while other research has focused on certain "structural" indicators, such as low child-to-staff ratios, small group sizes, well-educated teachers, or well-trained teachers (Gormley, 2007). A question still remains as to how to measure the various components of quality accurately and effectively, and to date, the findings of studies seeking to examine the role of classroom quality in relation to child outcomes, have been mixed (Early, Maxwell, Burchinal, et al., 2007; U.S. Department of Health and Human Services, 2002; and the NICHD Early Childcare Research Network and Duncan, G., 2003).

In sum, a review of the literature to date suggests that early childhood care and education programs can affect children's school readiness and long-term outcomes. However, there has not been a study with a rigorous experimental design and a nationally representative sample to evaluate how Head Start itself contributes to key outcomes. The Head Start Impact Study is the first such study. Following is a description of the congressional mandate which led to the study objectives and research questions.

The Congressional Mandate

The U.S. General Accounting Office (GAO) concluded in the late 1990s that (1) ". . .the body of research on current Head Start is insufficient to draw conclusions about the impact of the national program" (GAO, 1997) and (2) ". . .the Federal government's significant financial investment in the Head Start program, including plans to increase the number of children served and enhance the quality of the program, warrants definitive research studies, even though they may be costly" (GAO, 1998).

Based on the GAO's recommendations, and on the testimony of research methodologists and early childhood experts, Congress included in the 1998 reauthorization of Head Start a mandate that the U.S. Department of Health and Human Services (DHHS) determine the impact of Head Start on the children it serves.[18] The legislation also required the Secretary of DHHS to establish an Advisory Committee on Head Start Research and Evaluation. The Committee included 30 experts in areas of program evaluation and research, education, child care and early childhood programs, education policy, and economics. They set forth a framework for research on the impact of Head Start that would be both scientifically credible and feasible. The Committee acknowledged that the legislative language recommended the use of a rigorous methodology, including random assignment of children to Head Start and non-Head Start groups at a diverse group of sites, selected nationally and reflecting the range of Head Start quality across the country. To implement this design, DHHS competitively awarded a contract in October 2000 to Westat of Rockville, MD, and its team of collaborating partners, which currently include Chesapeake Research Associates, LLC; Abt Associates; the Urban Institute; and American Institutes for Research.

Study Objectives and Research Questions

Study Design Requirements

The congressional mandate required a study of the "impact" of Head Start, defining the term in the legislation as the "*. . .difference in an outcome for a participant in a program that*

[18] See Appendix A for the research-related amendments to the *Head Start Act* included in the 1998 reauthorization. All citations from this legislation can be found in this appendix.

would not have occurred without the participation in the program." Thus, impact means a difference between the outcomes observed for Head Start participants and what would have been observed for these same individuals had they not had the opportunity to participate in Head Start. Although the language in the legislation permitted different study designs, the Advisory Committee concluded that a randomized control trial (RCT) would provide the most reliable evidence of causal linkage between Head Start and intended outcomes. As will be discussed below, this was the study's design, in which a sample of 3- and 4-year-old Head Start applicants not previously served by the program[19] were randomly assigned either to a Head Start group (in which children and families received Head Start services) or to a control group (in which children were not granted access to Head Start but could receive any other available services chosen by their parents). Under this randomized design, a simple comparison of outcomes for the two groups yields an unbiased estimate of the impact of Head Start on child and parent outcomes. The advantage of this research design is that if random assignment is properly implemented with a sufficient sample size, program participants on average will not differ in any systematic way from non-participants except through their access to Head Start services.[20] This is true for both measured and unmeasured characteristics.

The legislation also suggested that the control group should represent the real world, i.e., it should be *". . .composed of—(i) individuals who participate in other early childhood programs (such as public or private preschool programs and day care); and (ii) individuals who do not participate in any other early childhood program."* In other words, the effects of Head Start were to be compared to the range of options that low-income families have for their preschool children and not against a no-services alternative, in which all children would spend all of their time at home with parents or other family members.

The legislation also clearly intended that the study be nationally representative, stating that: *"The Secretary shall ensure that the study focuses primarily on Head Start programs that operate in the 50 States, the Commonwealth of Puerto Rico or the District of Columbia and that*

---

[19] The Head Start Impact Study focuses on newly entering children to ensure that the estimated impacts are unaffected by previous program participation. Consequently, children who were returning to Head Start, as well as those previously enrolled in Early Head Start, were excluded from the study sample.

[20] More precisely, there will be differences between individuals in the two groups, but the expected or average value of these differences is zero except through the influence of Head Start (i.e., selection bias is removed by random assignment).

do not specifically target special populations" and that the selection of study participants should *". . .make use of random selection from the population of all Head Start programs. . .in selecting programs for inclusion in the research."* Accordingly, the study was designed with a nationally representative, probability-based sample of participating programs, centers, and children. Finally, Congress specified that the intent of the study was to examine program impacts over time, i.e., *". . .the impact of Head Start programs on participants on the date the participants leave Head Start programs, at the end of kindergarten, and at the end of first grade (whether in public or private school). . . ."*

Research Questions in Legislation

As specified in the authorizing legislation, the overall goal of the study was to determine, *". . .if, overall, the Head Start programs have impacts consistent with their primary goal of. . .increasing. . .school readiness."* Through the additional requirement to examine impacts over time, Congress also expressed interest in the impact of Head Start on children's later progress. With regard to school readiness, Congress specified that the study should address the impact of Head Start on:

- *". . .the growth and development of children in cognitive, emotional, and physical health areas";*
- *". . .families as the primary nurturers of their children. . . (to). . . ensure that children attain school readiness";*
- *". . .increasing access of children to such services as educational, health, and nutritional services, and linking children and families to needed community services."*

In other words, Congress envisioned a study that examined the direct impact of Head Start on child/family access to services and child developmental outcomes and the indirect impact of Head Start on children through the programs impact on their parents.

In addition to these overall impacts, Congress called for the consideration of *". . .possible sources of variation in impact of Head Start programs,"* such as*:*

- *"participant characteristics"*—characteristics of the study children and families*;*
- *"the age of the child on entering the Head Start program"* and *"the length of time a child attends a Head Start program;"*and

- *"other characteristics and features of the Head Start program (such as geographic location, location in an urban or a rural service area. . . ."*

A final specification in the original legislation is to determine *"how receipt of services. . . enriches the lives of children and families participating in Head Start programs."* In other words, the Congress wanted to understand how Head Start impacts the families as well as the children the program serves.

Following the legislative requirements and the Advisory Committee's recommendations, the following detailed research questions were developed to guide the study. First and foremost, the study had to estimate the impact of Head Start on children, parents, and the services they receive, at the national level. This focus led to the first set of overall program impact questions.

Overall National Impact Estimates. 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?

The Direct Impact of Head Start on Children

- What is the impact of Head Start on children's cognitive development preceding the start of school? What is the impact of Head Start on children's cognitive development during the early school years?
- What is the impact of Head Start on children's social-emotional development preceding the start of school? What is the impact of Head Start on children's social-emotional development during the early school years?
- What is the impact of Head Start on children's health status preceding the start of school? What is the impact of Head Start on children's health status during the early school years?

The Potential Indirect Impact of Head Start on Children Through Direct Impacts on Parents

- What is the impact of Head Start on parents' practices and support of their child's school readiness preceding the start of school? What is the impact of Head Start on parents' practices and support of their child's education?

In addition, the study aimed to examine the extent to which impacts vary across different groups of children, parents, and families. This goal led to the second set of program impact questions:

Variation in Impacts. Under what circumstances does Head Start achieve the greatest impact? What works for which children? What Head Start services are most related to impact?

Variation by Child Characteristics

- How do the estimated impacts of Head Start vary by child characteristics, such as child's age, primary language, special needs status, and academic ability?

Variation by Parent and Household Characteristics

- How do the estimated impacts of Head Start vary by parent and household characteristics, such as parents' race and ethnicity, depressive symptoms, and level of risk?

Variation by Community Characteristics

- How do the estimated impacts of Head Start vary by the characteristics of the community where the child lived at the time of application to Head Start, such as urbanicity?

The final research question focuses on the impact of Head Start on the nature and type of children's experiences:

Impacts on Children's Experiences

- What is the impact of Head Start on the settings, setting characteristics, and services that children experience prior to starting school? During the early school years?

Building on these research questions, the study design, including data collection instruments and procedures, was developed and tested during 2001 and 2002. Site selection and recruitment were conducted during 2002 along with the random assignment of children. Initial baseline data were collected in fall 2002 with subsequent annual waves of data collection in spring 2003, 2004, 2005, and 2006, following children through the end of 1st grade. A preliminary report on first-year findings was published in 2005 (U.S. Department of Health and Human Services, 2005). A report following children through 3rd grade is also planned.

Contents of Report

This report, including the Executive Summary, present the findings from the preschool years through children's 1st grade experience.

This document consists of the Executive Summary and nine chapters. Chapter 1 presents the study background, including a literature review of related Head Start research and the study purpose and objectives. Chapter 2 provides details about the study design and implementation. It discusses the experimental design, sample selection prior to random assignment, data collection, and data analysis. To provide a context in which to understand the impact findings, Chapter 3 examines the impact of Head Start on the services and child care settings that children experience prior to starting school. It also provides the impact of Head Start on the educational and child care settings, setting characteristics, and services that children experience during kindergarten and 1st grade. Chapters 4 through 7 present the impact of Head Start on children's outcomes and parenting practices for the years before school and then for kindergarten and 1st grade. Chapter 4 presents the impact of Head Start on children's cognitive development, Chapter 5 presents the impact of Head Start on children's social-emotional development, Chapter 6 presents the impact of Head Start on children's health status and access to health services, and Chapter 7 presents the impact of Head Start on parenting practices in the areas of educational activities, discipline practices, and school involvement. Chapter 8 examines variation in impacts by child characteristics, parent and family characteristics, and community characteristics. Chapter 9 provides an overall summary of the findings, implications for the Head Start Program, and unanswered questions.

Appendices in this volume include the Head Start Impact Study legislation, a list of the official Head Start Impact Study Advisory Committee members, the language decision form used to determine the language in which the child was assessed, and data tables that elaborate on the findings presented in the volume (e.g., Impact on Treated (IOT) findings). Finally, the findings from a sample of programs in Puerto Rico are provided in an appendix. Programs in Puerto Rico were included in the study with the intent that data on children in these programs would be analyzed along with the data on children in the 50 states and the District of Columbia, once children reached school-age. However, due to differences in instruction and differences in the interpretation of outcomes in the Spanish-dominant Puerto Rico context, these data could not be included in the main study analyses and are presented separately.

The *Head Start Impact Study Final Report* provides details on the methods used for sampling, the methods used for collecting and analyzing data, as well as the psychometric information for the data collection measures. Additionally, along with the Technical Report,

detailed tables for all analyses will be provided on the Administration for Children and Families, Office of Planning, Research, and Evaluation website at http://www.acf.hhs.gov/programs/opre/hs/impact_study/index.html.

Chapter 2: Study Design and Implementation

This chapter describes the design of the Head Start Impact Study, including procedures used to select the study sample and randomly assign eligible children, characteristics of the study sample, data collection procedures, and analysis methods used to derive the impact findings found in subsequent chapters. This chapter provides an overview of the study design. Additional details are provided in the Technical Report for the Head Start Impact Study. Copies of all data collection instruments can be found on the OPRE website: http://www.acf.hhs.gov/programs/opre/hs/impact_study/index.htm

The Experimental Design

As discussed in Chapter 1, the primary purpose of the Head Start Impact Study is to determine whether Head Start has impacts on participating children and their parents and whether any impacts vary among different types of children and families. By impact we mean a difference between the outcomes observed for Head Start participants and what would have been observed for these same individuals had they not participated in Head Start.

The critical question in designing this study was: "How do we determine what outcomes would have been observed if the children had not participated in Head Start?" In many studies, researchers have addressed this problem by comparing program participants to a similar group of children who, in the ordinary course of events, do not participate in Head Start. These non-participants might be drawn from a waiting list of children who applied to Head Start but did not receive services or from a group of low-income children who didn't apply to Head Start. However, by comparing families based on the actions that they took (e.g., attending Head Start or not, applying to Head Start earlier or later), even the best attempts at constructing such a comparable group of non-participants suffer from what evaluators call "selection bias." That is, families who seek out, or "select," Head Start for their children—or select it earlier—are likely to be different in important ways from those who do not, and thus, their children may have different outcomes independent of any effect of Head Start services. Because all of these differences cannot be accounted for, there is a risk of misinterpreting observed differences on a

particular outcome (e.g., emergent literacy) as a program impact when they may instead reflect intrinsic differences between participant and non-participant children and families.

To avoid this problem of selection bias, the Head Start Impact Study randomly assigned a sample of 3- and 4-year-old Head Start applicants not previously served by the program,[21] either to the Head Start group or to a control group. The Head Start group was allowed to enroll in Head Start, while the control group was not granted access to Head Start (but may have received similar services through other available programs chosen by their parents). Under this randomized design, a comparison of outcomes for the two groups yields an unbiased estimate of the impact of Head Start on children's school readiness and subsequent school success. If random assignment is properly implemented with a sufficient sample size, the two groups should not differ on average at time of random assignment. The only difference between the two groups from then on will be their access to Head Start services, and, therefore,[22] any differences observed in average outcomes after random assignment can be attributed to the effects of Head Start.

Sample Selection and Random Assignment

Most randomized studies are conducted in small demonstration programs or only in a small number of operating sites, usually those that volunteer to be included in the research. In contrast, the Head Start Impact Study is based on a nationally representative sample of both Head Start programs and children. First-time applicants to Head Start in fall 2002 were randomly selected from a nationally representative sample of Head Start programs.[23] This makes the study results generalizable to the Head Start program, not just the selected study sample. This approach responds both to the congressional mandate and to the recommendations of the Advisory Committee. Congress required that the study exclude Head Start programs specifically targeting special populations, thus, as discussed in Chapter 1, the study does not

---

[21] The Head Start Impact Study focuses on newly entering children to ensure that the estimated impacts are unaffected by previous program participation. Consequently, children who were returning to Head Start, as well as those previously enrolled in Early Head Start, were excluded from the study sample.

[22] More precisely, there will be differences between individuals in the two groups, but the expected value of these differences is zero except through the influence of Head Start (i.e., selection bias is removed by random assignment).

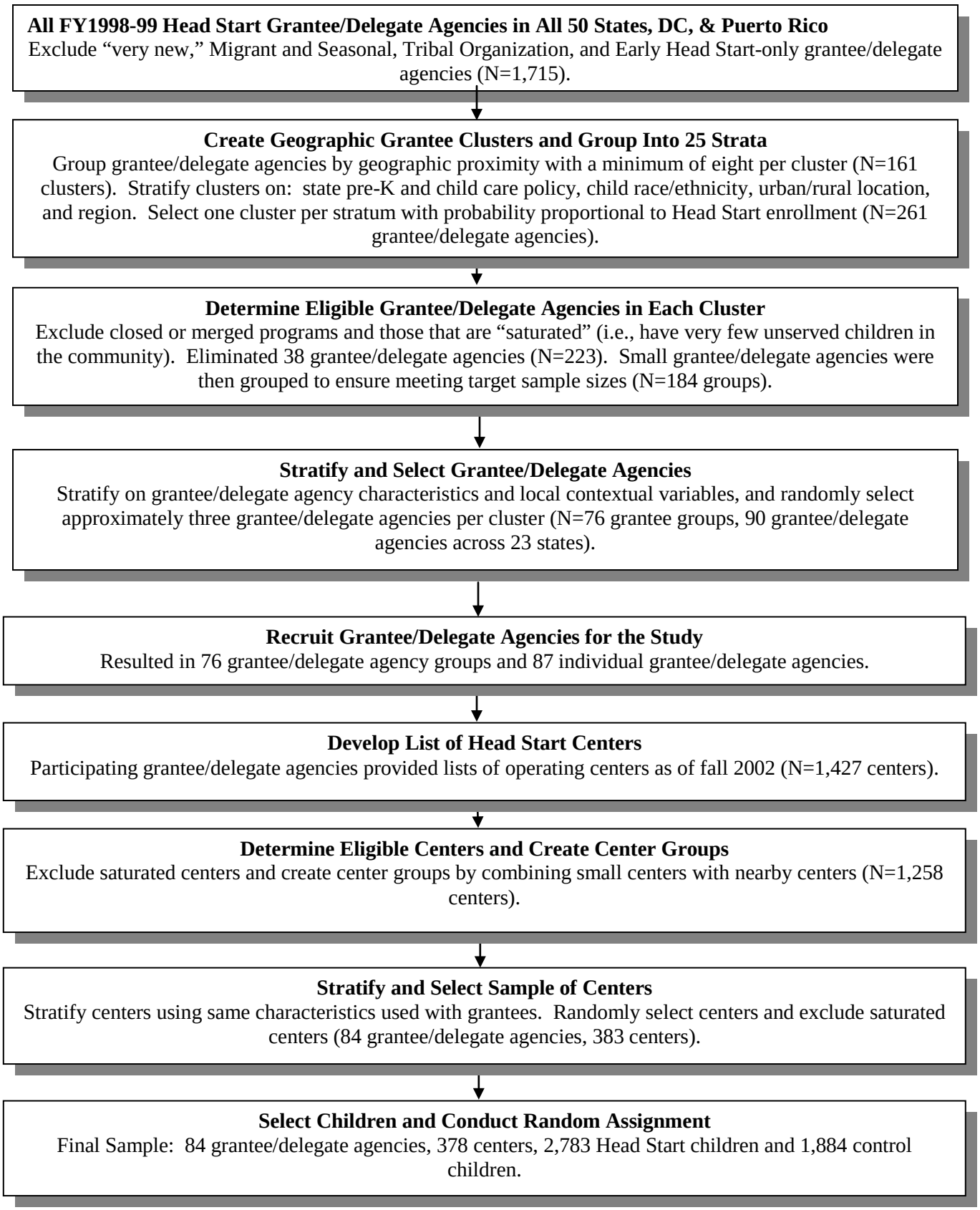
[23] Certain exclusions were made from the universe of all Head Start grantees in the country for reasons described below.

include the Migrant and Seasonal Head Start programs, Early Head Start (which serves pregnant women and children from birth to age three), or the American Indian/Alaska Native tribal programs. Programs in Puerto Rico were included in the study, but findings for these programs are reported separately. The study also excluded programs in communities that did not have more children eligible for Head Start than could be served with the existing number of funded slots. This constraint ensured that the study's need for a control group did not require slots to go unfilled. The study used a multi-stage sampling process to select a representative group of Head Start programs. The process, depicted in Exhibit 2.1, is described below:

1. **Identify grantee/delegate agencies.** The sampling process began by using the Head Start Program Information Report (PIR) to create a list of 1,715 Head Start grantee and delegate agencies operating in fiscal year (FY) 1998-99, after excluding (1) grantee/delegate agencies serving only special populations (migrant/seasonal and tribal Head Start programs, and sites serving only Early Head Start children), (2) grantees involved in the FACES 2000 study, and (3) as recommended in the Advisory Committee report (1999), grantees/delegate agencies that were "*extremely new to the program*."[24]
2. **Create, stratify, and select geographic clusters.** This pool of 1,715 Head Start programs was subsequently organized into 161 "geographic clusters" (to increase our ability to closely monitor random assignment and obtain high-quality data). The clusters were then grouped into 25 strata to ensure variation in factors such as region of the country, urban/rural location, race/ethnicity, and variation in state pre-kindergarten and child care policies. One cluster of programs was then randomly selected from each of the 25 strata with probability proportional to total enrollment. This resulted in a total of 261 grantee or delegate agencies in the sampled clusters (to improve efficiency, random subsampling was done in three very large urban clusters).
3. **Determine grantee/delegate agency eligibility.** To be eligible for inclusion in the study sample, grantee/delegate agencies had to have enough "extra" or additional newly entering applicants beyond their number of funded slots to allow for the creation of a non-Head Start control group. That is, the programs could not be serving all the eligible children in their community who wanted Head Start, a situation we refer to as "saturation." Ethically, random assignment could only be conducted in communities where Head Start programs were expected to be unable to serve all the eligible children seeking enrollment for fall 2002. This reduces the ability to generalize the results to some extent, as discussed later in the chapter. Eligibility was determined from information verified through telephone calls to all initially sampled 261 grantee/delegate agencies, augmented with information provided by Federal Regional Office staff and with data obtained from secondary sources such as local Child Care Resource and Referral Agencies, and the PIR. This

[24] Defined as in operation for fewer than two years.

Exhibit 2.1: Sample Selection Process for the Head Start Impact Study



screening process eliminated 28 grantees/delegate agencies (a reduction of 11 percent) found to be operating in saturated communities. Additionally, ten other grantee/delegate agencies had been closed or merged, further reducing the pool of eligible programs to 223 grantee/delegate agencies.

4. **Stratify and select grantee/delegate agencies.** Under a PPS (Probability Proportional to Size) sample design, the largest programs have the highest probability of being selected. To ensure the inclusion of the full range of Head Start grantee/ delegate agencies, smaller programs were combined with other agencies in the same cluster to form "grantee/delegate agency groups." The single grantee/delegate agencies, and the formed groups, were then stratified along several dimensions to ensure that programs selected represented the following conditions: urban location (central city, other urban, rural/small town), auspice (school based versus all other agency types), percentage Hispanic and percentage African American enrollment, program options offered (part-day only, full-day only, both), and the percentage of total enrollment represented by newly entering 3-year-olds. Approximately three grantee/delegate agencies or groups were randomly selected from each of the 25 strata with probabilities proportional to the number of newly entering children. This yielded a sample of 76 grantee/delegate agencies or groups comprising 90 individual grantee/delegate agencies, across 23 states.

5. **Recruit grantee/delegate agencies.** Senior project staff visited all 90 selected grantee/delegate agencies during summer 2001 to explain the study, verify information needed for study implementation, and to gain their agreement to participate in the Head Start Impact Study. Three agencies were dropped at this point—one had recently closed, and two were dropped due to an overlap with a study being conducted by the federally funded Head Start Quality Research Centers[25] Consortium—leaving 87 grantee/delegate agencies in 76 grantee/delegate agency groups (i.e., the overall number of grantee/delegate agency groups was not reduced).

6. **Develop list of Head Start centers.** Because administrative data do not identify individual Head Start centers, each of the 87 grantee/delegate agencies was asked to provide a list of all centers expected to be in operation for the 2002-03 program year and to validate basic data about the characteristics of children served, program options, and enrollment patterns in each center. This resulted in a list of 1,427 Head Start centers in the 87 grantee/delegate agencies (76 grantee groups) that could potentially be included in the Head Start Impact Study.

7. **Determine eligible centers and create center groups.** The center-level data were first used to eliminate 169 centers determined to be "saturated," as was done previously for grantee/delegate agencies. This step reduced the total eligible pool of centers from 1,427 to 1,258 across 84 separate grantee/delegate agencies in 76 grantee/delegate agency groups (a reduction of 11.8 percent and the loss of three grantee/delegate agencies, but no grantee groups). Next, small centers were combined with nearby centers to create "center groups."

[25] The Head Start Bureau (HSB) and the Office of Program, Research and Evaluation (OPRE) of DHHS awarded eight cooperative agreements under the Head Start Quality Research Center (QRC) Consortium II (2001-06) to study promoting approaches to the school readiness of Head Start children.

8. **Stratify and select a sample of study centers.** The resulting "center groups" were then stratified using the same characteristics used for the selection of grantee/delegate agencies (excepting those that do not vary within grantee/delegate agencies such as a region). Three center groups were selected from each eligible grantee/delegate agency, resulting in a main sample of 448 centers in 84 grantee/delegate agencies.

   More in-depth or up-to-date information on the initially sampled centers led to a determination that some were, in fact, ineligible for inclusion in the study. These included centers that: (1) had recently closed or had been merged with other centers; (2) served only Early Head Start children; (3) were in collaborations between Head Start and private preschool programs that could not subject their entire pool of applicants to random assignment; or (4) were, in fact, saturating their community with Head Start services. These findings resulted in the dropping of 103 initially sampled centers, but the addition of 38 replacement centers[26] to yield a final sample of 383 Head Start centers.

9. **Select children and conduct random assignment.** The Head Start grantee/delegate agencies and centers, when properly weighted, was designed to yield a sample of children that represented the national population of newly entering children and their families (with the exclusions noted above) for the 2002-03 program year. The sample of children included 2,783 Head Start children and 1,884 control children. Details on random assignment are described below.

For more details on the sample selection, analysis weights, and variance calculations, see the Technical Report for the Head Start Impact Study.

Random Assignment

At each of the selected Head Start centers, program staff provided study information to parents at the time enrollment applications were distributed. Parents were told that enrollment procedures would be different for the 2002-03 Head Start year and that some decisions regarding enrollment would be made using a "lottery-like" process. Children randomly assigned to the non-Head Start group were not to be admitted to Head Start during 2002-03. Those who were in the 3-year-old group, however, were told that they could re-apply for Head Start in 2003-04 and might be admitted if eligible.

Study staff worked with grantee/delegate agencies to ensure that parents received study information. These staff obtained data on all applications for the 2002-03 program year (to ensure that all applicants were considered for random assignment). Returning children, and a

[26] A "reserve" sample of an average of two center groups per program (a total of 237 centers) was also selected to be used as replacement sites if needed to achieve the expected overall study sample size of children. Thirty-eight of these centers were used. The final sample was 383 (448-103+38) centers.

small number of grantee-requested "high-risk" exclusions,[27] were eliminated from consideration for the study. Examples of such exclusions included children of homeless families, children in families with documented abuse and neglect, and children with severe disabilities, especially those disabilities that would make it difficult to assess these children's outcomes for the study (e.g., blindness). Each grantee was limited to one exclusion per center. In fact only 276 exclusions were taken out of approximately 18,000 applications.

At this point, 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 admission criteria. Study staff recorded basic information about each applicant and what was usually a numerical score determined by local staff that signified the relative need of individual children (e.g., in some agencies, a higher score indicated a greater need for Head Start and a corresponding higher priority for admission). Using these rankings, the list of newly entering children who would ordinarily have been enrolled was extended to add a specified number of children needed for the control group. The children added were those who would normally be next in line for admission if the initially targeted children could not be enrolled. Children were randomly selected from the entire list.

The goal was to randomly select, on average, 27 children from the expanded list at each of the sampled centers or center groups: 16 to be assigned to the Head Start group and 11 to be assigned to the control group. For an average center group, the 11 control group children represented about nine percent of total enrollment. In some cases, where fewer children than expected were actually available, a smaller sample of children was selected for the study.

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.

---

[27] This decision was made because: (1) there were ethical concerns about assigning very high-risk children to the control group, especially in situations where Head Start may provide their only option for early childhood services; (2) a previously conducted study demonstrated that the potential exclusion of those most severely in need affected cooperation when trying to recruit study sites; and (3) there were some children who could not be assigned to the control group because of placement by the local child welfare agency.

Consequently, the study included two separate samples: a newly entering 3-year-old group (to be studied through two years of Head Start participation, kindergarten and 1st grade), and a newly entering 4-year-old group (to be studied through one year of Head Start participation, kindergarten and 1st grade). The 3-year-old group was slightly larger than the 4-year-old group to protect against the likelihood of higher study attrition resulting from an additional year of longitudinal data collection for the younger children.[28]

Within the final set of 76 grantee/delegate agency groups (or 84 total grantees/delegate agencies), random assignment was attempted at a total of 383 randomly selected Head Start centers. Of these, random assignment could not be completed in only five centers (or 1.3 percent), resulting in a final sample of 378 centers with successful random assignment.[29] However, as noted above, the full desired sample could not be obtained at each center, resulting in the following situations:

- **Obtained Full Sample.** Random assignment was completed at 173 Head Start centers (45 percent) that provided the full expected sample of children.
- **Obtained Smaller Sample.** Random assignment was completed at 150 Head Start centers (39 percent) that provided a smaller than expected sample (i.e., because new application rates were lower than estimated).
- **Obtained Larger Sample.** Random assignment was completed at 55 Head Start centers (16 percent) for a larger number of children than originally planned, both to take advantage of situations where enrollments of new children were higher than expected and to compensate for other centers where new enrollments were unexpectedly low.

In total, 4,667 newly entering children were randomly assigned and included in the Head Start Impact Study (see Exhibit 2.2).

[28] This roughly equal sampling of 3- and 4-year-old applicants was done to obtain reliable estimates of program impacts for each age cohort, despite the fact that 4-year-olds represent about twice the proportion of all Head Start participants as do 3-year-olds. In large part, this is because the total of all 4-year-old participants includes both newly entering 4-year-olds plus returning children who began Head Start as 3-year-olds and who have turned 4 years of age in their second year of program participation.

[29] The five centers were excluded due to center closures and mergers.

Exhibit 2.2: Number of Children Randomly Assigned to Head Start and Control Groups, by Age Cohort

Age Cohort	Head Start Group	Control Group	Total Sample
3-year-olds	1,530	1,029	2,559
4-year-olds	1,253	855	2,108
Total	**2,783**	**1,884**	**4,667**

As indicated above, about 60 percent of the sample was assigned to the Head Start group, and about 40 percent was assigned to the control group. This imbalance reduces the precision of the impact estimates by less than two percent (compared to a balanced 50-50 design). However, it provided several important benefits: (1) it significantly increased the ability to recruit Head Start grantees and centers by decreasing the number of extra children needed for the control group, (2) the loss of sites due to saturation was decreased, and (3) the cost of data collection was decreased because Head Start group members require less effort to track and interview over time than children in the control group.

Representing the National Head Start Population

The population of interest was all newly entering 3- and 4-year-olds in all Head Start centers operating in 2002-03, except those serving only special populations (i.e., programs serving primarily only Migrant or Seasonal Farmworkers and their families, American Indian or Alaskan Native tribal populations, or Early Head Start children), or very new centers. Ideally, all such children would have the possibility of being included in the study, and the "coverage rate" would, therefore, be equal to 100 percent of all Head Start participants.

The major cause for undercoverage in the study was the requirement that the selected Head Start grantee/delegate agencies and centers have more eligible applicants than could be served at their current Federal funding level, to allow the creation of a control group. Programs that were serving essentially all the eligible children in the community (referred to as "saturated" programs or centers) could not be included in the study because creating control groups at these sites could have resulted in a reduction in the number of children being served by Head Start.

As noted above, there were four points in the sample selection process where grantee/ delegate agencies or centers were lost due to such saturation. First, some Head Start grantee/ delegate agencies were determined to be saturated before the sample was selected, and these

programs were, therefore, dropped from the sampling frame. Second, after the initial sample of grantees/delegate agencies was selected, some additional programs were found to be saturated and were also deleted from the sample. At this same point in the process, two additional programs were dropped from the sample because they were Head Start Quality Research Centers (QRC) and were excluded so as not to be overburdened. The third point at which saturated sites were dropped from the sample was during the selection of Head Start centers. As with grantees/ delegate agencies, some centers were initially determined to be saturated and were considered to be ineligible for inclusion and deleted from the study sample. Some centers were determined to be saturated during later attempts to conduct random assignment and also had to be dropped from the study sample.

Taking into account all of these opportunities for Head Start grantees/delegate agencies and centers to be deleted, the estimated weighted national coverage rate[30] for spring 2003 data was 84.5 percent, meaning that the study sample was representative of 84.5 percent of the total universe of all newly entering 3- and 4-year-olds across the country. (The small number of grantees/delegate agencies and centers that was found to be closed or merged into another program or center was properly considered as ineligible, not as non-covered.)

As discussed above, additional under-coverage of children occurred because grantee-requested "high-risk" children were excluded from the study. The coverage rate of 84.5 percent cited above does not account for these few exclusions. These exclusions have negligible effect on the overall coverage rate, however, as there were only 276 exclusions out of approximately 18,000 applications received in the targeted programs.

Finally, there was some under-coverage due to a number of sampled centers found to be "partially saturated," that is, there were enough applicants at the center to permit some children to be assigned to the control group, but the number available was insufficient to allow the selection of the full targeted sample. In such situations, available children were sampled and either (a) additional treatment and control group children were selected from another "reserve"

[30] The weight that is used for this estimate accounts for the probability of selection for each program and center and also weights the contribution of programs and centers according to the size of their enrollment. An unweighted coverage rate can also be calculated, but this is a less useful measure of coverage as it estimates the proportion of children in the *sample*, not the universe of children served by Head Start nationally who are in programs and centers that are not saturated.

center, or/and (b) a larger sample of children was selected from another already included center in the same geographic cluster to make up for the shortage of study children.[31]

The Success of Random Assignment

The extent to which random assignment was successful is assessed from two perspectives. First, the characteristics of children randomly assigned to the Head Start and non-Head Start groups are compared using information collected for each child at the time of random assignment. Then, the extent to which children complied with their assigned status is examined, i.e., to what extent did children assigned to the Head Start group actually receive some Head Start services, and to what extent did children assigned to the control group receive any Head Start services?

Comparing Head Start and Non-Head Start Children at Baseline

Exhibit 2.3 provides, separately for the 3- and 4-year-old age groups, a comparison of children randomly assigned to the Head Start and control groups using weighted data[32] on all characteristics that were measured and available at the time of random assignment. These data were drawn from parental applications for Head Start.

As the exhibit shows, there are no statistically significant differences between the two randomly assigned groups. This suggests that the initial randomization was done with high integrity and that the samples can provide the necessary confidence in the validity of the impact estimates.

Although not related to the success of random assignment, it is interesting to note that 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. This difference shows that newly entering 3-year-olds were relatively evenly distributed between the Black and Hispanic groups (32.8% vs. 37.4%), while about half of newly entering 4-year-olds were Hispanic (51.6% vs. 17.5% Black). This distribution for newly entering 4-year-olds is

---

[31] Weighting procedures were used to account for the under-coverage attributable to these factors. Details can be found in the Technical Report for the Head Start Impact Study.

[32] The weights used are the same as those used for all the analyses discussed in this report. Details are provided in the Technical Report for the Head Start Impact Study.

Exhibit 2.3: Comparison of Head Start and Control Groups: Child and Family Characteristics Measured Prior to Random Assignment (Weighted Data)

Characteristic	Head Start Group	Control Group	Difference: Head Start – Control
Child Gender:			
3-Year-Old Cohort			
Boys	48.5%	48.9%	-0.4%
Girls	51.5%	51.1%	0.4%
4-Year-Old Cohort			
Boys	51.1%	49.4%	1.7%
Girls	48.9%	50.6%	-1.7%
Child Race/Ethnicity:			
3-Year-Old Cohort			
White	24.5%	26.6%	-2.1%
Black	32.8%	31.8%	1.1%
Hispanic	37.4%	35.7%	1.6%
Other	5.3%	5.9%	-0.6%
4-Year-Old Cohort			
White	26.7%	23.3%	3.4%
Black	17.5%	17.0%	0.5%
Hispanic	51.6%	53.8%	-2.1%
Other	4.1%	5.9%	-1.8%
Child Language:			
3-Year-Old Cohort			
English	71.1%	69.9%	1.2%
Spanish	24.8%	24.0%	0.8%
Other	3.9%	5.7%	-1.8%
Missing	0.2%	0.4%	-0.2%
4-Year-Old Cohort			
English	57.1%	56.4%	0.8%
Spanish	39.3%	40.8%	-1.5%
Other	3.2%	2.3%	0.8%
Missing	0.4%	0.5%	-0.1%
Parent Language:			
3-Year-Old Cohort			
English	74.8%	74.8%	0.0%
Spanish	23.1%	22.0%	1.1%
Other	1.5%	1.7%	-0.2%
Missing	0.6%	1.5%	-0.9%
4-Year-Old Cohort			
English	59.5%	58.4%	1.1%
Spanish	37.8%	39.5%	-1.7%
Other	0.9%	0.5%	0.5%
Missing	1.8%	1.6%	0.2%
Child Income Eligible:			
3-Year-Old Cohort			
No	7.7%	6.7%	1.0%
Yes	91.4%	91.9%	-0.6%
Missing	0.9%	1.4%	-0.5%
4-Year-Old Cohort			
No	6.0%	10.1%	-4.0%
Yes	91.8%	87.9%	3.9%
Missing	2.2%	2.1%	0.1%

Notes: (1) Due to rounding, the sum of the percents may not equal 100 percent; (2) Data source: Roster information used at time of random assignment; (3) *t*-tests of the difference between the Head Start and non-Head Start percentage in each row were run for each characteristic; no statistically significant differences were found. With large samples, differences in means for 0/1 variables (e.g., 1=boys, 0=girls) have approximately normal distributions and follow the *t* distribution once divided by their standard errors.

similar to the 4-year-old distribution in data from the Head Start National Reporting System (HSNRS), 2003 data.[33] This ethnic difference is also reflected in the age-group differences in child and parent language.

Deviations From Random Assignment

Random assignment rarely results in perfect adherence to the assigned program status. In the current study, one would expect some children assigned to the Head Start group not to participate in the program (referred to as "no-shows"), and some of the children assigned to the non-Head Start group to enroll in the program (referred to as "crossovers"). During program recruitment, Head Start grantees and centers described no-shows as a common occurrence in ordinary program operations, with rates among enrolled children often in the double-digits. Consequently, it is not surprising that some families who were randomly assigned to the Head Start group subsequently opted for a different care setting for their child.[34]

Similarly, although every effort was made to maintain the integrity of the control group, perfect conditions could not be implemented. In some instances, local staff intentionally enrolled control group children into Head Start. More commonly, parents simply applied to another nearby Head Start program, especially in densely populated areas with Head Start programs operating in proximity. Due to confidentiality restrictions, information on study participants was not shared with programs not involved in the study, so control group families were not prevented from being served by other Head Start programs.

For analysis purposes (as explained below), it is only the degree of compliance with the random assignment design *in the first year of the study* that matters, since this was the one year in which the study sought to have all Head Start group children—and none of the control group children—participate in Head Start. Exhibit 2.4 provides information on the incidence of Head Start group no-shows and control group crossovers by age group in that year. In the exhibit, a child in the Head Start group is considered a no-show if it was determined that he/she did not participate in Head Start at any time during the 2002-03 program year. A child in the control

---

[33] The HSNRS gathered information about the progress of approximately 430,000 Head Start children in the areas of early literacy and numeracy skills at the beginning and end of the Head Start year. Data for this analysis come from HSNRS 2003.

[34] Chapter 3 presents a breakdown of the types of settings children attended.

group was deemed a crossover if he/she participated in Head Start at any time during the 2002-03 program year. This determination was based on information from parent surveys, checking Head Start enrollment in fall 2002, and the care setting identified at the time of the child's fall and spring assessments. No-shows accounted for 15 and 20 percent of the full randomly assigned Head Start sample for children in the 3- and 4-year-old cohorts, respectively; properly weighted, crossovers accounted for 17 and 14 percent of the randomly assigned control group.

Exhibit 2.4: The Incidence of No-Show and Crossover Behavior for the Sample as Randomly Assigned, by Age Cohort (Weighted Data)

Sample Group	Some Year 1 Head Start Participation	No Year 1 Head Start Participation	Total
All Randomly Assigned (N=4,667):			
3-Year-Old Cohort			
Head Start Group	85.1%	14.9%	100%
Control Group	17.3%	82.7%	100%
4-Year-Old Cohort			
Head Start Group	79.8%	20.2%	100%
Control Group	13.9%	86.1%	100%

Subsequent sections of this chapter describe how impact estimates are adjusted to account for these occurrences. Violations of random assignment that extend Head Start's services to some children in the control group and reduce the exposure to Head Start among the treatment group make it harder to detect any impact of the program with the available sample size. Estimates of the size of Head Start's effects on participants will also be biased downward absent special adjustment. Because the bias is downward, we have full confidence that statistically significant impacts are real and important. The downside, of course, is that some true impacts of Head Start may be overlooked because of the bias. To address this, the impact analyses presented throughout the report show the magnitude of estimated impacts with and without adjustments for the no-shows and crossovers in the sample (these adjustments are described in Chapters 4, 5, 6, and 7 as well as in the Technical Report for the Head Start Impact Study).

Data Collection and Data Sources

Data collection began in fall of 2002[35] and continued through the spring of 2006, following children from entry into Head Start through the end of the preschool years, end of kindergarten, and end of 1st grade.

The data collection procedures and measures used are summarized below. Additional details can be found in the Technical Report for the Head Start Impact Study.

Data collection included the following components.

- **Direct Child Assessments.** Child assessments arguably provide the best and most direct measures of the cognitive development of study children and the extent to which they are educationally ready for success in school. The child assessment battery used in the Head Start Impact Study focused on language and literacy, including children's vocabulary knowledge, reading and writing skills and achievement, oral comprehension and phonological awareness, and math skills and achievement. The 45- to 60-minute child assessment battery was typically administered one-on-one by specially trained assessors in the child's main care setting during the preschool years (i.e., where the child spent the most time Monday through Friday between the hours of 9 am and 3 pm) and in the child's home during the kindergarten and 1st grade years.

  At the start of the study in fall 2002, information was collected on each child's language ability. To determine a child's language for assessment, assessors asked the child's main child care provider three questions to determine the language that would be used to assess the child: (1) What language does child speak most often at home; (2) What language does the child speak most often at this child care setting; and (3) What language does it appear this child prefers to speak? Children were tested in the language in which at least two of the three responses were the same. The language decision form is presented in Appendix C.

  For children requiring assessment in Spanish, the assessor administered a bilingual child assessment in fall 2002 that included the complete Spanish assessment battery and two English tests the Peabody Picture Vocabulary Test (PPVT) and the Woodcock-Johnson III (WJIII) Letter-Word Identification test. In spring 2003, and in all subsequent data collection periods, those initially identified Dual Language

[35] Fall 2002 data collection was completed between the end of September and mid-November for the majority of children and parents (although a small number did extend into December). The discussion of analysis procedures in this chapter and in the Technical Report for the Head Start Impact Study detail how this late baseline data collection is handled in the analysis of program impacts.

Learners (DLL)[36] were given the complete English assessment battery plus two Spanish tests—the Test de Vocabulario en Imágenes Peabody (TVIP) and the Batería Woodcock-Muñoz (WM) Identificacón de letras y palabras. One exception to this rule was Puerto Rico where, because all instruction is in Spanish, children were assessed with the complete Spanish assessment battery at each data collection point. If the child's primary language was other than English or Spanish (e.g., Creole, Arabic), the assessor asked the main care provider if the child could understand and answer questions in English. If yes, the child was assessed using the English assessment battery. If no, and the assessor was not fluent in the child's language, an interpreter was used. In fall 2002, four tests (McCarthy Draw-A-Design, Color Names and Counting, Leiter-R (adapted), and Story and Print Concepts) were administered to these children. In spring 2003, and in all subsequent data collection periods, these children were all tested using the complete English assessment battery.

- **Parent Interviews.** In-person interviews were typically conducted in the home of each study child with a parent or primary caregiver living with, and responsible for raising, the child at the fall 2002 baseline point and at each of the subsequent spring 2003, 2004, 2005, and 2006 follow-up data collection waves. It was possible that the parent or primary caregiver could change over time, but this occurred for a very small percentage of the children. Parent interviews were available in both English and Spanish versions, and bilingual English/Spanish speakers were hired for areas with Spanish-speaking families. For other languages, either interviewers/assessors fluent in these languages were hired or other local resources were asked to identify interpreters to aid in completing the parent interviews.

  Information collected during the interviews included: (1) parents' report of a variety of child-specific information, including the child's demographic characteristics, behavior, developmental accomplishments, and disabilities; (2) parental characteristics such as education, employment, and reported depressive symptoms; (3) household characteristics, such as household risk, household members and income; (4) parent-child activities and interactions such as reading to the child; (5) parenting practices such as safety practices and parenting styles; (6) the child's experiences during preschool and early elementary school years, including parent communication and involvement with school; and (7) community characteristics such as urbanicity.

  In addition, in the winter of 2003, and in the fall of each subsequent year, a 10-minute telephone interview was conducted with the parent/primary caregivers to obtain up-to-date contact information and information regarding the child's current preschool, child care, or school placement to determine the appropriate setting for the spring data collection waves. If parents could not be reached by telephone, in-person interviews were conducted to collect this information.

[36] Dual Language Learners (DLL) are children learning two (or more) languages at the same time, as well as those learning a second language while continuing to develop their first (or home) language. These children are also often referred to as Limited English Proficient (LEP), bilingual, English language learners (ELL), English learners, and children who speak a language other than English (LOTE).

- **Teacher Surveys and Child Ratings.** Additional information was obtained from teachers and other care providers (e.g., family day care providers) who completed self-administered questionnaires to rate each of the study children who were in their classroom or care (Teacher/ Care Providers' Child Reports). Teachers also completed questionnaires, and care providers were interviewed in person, to obtain information about them, the nature of the setting in which they worked, and the types of services they provided to the selected study children. Each of these activities is described below:

  - **Teacher's/Care Provider's Child Reports (TCRs).** Teacher/other care provider ratings of children's accomplishments and behavior are an important source of information about children's learning and behavior because teachers and care providers see children over extended periods of time in different settings, providing for appraisals of children's skills and competence in those settings. Moreover, these reports—while not as objective as direct assessment or observations by impartial observers—can be important in and of themselves because they influence the way these individuals interact with the children. During the preschool years, teachers and other care providers were asked to rate each of the children in their classroom or care who were participating in the study. Ratings of teacher/provider relationship with the child, child's behavior, and child's classroom performance were provided. In kindergarten and 1st grade, teachers were asked to rate each of the study children in their classroom on their relationship with the child, the child's conduct in their classroom, academic skills, school accomplishments, and health and developmental concerns.

  - **Teacher Surveys and Care Provider Surveys.** During the preschool years, Head Start teachers and teachers in other center-based programs were asked to complete a survey. The survey included questions on teacher demographics, such as education and years of experience, curriculum used, type and frequency of language and math activities used in the classroom, mentoring, parental contact methods, and beliefs about working with and teaching children. To obtain comparable data from children not in center-based programs, an "other care provider" interview was used. In addition to the information included in the teacher survey, the other care provider interview collected additional information on the types of services available to the study children in their care. In kindergarten and 1st grade, the teacher survey included questions about teacher's characteristics (e.g., training and education), the classroom environment (e.g., number of children, race/ethnicity of children in the class, the number of children eligible for free or reduced-price lunch, classroom attendance and behavior), and the type and frequency of language and math activities used in the classroom.

- **Head Start and Elementary School Experiences.** Information was obtained on the experiences of children and the services they received during their preschool years (when they were in Head Start or other child care environments), as well as during their kindergarten and 1st grade years.

  For the preschool year, in-person interviews were conducted with directors of the Head Start and non-Head Start centers that study children attended. To further measure quality of care, direct observations of classrooms and family day care homes

were conducted. The teacher survey (described above) also provided information on the teacher qualifications and classroom environments that children attended.

- **Center Director Setting Interviews.** This in-person interview was used to collect information on the operation and quality of Head Start and non-Head Start center-based programs. Issues addressed in this interview included staffing and recruitment, teacher education initiatives and staff training, parent involvement, curriculum, classroom activities and assessment, home visits, kindergarten transition, and demographic information about the director.
- **Care Setting Observation.** Direct observations of care setting and quality were used for children in center-based and family day care home programs, including those participating in Head Start. These tools provide direct measures of the extent to which Head Start centers, and other child care programs, employ skilled teachers and provide developmentally appropriate environments and curricula for their pupils. Trained observers conducted observations in classrooms and centers attended by the sampled children. Observers spent four hours in each class to ensure observation of a major portion of the daily schedule and a variety of classroom and center activities. For the elementary school years, data about the school environment were gathered from secondary data sources as well as through teacher survey information. The secondary data sources included:
  - The Department of Education's **Common Core of Data (CCD)** (www.nces.ed.gov/ccd) and **Private School Universe Survey (PSS)** (www.nces.ed.gov/surveys/pss) were used for poverty and race/ethnicity distributions in public and private schools, for the kindergarten and 1st grade years.
  - The Great Schools Database (www.greatschools.net) was used for data on school reading and math proficiency levels for the kindergarten and 1st grade years.

Response Rates

Exhibits 2.5 (4-year-old cohort) and 2.6 (3-year-old cohort) present response rates for all study instruments administered over the entire data collection period from fall 2002 through spring 2006. Some instruments (indicated as NA) were not administered in all data collection periods. Parent and child assessment response rates represent the number of interviews/ assessments completed, i.e., the percentage of the randomly assigned sampled population that completed the interview or assessment. For each year, the response rate is calculated on the entire randomly assigned sampled population, not just on those with completed responses the previous year.

Exhibit 2.5: Treatment and Control Response Rates for All Study Years and for All Study Instruments for the 4-Year-Old Cohort

Instruments	Fall 2002 (Baseline)		Spring 2003 (End of Head Start Year)		Spring 2004 (Kinder-garten)		Spring 2005 (1st Grade)	
	T	C	T	C	T	C	T	C
Child Assessment	86%	77%	87%	77%	81%	74%	79%	73%
Parent Interview	90%	84%	85%	79%	82%	75%	82%	75%
Teacher/Care Provider Survey*	NA	NA	90%	70%	64%	68%	78%	81%
Teacher/Care Provider's Child Reports*	NA	NA	90%	70%	64%	68%	78%	81%
Center Director Interviews	NA	NA	91%	73%	NA	NA	NA	NA
Classroom Observations*	NA	NA	92%	68%	NA	NA	NA	NA

*Base for these response rates is those children with both a parent survey and child assessment.

Exhibit 2.6: Treatment and Control Response Rates for All Study Years and for All Study Instruments for the 3-Year-Old Cohort

Instruments	Fall 2002 (Baseline)		Spring 2003 (End of Head Start Year)		Spring 2004 (End of Age 4 Year)		Spring 2005 (Kinder-garten)		Spring 2006 (1st Grade)	
	T	C	T	C	T	C	T	C	T	C
Child Assessment	87%	76%	89%	80%	87%	79%	82%	77%	81%	74%
Parent Interview	93%	84%	88%	81%	86%	79%	85%	79%	85%	76%
Teacher/Care Provider Survey*	NA	NA	88%	64%	87%	79%	82%	84%	86%	88%
Teacher/Care Provider's Child Reports*	NA	NA	88%	64%	87%	79%	82%	84%	86%	88%
Center Director Interviews	NA	NA	86%	81%	78%	73%	NA	NA	NA	NA
Classroom Observations*	NA	NA	91%	66%	87%	84%	NA	NA	NA	NA

*Base for these response rates is those children with both a parent survey and child assessment.

Data from the Teacher/Care Provider surveys, Teacher/Care Provider child reports, Center Director interviews, and classroom observation response rates are calculated at the child level and are conditioned on the percentage of children for whom there is both a parent interview and child assessment.[37] Furthermore, in the Head Start years (2003 and 2004 for the 3-year-old cohort and 2003 for the 4-year-old cohort), response rates for children's teachers, directors, and classrooms are calculated only for children who experienced non-parental care, since children in parental care did not have respondents to these instruments. Further information about response rates is provided in the Technical Report for the Head Start Impact Study.

There were some differences in response rates between the Head Start and control groups for both cohorts. Control group response rates for parent interviews and child assessments were typically about seven to eight percentage points lower than Head Start group response rates. For other instruments—teacher survey and teacher child report—the pattern runs the other way, with response rates slightly higher for control group children once children enter school. The results with the greatest unevenness are for the teacher instruments and classroom observations in the first preschool years, where Head Start group response rates exceed those of the control group by approximately 20 percentage points. Response rates for these particular instruments were approximately 90 percent for the Head Start group as compared to approximately 70 percent for the control group.

Differential parent and child response rates led to few significant differences in the characteristics of the Head Start and control groups, as discussed later in this chapter. However, to address any disparities, as part of the weighting procedures, separate nonresponse adjustment factors were applied to all instruments. To the extent that nonrespondents and respondents within a weighting category have similar impacts from Head Start, the application of these adjustment factors reduces the bias in the impact estimate due to nonresponse. Further, the use of baseline covariates in the impact estimation adds further control over such nonresponse bias. A detailed discussion of nonresponse adjustment is provided in the Technical Report for the Head Start Impact Study.

---

[37] Response rates were conditioned on having both a child assessment and parent interview because a child's care setting could only be determined by parents providing information about where their child was receiving services. Additionally, parents needed to give permission for children to be assessed and for data to be collected about the children's classroom.

Measures

The measures used in this study, described in this section, fall into three categories: (1) child and family demographics collected at baseline that were used as covariates in the impact analyses and also used to form child and family subgroups, (2) child and family outcome measures, i.e., the variables on which program impacts were estimated, and (3) characteristics of the preschool and early elementary school experiences of the participating children. Each area is described below with details provided on individual measures used in these analyses.

Child and Family Demographics

The following measures were created from data collected at baseline and used as covariates, to create subgroups, for the impact analysis and to describe characteristics of children who attended one or two years of Head Start:

- **Child's race/ethnicity**—based on information provided by the fall 2002 parent interview where respondent was specifically asked about child's Spanish, Hispanic, or Latino origin and then also asked about the child's race. Missing information was supplemented with roster data provided by the child's Head Start center. The measure is presented in three categories, White/Other, Black, and Hispanic.
- **Child's gender**—based on information provided by the fall 2002 parent interview; missing information was supplemented with roster data provided by the child's Head Start center. Measure is presented as a dichotomous variable, male or female.
- **Individual education plan (IEP) status in spring**—based on parent report of whether or not child had an IEP. A dichotomous variable was created with zero for no IEP and one for having an IEP in spring 2003 in order to define the special needs subgroup. For analyses of which children returned for a second year of Head Start, change in IEP status from fall 2002 to spring 2003 was also examined.
- **Child's pre-academic skills**—based on whether the child scored in the lowest quartile of the study population on the Woodcock-Johnson III Pre-Academic Skills (comprising three tests, i.e., Letter-Word Identification, Spelling, and Applied Problems) at the time of the baseline assessment (fall 2002). Two subgroups were created using this test score, the child was in the lowest quartile group, or the child was not in the lowest quartile group.
- **Biological father lives in household**—based on responses to the spring 2003 parent interview asking whether child's biological father lived in the household. A dichotomous variable was created, and a "yes" indicated that the biological father lived in household and "no" indicated that he did not.

- **Grandparent lives in the household**—based on household delineation obtained from spring 2003 parent interview. Respondent was asked to delineate all people living in the household and how each person was related to the study child. A dichotomous variable was created, and a "yes" indicated that either or both grandparents lived in the household, and a "no" indicated that neither grandparent lived in the household.
- **Number of adults over 18 in the household**—based on household delineation obtained from spring 2003 parent interview. Respondent was asked to delineate all people living in the household and how that person was related to the study child. Respondent was also asked to provide the age of the person. A variable was created which counted all people over age 18 living in the household and the percentage of households with one, two, or three or more people over age 18 is presented.
- **Number of children under six in the household besides the study child**—based on household delineation obtained from spring 2003 parent interview. Respondent was asked to delineate all people living in the household and how each person was related to the study child. Respondent was also asked to provide the age of the person. A variable was created which counted all children under six living in the household and percentage of households with any other children younger than school age (not including study child) as compared to no other children younger than school age is presented.
- **Home language**—based on information provided by the fall 2002 parent interview where respondent was asked the language spoken most frequently to the study child at home. Missing information was supplemented with roster data provided by the child's Head Start center. A dichotomous variable of not English and English was created. The vast majority of non-English households was Spanish speaking.
- **Whether family moved in the past 12 months (from spring 2003-2004)**—based on information provided in the spring 2004 parent interview. Respondents were asked the number of moves the family had made in the last year. This period was used to cover the time closest to parent's decision time for a second year of Head Start. A dichotomous variable of no moves and one or more moves was developed.
- **Family monthly income range**—based on information provided in the spring 2003 parent interview. If respondent was unable to provide actual monthly income, he/she was asked to indicate where his/her income fell in one of seven categories—ranging from less than $250 a month to over $2500 a month.
- **Economic difficulty in the past three months**—based on information provided in the spring 2003 parent interview. A dichotomous variable (yes/no) was created based on whether parent answered "yes" to having difficulty in the past three months—paying rent, paying electric and heating bills, buying food for the family, buying clothes for the children. A yes to any of these responses was characterized as having economic difficulty.
- **Father's employment status**—based on father's employment status as reported in the spring 2003 parent interview. Responses were collapsed into three categories—full time (35 hours or more per week), part time, and not working.

- **Mother employment status**—based on mother's employment status as reported in the spring 2003 parent interview. Responses were collapsed into three categories—full time (35 hours or more per week), part time, and not working.
- **Biological mother recent immigrant**—based on response to question in the fall 2002 parent interview that asks the mother "How many years have you lived in the United States?" A recent immigrant was considered living in the United States for less than ten years. A dichotomous variable was created, and a "yes" meant that mother was a recent immigrant and a "no" meant not a recent immigrant.
- **Mother's age**—based on mother's date of birth, which was reported by the mother in the fall 2002 parent interview and then calculated as of the date of the interview. Mother's age was collapsed into four categories—under 20 years old, 20-30 years old, 31-40 years old, and over 40 years old.
- **Mother teenager at birth of study child**—based on calculation of mother's age on study child's date of birth. If birth mother was under 20 years old when study child was born, then she was classified as a teenager at the birth of the study child.
- **Mother's marital status**—based on mothers' report in the spring 2003 parent interview. Mothers' responses were collapsed into three categories—never married, married, and separated/divorced/widowed.
- **Mother's highest level of education attained**—based on mothers' report in the spring 2003 parent interview. Mothers' responses were collapsed into three categories—less than high school, high school diploma or GED, and beyond high school.
- **Mother reported depressive symptoms**—determined from responses to the spring 2003 parent/caregiver interview using the shortened version (12 items) of the Center for Epidemiologic Studies-Depression scale (CES-D).[38] Four subgroups were created from the scale: (1) no depressive symptoms (score of 0-4), (2) mild depressive symptoms (score of 5-9), (3) moderate depressive symptoms (score of 10-14), and (4) severe depressive symptoms (score of 15-36).

Child and Family Outcome Measures

Outcome measures were developed in four domains—child cognitive development, child social-emotional development, health, and parenting practices. The selection of these domains was guided by several factors. First, it was important to measure the school readiness skills that are the focus of the Head Start program. The Head Start performance measures and conceptual framework (U.S. Department of Health and Human Services, 2001) indicate that children enrolled in Head Start should demonstrate improved emergent literacy, numeracy, and language

[38] Seligman, M.E.P. (1993). *What You Can Change…And What You Can't**. New York: Ballantine Books. The four depression categories are reported on page 101 in the above reference for the 20-item CES-D. The cut points were proportionately adjusted for the shortened version of the CES-D for use in Early Childhood Longitudinal Study-Birth Cohort (ECLS-B), Family and Child Experiences Survey (FACES), and HSIS.

skills. The framework also stresses that children should demonstrate positive attitudes toward learning and improved social and emotional well-being, as well as improved physical health and development.

Second, domains were selected to reflect the program's whole child model, i.e., school readiness is considered to be multi-faceted and comprising five dimensions of early learning: (1) physical well-being and motor development, (2) social and emotional development, (3) approaches toward learning, (4) language usage, and (5) cognition and general knowledge (Kagan, Moore, & Bredekamp, 1995). The whole child model also was recommended by the Goal One Technical Planning Group of the National Education Goals Panel (Goal One Technical Planning Group, 1991, 1993).

Third, in 2002, the National Institute of Child Health and Human Development (NICHD), the Administration on Children, Youth and Families (ACYF), and the Office of the Assistant Secretary for Planning and Evaluation (ASPE) within the U.S. Department of Health and Human Services (HHS) convened a panel of experts to discuss the state of measurement and assessment of learning and development in early childhood. Language, early literacy, and mathematics were the primary cognitive domains identified by the experts as important to early childhood development.

Based on these factors and the advice from the experts from the Head Start Impact Study working groups and Advisory Committees, measures were selected to assess the cognitive, social-emotional, and health outcomes of children. Considering the major emphasis Head Start places on parent education and involvement, a fourth domain, parenting practices, was also included. The selected measures are summarized below, organized by the four domains.

Cognitive Domain

The cognitive test battery consists of both standardized tests developed by recognized test publishing companies and non-standardized tests developed for use in the Head Start Family and Child Experiences (FACES) project. As the children developed, new tests were added to the child assessment battery; existing tests were extended to include more difficult items; and, in some cases, preschool-level tests were dropped as the children entered elementary school. Each of the tests is described briefly below; additional details on the assessments, including test

administration, adaptations, scoring, and reliability are provided in the Technical Report for the Head Start Impact Study.[39] Exhibit 2.7 presents all the cognitive domain measures from direct assessment and denotes the year in which they were administered. Measures related to educational performance, obtained from teacher and parent reports, are also included in this exhibit.

Language and Literacy: Vocabulary

- **Peabody Picture Vocabulary Test (PPVT), Third Edition.** The PPVT measures receptive vocabulary, i.e., listening comprehension for the spoken word in standard English (published reliability = 0.95). The child is instructed to look at four pictures and point to the one best representing the meaning of the stimulus word presented orally by the assessor. The *Test de Vocabulario en Imágenes Peabody* (TVIP) was used with the Spanish-speaking children (published reliability = 0.93). An adaptive shorter version of the PPVT and the TVIP was used for the Head Start Impact Study. The adaptive version was first used in the 1997 FACES project. The Technical Report for the Head Start Impact Study provides further details on the adaptation.
- **Color Names.** This task was adapted for use in FACES from the Color Concepts task included in *The CAP Early Childhood Diagnostic Instrument* (Mason & Stewart, 1989). The task measures color recognition (Color Names) by asking the child to point to different color bears and identify each of 10 colors by name. This test was translated into Spanish for use in FACES and also used in the Head Start Impact Study. FACES reported the reliability for Color Names as 0.95 (fall 2000) and 0.94 (spring 2001).

Language and Literacy: Oral Comprehension

- **Woodcock-Johnson III Tests of Achievement: Oral Comprehension.** This test measures the ability to comprehend a short spoken passage and to provide the missing word based on syntactic and semantic clues. The test requires the child to use listening, reasoning, and vocabulary skills. The assessor reads an analogy or passage with one word missing; the child is asked to respond orally with the correct word that completes the passage or analogy. The published median reliability is 0.80 in the 5-19 age range. No Spanish test was used for this measure.

[39] Two measures for the 3-year-old cohort have reliability (Cronbach alpha) less than 0.60. These are the Parent Emergent Literacy Scale at the end of the Head Start year and the Woodcock-Johnson Writing Samples at the end of 1st grade. Reliability was greater than 0.60 for the 4-year-old cohort. Reliabilities are provided for all measures in the Technical Report for the Head Start Impact Study.

Exhibit 2.7: Cognitive Domain Measures From Direct Assessment and Teacher Report and Year That Measure Was Administered

Construct & Test	What Is Measured	Year Measured		
		Preschool	K	1st Grade
COGNITIVE DOMAIN				
Vocabulary				
Peabody Picture Vocabulary Test III (PPVT adapted)	Vocabulary knowledge and receptive language	X	X	X
Color Names	Color identification	X		
Test de Vocabulario en Imágenes Peabody (TVIP adapted)	Vocabulary knowledge and receptive language	X	X	X
Oral Comprehension				
Woodcock-Johnson III Oral Comprehension	Oral comprehension using syntactic and semantic clues	X	X	X
Phonetic Awareness				
Preschool Comprehensive Test of Phonological and Print Processing: Elision (CTOPPP)	Phonetics of words, syllables, and phonemes	X	X	
Woodcock-Johnson III Word Attack	Phonetic and structural skills		X	X
Pre-Writing				
McCarthy Draw-A-Design	Perceptual motor skills	X		
Pre-Reading/Reading				
Letter Naming	Ability to recognize letters of the alphabet	X	X	
Woodcock-Johnson III Letter-Word Identification	Letter and word identification skills	X	X	X
Bateria R Woodcock-Muñoz Identificación de letras y palabras	Letter and word identification skills	X	X	X
Woodcock-Johnson III Passage Comprehension	Word recognition and reading comprehension using syntactic and semantic clues			X
Woodcock-Johnson III Spelling	Early writing and spelling	X	X	X
Woodcock-Johnson III Writing Samples	Writing			X
Batería R Woodcock-Muñoz Dictado	Early writing and spelling	X	X	X

Exhibit 2.7: Cognitive Domain Measures From Direct Assessment and Teacher Report and Year That Measure Was Administered (continued)

Construct & Test	What Is Measured	Year Measured		
		Preschool	K	1st Grade
COGNITIVE DOMAIN (continued)				
Math				
Woodcock-Johnson III Applied Problems	Analyze and solve math problems	X	X	X
Woodcock-Johnson III Quantitative Concepts – Concepts and Number Series	Knowledge of math concepts, symbols and vocabulary, counting, identifying numbers and shapes, and identifying number patterns		X	X
Woodcock-Johnson III Calculation	Mathematical computations			X
Counting Bears	One-to-one correspondence	X		
Batería R Woodcock-Muñoz Problemas Aplicados	Analyze and solve math problems	X	X	X
School Performance				
Grade Promotion (Parent)	Grade promotion		X	X
Academic Ratings (Teacher)	Rating of academic skills, school accomplishments		X	X

Language and Literacy: Phonological Awareness

- **Preschool Comprehensive Test of Phonological and Print Processing (CTOPPP): Elision.** The CTOPPP Elision measures the ability to remove words, syllables, and sub-syllables as part of words or compound words. Both multiple choice and free-response items are included in the test to create a new word. The child is asked to respond by pointing to pictures and verbally to the assessor's oral directions (e.g., Say seesaw without see). No published reliability is available. The instrument was translated for the Spanish version.
- **Woodcock-Johnson III Tests of Achievement: Word Attack.** This test measures the child's ability to apply phonic and structural analysis skills to the pronunciation of printed nonsense words. The initial items require the child to produce the sounds for a single letter. The remaining items require the child to read aloud nonsense words that become increasingly more difficult. The published median reliability is 0.87 in the 5-19 age range. No Spanish test was used for this measure.

Language and Literacy: Pre-Reading and Reading

- **Letter Naming.** This task was modified for use in FACES from a test used in the Head Start Quality Research Center's (QRC) curricular intervention studies. The Letter Naming task measures the child's ability to recognize the letters of the alphabet. The letters of the alphabet are divided into three plates with the easiest

letters printed on the first plate. Children are asked to identify each letter on the plate. No published reliability is available. This task was translated into Spanish for use in the Head Start Impact Study. Although this task was administered in English to the bilingual children, responses in English or Spanish were acceptable.

- **Woodcock-Johnson III Tests of Achievement: Letter-Word Identification.** The Letter-Word Identification test measures letter and word identification skills. The initial items involve symbolic learning or the ability to match a rebus (pictographic representation of a word) with an actual picture of the object. The remaining items measure a child's reading identification skills in identifying isolated letters and words as they appear in the test easel. The published median reliability is 0.91 in the 5 to 19 age range. The Batería-R Woodcock-Muñoz Pruebas de aprovechamiento-Revisada Identificación de letras y palabras was used for the Spanish and bilingual test administration.
- **Woodcock-Johnson III Tests of Achievement: Passage Comprehension.** This test measures the child's ability to match a pictographic representation of a word (rebus) with the actual picture of the object and to read a short passage and identify a missing key word based on the passage context. The items become more difficult by removing pictures and increasing passage length, level of vocabulary, and the complexity of semantic and syntactic clues. The published median reliability is 0.83 in the 5-19 age range. No Spanish test was used for this measure.
- **Woodcock-Johnson III Tests of Achievement: Spelling.** The Spelling test measures the child's ability to correctly write orally presented letters and words. For the initial items, pre-writing skills are measured through tasks such as drawing lines and copying letters. As the items progress in difficulty, the child is asked to write specific upper and lower cases of the alphabet and specific words. The published median reliability is 0.90 in the 5-19 age range. The Batería-R Woodcock-Muñoz Pruebas de aprovechamiento-Revisada Dictado was used for the Spanish test administration.
- **Woodcock-Johnson III Tests of Achievement: Writing Samples.** This test measures the child's ability to respond in writing to a variety of demands such as completing written passages or writing responses to pictures. The child is asked to respond to simple tasks such as completing the sentence, "My name is _____" to more complex tasks such as writing a sentence to describe a picture (e.g., picture of a bird in a cage singing). The published median reliability is 0.84 in the 5-19 age range. No Spanish test was used for this measure.

Pre-Writing

- **McCarthy Scales of Children's Abilities: Draw-A-Design Task.** The Draw-A-Design task is a measure of perceptual motor skills and pre-writing. The child is asked to draw a series of increasingly complex figures. The reliability for the Perceptual-Performance subscale, of which the Draw-A-Design is one component, is 0.84. The task was translated into Spanish for use in FACES for the 1997 cohort and also used in the Head Start Impact Study.

Math

- **Counting Bears.** This task was adapted for use in FACES from the counting tasks included in *The CAP Early Childhood Diagnostic Instrument* (Mason & Stewart, 1989). The task measures early numeracy skills of counting and one-to-one correspondence (counting). The child is asked to count 10 pictures of bears and arrive at the correct sum. This test was translated into Spanish for use in FACES and also used in the Head Start Impact Study. FACES reported the reliability for the combined tests of Color Names and Counting Bears as 0.95 (fall 2000) and 0.94 (spring 2001). No separate published reliability is available for Counting Bears.
- **Woodcock-Johnson III Tests of Achievement: Applied Problems.** This test measures the child's ability to analyze and solve practical math problems. To solve the problems that are read by the assessor to the child, the child must recognize the procedure to be followed and then count and/or perform simple calculations. The published median reliability is 0.92 in the 5-19 age range. The Batería-R Woodcock-Muñoz Pruebas de aprovechamiento-Revisada Problemas aplicados was used for the Spanish test administration.
- **Woodcock-Johnson III Tests of Achievement: Quantitative Concepts.** This test consists of two subtests: Concepts and Number Series. Concepts measures the child's understanding of counting; identifying numbers, shapes and sequences; and knowledge of mathematical terms and formulas. Number Series measures the child's ability to look at a series of numbers, determine the pattern, and provide the missing number in the series. The published median reliability is 0.90 in the 5-19 age range. No Spanish test was used for this measure.
- **Woodcock-Johnson III Tests of Achievement: Calculation.** This test measures the ability to perform mathematical computations. The initial items require the child to write single numbers. The items progress in difficulty from basic operations to geometric, trigonometric, logarithmic, and calculus operations. The calculations involve whole numbers, percents, fractions, decimals, and negative numbers. The published median reliability is 0.85 in the 5-19 age range. No Spanish test was used for this measure.

School Performance Measures

The tests included in the direct child assessment battery are described above. Other measures of children's cognitive skills include the following:

- **Teacher report of academic skills.** Each child was rated on three academic skills (language and literacy, science and social studies, and mathematical skills) by his/her teacher. The child was rated as compared to other children at the same grade level using a five point scale ranging from one (far below average) to five (far above average). For the analysis, the scores were collapsed to zero (far below average and below average) and one (average, above average, and far above average).

- **Teacher report of school accomplishments.** Each child was rated by his/her teacher on a series of items that described the child's skills, knowledge, and behaviors focusing on language and literacy and mathematics. The child was rated using a five-point scale that reflected the degree to which the child acquired the demonstrated skills, knowledge, and behaviors, ranging from one (not yet) to five (proficient). More complex skills, knowledge, and behaviors were added to the 1st grade list.
- **Parent report of promotion.** Parents were asked the grade level of their child. This information was confirmed with the teacher-reported expected promotion of the child. Overall there was consistency between the two reports. Parent data were used because the response rate was higher for parents than teachers.
- **Parent emergent literacy scale (PELS).** PELS is a parent-report on five literacy items originally developed for use in FACES 2000: child can recognize most/all of the letters of the alphabet; child can count to 20; child pretended to write his/her name in the last month; child can write his/her first name; and child can identify the primary colors.

In addition to the individual tests, five Woodcock-Johnson III composite measures, derived from the results of the individual tests described above, were also used as outcome measures to provide a more multifaceted assessment of children's ability:

- **Pre-Academic Skills.** This cluster measures pre-reading skills, letter and word identification skills, developing mathematics skills, and skill in written production. The tests included in the cluster are Letter-Word Identification, Spelling, and Applied Problems. The published median reliability is 0.97 for four- and five-year-olds and 0.98 for six-year-olds. The Pre-Academic Skills Cluster was available for the Spanish Administration.
- **Basic Reading Skills.** This cluster measures sight vocabulary, phonics, and structural analysis. Tests included in the cluster are Letter-Word Identification and Word Attack. The published median reliability is 0.93 in the 5-19 age range. This composite was not available for the Spanish assessment.
- **Math Reasoning.** This cluster measures mathematical problem solving, analysis, reasoning, and vocabulary. Tests included in the cluster are Applied Problems and Quantitative Concepts. The published median reliability is 0.95 in the 5-19 age range. This composite was not available for the Spanish assessment.
- **Academic Skills.** This cluster is an aggregate measure of reading decoding, math calculation, and spelling of single-word responses. Tests included in the cluster are Letter-Word Identification, Calculation, and Spelling. The published median reliability is 0.95 in the 5-19 age range. This composite was not available for the Spanish assessment.
- **Academic Applications.** This cluster measures the application of academic skills to academic problems. Tests included in the cluster are Passage Comprehension, Applied Problems, and Writing Samples. The published median reliability is 0.94 in the 5-19 age range. This composite was not available for the Spanish assessment.

Finally, four tests were administered to the study children but were not used in the final analysis because they were subsequently found to have problematic psychometric properties. These tests are:

- **Leiter Revised: Sustained Attention Task.** This task measures the child's ability to pay sustained attention to a repetitive task and to pay attention to detail. This is a timed test with a targeted picture at the top of each page. The child is asked to cross as many of the target pictures as possible during the allotted time. The targeted pictures are interspersed among non-target pictures. The Attention Sustained task is one of 10 tasks in the Attention and Memory battery. An adaptive shorter version of the task was used for the Head Start Impact Study. The overall published reliability is 0.83. The task is a nonverbal task, but the directions were translated into Spanish for use in FACES and also used in the Head Start Impact Study. Due to low reliability, data from this test are not included in analysis.
- **Story and Print Concepts.** This test was adapted for use in FACES from the Story and Print Concepts task included in *The CAP Early Childhood Diagnostic Instrument (*Mason & Stewart, 1989). This test measures emerging literacy relative to knowledge of books and print concepts. For this test, the assessor reads a book to the child asking questions as the book is read. FACES reported reliabilities for the constructs measured in the Story and Print Concepts subtests in fall 2002 and spring 2001 were: Book Knowledge (0.57 and 0.59); Print Conventions (0.73 and 0.74); and Comprehension (0.43 and 0.41). This test was translated into Spanish for use in FACES and also used in the Head Start Impact Study. The books used to assess the child's story and print concepts were as follows: (1) Alborough, J. (1992). ***Where's My Teddy?*** Cambridge, MA: Candlewick Press. (English version) and Alborough, J. (1995). ***¿Dónde Está Mi Osito?*** (translated by M. Castro) Miami, FL: Santillana USA Publishing Company, Inc. (Spanish version). Due to low reliability, data from this test are not included in analysis.
- **Writing Name Task.** This task was modeled after the Name Writing tasks in *The CAP Early Childhood Diagnostic Instrument (*Mason and Stewart, 1989) and the Writing Samples test in the *Woodcock-Johnson III Tests of Achievement* (2001). The task measures the child's basic writing skills. For this task, the child is asked to write his or her name. No published reliability data are available. This task was translated into Spanish for use in the Head Start Impact Study. Ninety-eight percent of the children could write their name, so there was no meaningful variation, and data were not included in the analysis.
- **Preschool Comprehensive Test of Phonological and Print Processing (CTOPPP): Print Awareness.** The CTOPPP Print Awareness, adapted for this study, measures the recognition of letter symbols and sounds. The child is asked by the assessor to point to a letter (letter discrimination) that represents the stimulus sound provided orally by the assessor. Additional items measure print concepts, word discrimination, letter-name identification, letter-name identification free response, and letter-sound identification free response. No published reliability data are available. The instrument was translated into Spanish for use with bilingual children.

Social-Emotional Domain

Measures for this domain include parents' report of child behavior and parent and child relationships as well as teacher report of classroom behavior:[40]

- **Social skills and positive approaches to learning.** Parents were asked to rate their child's social skills and positive approaches to learning. The measure assesses social skills focused on cooperative and empathic behavior and children's approaches to learning such as curiosity, imagination, openness to new tasks and challenges, and having a positive attitude about gaining new knowledge and skills. Examples of the questions asked include: "Makes friends easily," "Comforts or helps others," "Accepts friends' ideas in sharing and playing," "Enjoys learning," "Likes to try new things," and "Shows imagination in work and play." The scale contains seven items, with each item scored from zero (not true) to two (very true), and the scale scores can range from zero to 14. The scale is based on an instrument used in FACES and is based on a modified Achenbach Classroom Behavior Checklist (CBCL) (Achenbach, Edelbrock, and Howell, 1987).
- **Social competencies checklist.** Parents were asked to provide information on social competencies using a 12-item Social Competencies Checklist, also used in FACES 2000. For each item, the parent was asked to report whether the child engaged in that behavior or exhibited that attribute "regularly" or "very rarely or not at all." Examples of the items included: "Shares newly learned ideas," "Takes care of personal belongings," "Helps with simple household tasks," and "Notices when others are happy, sad, angry." The total scale score could range from zero (all items rated "rarely or not at all") to 12 (all items rated "does regularly") (Developing Skills Checklist, 1990).
- **Problem behavior of children.** Parents were asked to rate their children on items dealing with aggressive or defiant behavior such as, "Hits and fights with others," "Has temper tantrums or hot temper," and "Is disobedient at home." Other items dealt with inattentive or hyperactive behavior, including, "Can't concentrate, can't pay attention for long," and "Is very restless and fidgets a lot." A third set of items dealt with shy, withdrawn, or depressed behavior, e.g., "Feels worthless or inferior," and "Is unhappy, sad, or depressed." For each item, the parent was asked to judge whether the behavioral description was "not true," "sometimes true," or "very true" of the child. We constructed four measures from the instrument described below:
  - The **Total Behavior Problem** scale derived from parent ratings contained 14 rating items, and the total scale score could range from zero (all items marked "not true") to 28 (all items marked "very true").
  - The **Aggressive Behavior** subscale contained four items, and could range from zero to eight.

[40] Social emotional measures with a reliability below 0.60 are indicated in the main impact tables in Chapter 5. All reliabilities are provided in the Technical Report for the Head Start Impact Study.

- The **Hyperactive Behavior** subscale contained three items, and scores could range from zero to six.
- The **Withdrawn Behavior** subscale contained three items, and scores could range from zero to six.

These scales were also used in FACES 2000. The mean scores obtained in the Head Start Impact Study were very comparable to mean scores obtained from parents of an independent national sample of Head Start children in FACES (U.S. Department of Health and Human Services, 2003).

- **Child-Parent Relationship.** Parents were asked to rate their child's relationship with them using a five-point response format ("definitely does not apply," "not really," "neutral or not sure," "applies sometimes," and "definitely applies"). The scale includes items such as, "My child and I always seem to be struggling with each other," "If upset, my child will seek comfort from me," and "My child values his/her relationship with me." The 15-item instrument generates three dimensions: closeness, conflict, and positive relationships.
  - The **Closeness** dimension, focusing on positive effect, consists of seven items, such as, "It is easy to be in tune with what my child is feeling," and "When I praise my child, he/she beams with pride." Scores can range from seven to 35.
  - The **Conflict** dimension is measured by eight items that indicate the level to which the parent and child are at odds with each other, such as, "My child easily becomes angry with me," and "When my child is in a bad mood, I know we're in for a long and difficult day." Scores can range from eight to 40.
  - The **Positive Relationship** dimension is a measure of the overall relationship between the child and the parent and combines the closeness and conflict items (reversed). The score for the 15 items can range from 15 to 75. This scale, modified from the Student-Teacher Relationship Scale by changing "this" child to "my" child in the items (e.g., If upset, this (my) child will seek comfort from me). The modified instrument measures the parent perspective on the closeness, conflict and positive relationship dimensions (Pianta, 1992).
- **Student-Teacher Relationship.** This instrument developed by Robert Pianta (1996) includes three scales similar to the child-parent relationship scales discussed above, i.e., closeness, conflict, and total positive relationship. Both a short form and a long form are available. The shortened version of the instrument was used for the Head Start Impact Study. The teacher is asked to rate the child on 15 items, such as, "If upset, this child will seek comfort from me," or "This child easily becomes angry at me." The teacher rates the child on each item using a five-point response format ranging from one (definitely does not apply) to five (definitely applies). The closeness scale contained seven items, and the scores could range from seven to 35. The conflict scale contained eight items, and the scores could range from eight to 40. The total positive relationship scale contained 15 items, and the scores could range from 15 to 75.

- **Adjustment Scales for Preschool Intervention (ASPI).** The ASPI is based on the Adjustment Scales for Children and Adolescents (ASCA). The ASPI (Lutz, Fantuzzo, & McDermott, 2000; 2002) is a teacher-reported instrument designed to measure emotional and behavioral adjustment within usual classroom situations. It consists of 24 classroom situations that provide 144 descriptors of both typical and problem classroom behavior. The teacher is asked to select all behavior descriptions that match a child's behavior to a specified classroom situation over the past two months. The raw score is the sum of the behavior descriptions checked for each dimension. Raw scores are converted to *t*-scores for the analysis. A *t*-score of 60 or higher on the ASPI is empirical documentation of a problem (e.g., aggressive behavior, withdrawn behavior, etc.).[41] As expected, the number of children with a *t*-score of 60 or higher is usually small. Five behavioral dimensions can be measured using the ASPI.
  - The **Aggressive** dimension consists of 22 descriptions of rough or aggressive behaviors, such as "Overly rough with other children in games" and "Answers back aggressively, makes threats or creates a disturbance when corrected."
  - The **Withdrawn-Low Energy** dimension consists of 18 descriptions of behavior related to disconnection from activity, primarily due to low energy levels, such as, "Too lacking in energy to be troublesome" and "Sits lifelessly most of the time during teacher-directed activities."
  - The **Socially Reticent** dimension characterized by shy or hesitant behaviors is composed of 12 items, such as, "Needs encouragement to join in games" and "Tends to have untalkative moods."
  - The **Oppositional** dimension includes 11 items related to moody or controlling behavior, such as, "Answers questions except when in bad mood" and "Tells on others to gain teacher's favor."
  - The **Inattentive/Hyperactive** dimension consists of 10 items that describe inattention, impulsivity, or high levels of activity, such as, "Answers questions before taking time to think" and "Constantly restless (changes position, etc.)."

  In addition to the behavioral dimensions, three situational dimensions related to problem behavior were also measured using the ASPI:
  - The **Problems with Structured Learning** dimension consists of seven structured classroom situations and 40 behavioral responses associated with *problem behavior,* such as class involvement ("Seldom gets involved in any class activities," "Listless, seems unmotivated," "At times does not participate in activities," or "Only gets involved with the help of adults") or how the child sits during whole-group, teacher-directed activities ("Sits lifelessly most of the time," "Sits meekly, seems afraid to budge," "Doesn't stay seated when he/she should," "Constantly restless," "Experiments with unusual sitting positions, climbs on

[41] All ASPI raw scores were converted to *t*-scores derived from the developer's original ASPI standardization sample. The ASPI standardization sample consisted of Head Start children in a pre-kindergarten Head Start program in a large urban school district.

classroom furniture, etc.," or "Bothers other children"). Other classroom situations measured in this dimension include taking part in games with other children, maintaining friends, paying attention in the classroom, engaging in free play/individual choice, and working with his/her hands.

- The **Problems with Peer Interaction** dimension consists of six peer situations and 24 behavioral responses associated with *problem behavior* in the classroom, such as telling the truth ("Tells tall tales about him/her or family," "Occasionally lies to avoid blame or punishment," or "Doesn't hesitate to lie") or how the child behaves in the classroom ("Misbehaves when you are attending to others," "Never any trouble because he/she is so timid," "Too lacking in energy to be troublesome," or "Does things in front of you he/she knows are wrong"). Other problem behaviors measured in the Peer Interaction dimension include reaction to correction, respect for other's belongings, behavior when standing in line, and interaction with others of his/her age.
- The **Problems with Teacher Interaction** dimension includes six classroom situations directly involving teachers and 30 behavioral responses where *behavior problems* may occur, such as helping the teacher with jobs ("Asks to be given jobs but often doesn't finish them or do them according to classroom standards," "Helps unless in a bad mood," "Appears too withdrawn to come forward," or "Causes a disturbance when not chosen for jobs") or how the child seeks the teacher's help ("Too lethargic to ask," "Seeks help when not needed," "Not shy but never seeks help," or "Too timid to ask"). Other classroom situations measured with this dimension include greeting the teacher, answering teacher questions, talking to the teacher, and general manner with the teacher.

Health Domain

Health measures were based on parent report and include the following:

- **<u>Receipt of health care services</u>.** Parents were asked to report on whether the child had received or had access to, two health care services.
  - **Whether the child has health insurance.** Parents were asked if the child was currently covered by Medicaid or a state health insurance program or by health insurance through their job or the job of another employed adult.
  - **Whether the child has received dental care.** Parents were asked if the child had seen a dentist since September.
- **<u>Child's health status</u>.** Parents were asked to report on their child's current health status:
  - **Child's health status (excellent or very good).** Parents were asked if, overall, the child's health was excellent, very good, good, fair, or poor. A dichotomous variable was developed—those who reported that their child's health was excellent or very good and those who reported that their child's health was good, fair, or poor.

- **Whether the child needs ongoing medical care.** Parents were asked if their child had an illness or condition that requires regular ongoing medical care.
- **Whether child received medical care for an injury in the last month.** Parents were asked how many times their child, in the last month, had seen a doctor or other medical professional or visited a clinic or emergency room for an injury. This outcome was coded yes if the parent reported any such occurrences in the last month.

Parenting Practices Domain

Parenting practices measures were based on both parent and teacher report and include the following measures:[42]

- **Educational activities.** Parents were asked to report on the types of educational activities they did with their child:
  - **Reading to the child at home.** Parents reported on the item, "How many times have you or someone in your family read to [CHILD] in the past week?" Possible responses range from one (not at all) to four (every day).
  - **Family cultural enrichment activities.** Parents reported on a seven-item checklist of activities the parent, or another family member, may have done with the child during the past month. The seven activities include: (1) going to a movie; (2) play or concert; (3) art gallery or museum; (4) playground, park, or zoo; (5) participating in community, ethnic, or religious event; (6) talking about family or cultural heritage; and (7) going on errands. A total score was computed by summing the number of different activities the parent and child participated in together, with a possible score of zero (none) to seven (all).
- **Discipline practices.** Parents reported on the following:
  - **Use of physical discipline.** Parents reported on the item, "Sometimes children mind pretty well and sometimes they don't. Have you spanked [CHILD] in the past week for not minding?"
  - **Use of time out.** Parents reported on the item, "Have you used 'time out' or sent [CHILD] to his/her room in the past week for not minding?"
- **Parental safety practices.** Parents reported on a 10-item scale that assessed how often each of 10 different safety precautions were used, including keeping harmful objects out of reach, using car seats, supervising the child during bath time, and having a first aid kit and working smoke detector at home. Possible responses ranged from one (never) to four (always).
- **Parenting styles.** The parents were asked to respond to selected items from the Child-Rearing Practices Report (CRPR) (Block, 1965). Parents were asked to

[42] Parenting practice measures with a reliability below 0.60 (i.e., parental safety practices scale, family cultural enrichment scale, and parent participation in school) are indicated in the main impact tables in Chapter 7. All reliabilities are provided in the Technical Report for the Head Start Impact Study.

respond to items, such as, "I do not allow my child to get angry with me" and "I am easygoing and relaxed with my child" using a Likert scale that ranged from one (exactly like you) to five (not at all like you). The parenting styles identified for the analysis, and described in Chapter 7 are: authoritative, authoritarian, permissive, and neglectful.

- **Teacher report of parent participation in school activities.** Teachers were asked two questions: "Have one or both of the child's parents (or guardians) attended open house meetings, back-to-school nights, or class events, such as a class play or recital, this year?" and "Have one or both of this child's parents (or guardians) acted as volunteers or helped out with class activities or class trips this year?"
- **Teacher report of school contact/communication.** Teachers were asked two questions: "How often have this child's parents (or guardians) initiated contact with you to find out how things were going with the child or to offer help with class activities?" and "How often have you had to contact or tried to contact this child's parent(s) or guardians about behavior or schoolwork problems this child has been having?" The response categories ranged from zero (not at all) to four (about once a month or more often).

A summary of the social-emotional, health, and parenting measures is presented in Exhibit 2.8.

Exhibit 2.8: Social-Emotional, Health, and Parenting Practice Domain Measures from Parent and Teacher Report and Year That Measure Was Administered

Construct & Test	What Is Measured	Year Measured		
		Preschool	K	1st Grade
SOCIAL-EMOTIONAL DOMAIN				
Problem Behaviors				
Adapted Child Behavior Checklist (parent report)	Total problem behavior, hyperactive behavior, aggressive behavior, and withdrawn behavior	X	X	X
Social Skills and Social Competencies				
Developing Skills Checklist	Parents rate children on a number of social skills	X	X	X
Social Skills and Positive Approaches to Learning	Parents rate children on social skills and learning behaviors	X	X	X
Approaches to Learning				
Adjustment Scales for Pre-School Intervention (ASPI)	Emotional and behavioral adjustment in the classroom	X	X	X
Parent-child relationship and student-teacher relationship				
Teacher/Parent Child Relationship Scale	Child temperament and behavior	X	X	X

Exhibit 2.8: Social-Emotional, Health, and Parenting Practice Domain Measures from Parent and Teacher Report and Year That Measure Was Administered (continued)

		Year Measured		
Construct & Test	**What Is Measured**	**Preschool**	**K**	**1st Grade**
HEALTH DOMAIN				
Receipt of Health Care Services	Parent report of child's use of and access to dental care and health insurance	X	X	X
Health Status	Parent report of child's overall health status, child needs ongoing care, child has had care for injury	X	X	X
PARENTING PRACTICES DOMAIN				
Educational Activities with Child	Parent report on frequency of reading to child at home, cultural enrichment activities done with child, summer activities	X	X	X
Discipline Practices	Parent report of use of physical discipline and use of timeout	X	X	X
Safety Practices	Parent report on 10 item scale. How often used 10 different safety precautions from seat belts to supervising during bath time	X	X	
School Contact and Communications and Parent Participation	Parent report on receipt of information on child progress, parents' comfort at school, parent volunteering, and teacher report of parent involvement		X	X
Parenting Styles	Degree to which warmth and control exist in parent-child relationship	X	X	X

Preschool and Elementary School Program Measures

The preschool and elementary school measures are used in this report to describe children's experiences as they were followed over the course of this study from application to Head Start through the end of 1st grade.

Preschool Measures

The preschool measures were selected both because of their relevance to the Head Start Performance Standards and because of an important role they may play in children's subsequent development and overall preschool experiences. Measures are presented across six constructs: type of preschool setting attended by children, center environment, teacher training and qualifications, classroom environment, classroom activities, and overall quality of children's experience. The same measures were used for Head Start and other center-based programs. When applicable, similar measures were used for family day care homes. Each measure is described below.

- **Type of child care setting.** Information was obtained from parent interviews each spring to identify a focal child care setting for each study child. The focal setting is defined as the child care setting where the child spent a minimum of five hours between Monday and Friday and the hours of 8 am and 6 pm. Settings include: center-based program (including Head Start), non-relative's home, relative's home, non-parental care in the child's own home by a non-relative, non-parental care in the child's own home by a relative, and parent care. Head Start is always defined as the focal setting for children enrolled in the program, whether the child was in the Head Start or control group.

  In addition, parents were asked several questions about their child's setting:

  - **Stability in child care relationships.** Based on a question in the parent interview that asked in spring 2003 whether there was stability in his/her child care relationships—never, sometimes, often, or always.
  - **Too much turnover in care providers.** Based on a question in the parent interview that asked in spring 2003 whether there was too much turnover in care providers—never, sometimes, often, or always.
  - **Child was in a familiar place.** Based on a question in the parent interview that asked in spring 2003 whether their child was in a familiar child care setting with people (he/she) knows—never, sometimes, often, or always.

- **Center environment.** Information on the operation and overall environment of Head Start and non-Head Start center-based preschool programs was collected from interviews with center directors. Where applicable (and noted below), similar information was collected from the day care home provider. Measures were selected because of their potential effect on children's learning experiences and include:

  - **Center size.** The median capacity size of centers attended by study children was 50. To differentiate whether children attended large or small centers, a dichotomous variable was created indicating whether the center served more than 50, or 50 or fewer children. Size was also collected from day care home providers.

- **Center affiliation.** Center affiliation is defined as whether the center is affiliated with a school or not affiliated with a school. This dichotomous split is highlighted because of the recent attention being paid to the effectiveness of school-based pre-K programs.
- **Part-day/full-day program.** At the time of random assignment (fall 2002) centers were asked if they offered part-day Head Start, full-day Head Start, or a combination of both programs. The measure is based on the Head Start definition of part-day/full-day. This measure comprises three groups—part-day only, full-day only, or both and represents the percentage of children in centers where the options are offered.
- **Director qualifications.** The center director provides overall program leadership and management. His/her qualifications are measured by education level (whether the director had at least a bachelor's degree) and tenure (the length of time the director has been in his/her current position—greater than four years vs. four years or less). Four years was selected as an indication of substantial stability in the position.
- **Teacher training and mentoring.** To capture the availability of teacher training, the center director was asked how often center-wide teacher training was provided (weekly, monthly, once every few months, about once a year, every few years). The director also was asked about the availability of mentor teachers to work with teachers. The two measures are presented as: (1) whether training is provided at least monthly and (2) whether center provides any teacher mentoring.[43]
- **Teacher turnover.** Teacher stability and consistency are measured by whether the center has a low proportion of new lead teachers each year. Low proportion of new lead teachers is defined as less than or equal to 20 percent based on an average turnover rate at centers.
- **Curriculum use.** Head Start Performance Standards mandate the use of a curriculum, consistent with developmentally appropriate early childhood care and education. Directors and day care home providers were asked if a curriculum was used center wide, and if so, the name of the curriculum.
- **Services available for children and families.**[44] In keeping with the overall goal of Head Start to serve the "whole child," Head Start Performance Standards require that grantee and delegate agencies work collaboratively with all participating parents to identify and then either refer or provide services and resources that are responsive to each family's goals and interests. Additionally,

[43] These measures describe the overall availability of training at the center. Training of the individual teachers of study children is covered in a separate variable under Teacher/Care Provider Qualifications and Training.

[44] These data come from Center Director and Care Provider interviews, and although the data are combined here, the actual wording of the questions differed somewhat across the two data collection activities. **Services to Children:** The center director interview asked: "What services does the center provide to children?" The care provider interview asked: "Do you provide or arrange any of the following services for children in your care?" **Services to Families:** The center director interview asked: "What services does the center provide to other family members?" The care provider interview asked: "Since September, have you helped the families of children you care for obtain any of the following services?"

the performance standards outline health, mental health, and nutrition services to be available to children and families. To reflect these standards, directors were asked whether a number of services were available to children and their families. Children's services include: (1) mental health services, (2) health services, (3) hearing/vision screening referrals, and (4) nutrition services. Family services include: (1) home adult education/literacy, (2) family counseling, (3) job training, (4) help with medical care, (5) dealing with family violence, (6) housing, (7) utilities, (8) alcohol/drug abuse treatment or counseling, (9) food and nutrition assistance, (10) income assistance, (11) foster care payments, and (12) home visitation. Two measures were created: (1) the percentage of children in centers that provide each individual service; and (2) the number of services divided into three categories—0-8, 9-14, and 15 services.

- **Competition from other preschools.** To provide further understanding of other child care resources available to parents, directors were asked about the extent of competition from other preschools in the community and whether the center was always filled to capacity. Two measures were created: (1) a dichotomous variable differentiating always filled or never/rarely filled and (2) a dichotomous variable indicating competition/no competition.
- **Respect for family culture.** Based on a question to parent which asked him/her to rate whether they were very dissatisfied, somewhat dissatisfied, somewhat satisfied, or very satisfied with how the Head Start center supported and respected his/her family's culture and background.
- **Helping child grow and develop.** Based on a question in the parent interview that asked them to rate whether they were very dissatisfied, somewhat dissatisfied, somewhat satisfied, or very satisfied with how the center helped the study child to grow and develop.
- **Open to ideas and participation.** Based on a question in the parent interview that asked them to rate whether they were very dissatisfied, somewhat dissatisfied, somewhat satisfied, or very satisfied with how much the center was open to their ideas and participation.
- **Child felt safe/secure in center.** Based on a question in the parent interview that asked about their child's experience at the center and to identify whether the study child felt safe and secure at the center—never, sometimes, often, or always.
- **Child received individual attention.** Based on a question in the parent interview that asked about the child's experience at the center and to identify whether the study child received individual attention—never, sometimes, often, or always.
- **Teacher was open to new information and learning.** Based on a question in the parent interview regarding whether the teacher was open to new information and learning—never, sometimes, often, or always.
- **Parental involvement.** Parents were asked about how often they volunteered or observed in the classroom; attended parent-teacher conferences; attended parent education meetings or workshops; attended or helped with activities such as fieldtrips; or participated in fundraising, Policy Council, any other activities. A

continuous variable was created (zero to five) to show how many of these activities the parent did.

- **Child in urban or rural Head Start center.** To derive this variable, families were first linked to the Head Start center where they were randomly assigned. Then each Head Start center's address was geocoded and matched to criteria established by the U.S. Census to determine whether it was located in a Census-defined urbanized area. If so, the children were classified as urban; if not, they were classified as rural.

- **Teacher/care provider qualifications and training.** The current Head Start reauthorization act (P.L. 110-134) requires that by 2013, at least 50 percent of Head Start teachers nationwide in center-based programs have at least (1) a bachelor's degree in early childhood education (ECE), or (2) a bachelor's degree and coursework equivalent to an ECE major plus experience in teaching preschool children. It also requires that all Head Start assistant teachers nationwide in center-based programs have a Child Development Associate's degree (CDA) or be working toward an associate's or bachelor's degree by 2013. These requirements were designed to improve the quality of Head Start programs, reflecting the importance of teachers' educational attainment, training, and specialized credentials in children's preschool experiences. The teacher qualification variables represent study children's teachers rather than all the teachers in a center. Also included in these measures are children's care providers (non-center-based care providers) if the children are in child care homes. The measures used are as follows:

  - **Teacher education.** Teacher education is measured at various levels of educational completion: (1) had college ECE courses or obtained CDA, (2) only obtained CDA, (3) highest educational attainment was associate's degree, (4) obtained a bachelor's degree or higher, and (5) attained bachelor's degree or higher in ECE. For some analyses, a dichotomous variable was used dividing the sample into two groups—children who were in classrooms that had teachers with a BA and above and those in classrooms with teachers whose education level was below a BA.

  - **Teacher training and mentoring.** Separate teacher training and mentoring measures are provided. There was no information on the quality of teachers' training. However, information on the frequency was included. Teacher training is defined as having received at least 25 hours of training in the past year vs. less than 25 hours. Teacher mentoring is defined as having received mentoring at least once a month vs. less than once a month.

- **Classroom and child care home environment.** The classroom environment construct is intended to capture overall quality of the preschool classrooms that children attended. Preschool programs are typically rated on two important dimensions of quality—process characteristics (e.g., nature of teacher-child interactions, use of curriculum, schedule of activities, and use of instructional materials) and structural indices such as staff-child ratio and group-size (Phillips et al, 2001). The classroom environment measures presented in this chapter include the following:

- **ECERS-R/FDCRS.** Classroom quality was measured, using the Early Childhood Environmental Rating Scale-Revised (ECERS-R) (Harms et al, 1998) for children who were in centers and the Family Day Care Rating Scale (FDCRS) (Harms and Clifford, 1989) for children who were in child care homes. Classroom observations conducted by study staff gave the classroom a rating ranging from a low of one ("inadequate") to a high of seven ("excellent") on 37 items covering six subscales: (1) adequacy of space and furnishings; (2) personal care routines; (3) language and reasoning, including materials available and activities used; (4) range of activities that are used and available; (5) interactions, including both staff-child and child-child interactions; and (6) program structure, including the use of a daily schedule. An overall average score and scores for each individual subscale were computed for all classrooms attended by study children and the percentage of children in classrooms with overall average scores of five, six, and seven are presented.[45] For some analyses, a dichotomous variable was created indicating poor quality (a score of three or less) or better than poor quality (score greater than three).

- **Arnett Caregiver Interpretation Scale.** In addition, the classroom/child care home observers completed the Arnett Caregiver Interaction Scale (Arnett, 1989), with ratings ranging from one ("not at all") to four ("very much") on traits and interactions: (1) greater teacher sensitivity, (2) responsiveness, (3) encouragement of children's independence, and (4) lower levels of punitiveness and detachment.[46] An overall average and subgroup scores were computed for all classrooms attended by study children, and the percentage of children in classrooms with an overall average score of three and four are presented. For some analyses, a dichotomous variable was created dividing centers on a cut point of an overall score of four versus a score below four.

- **Child/staff ratio.** The child/staff ratio is based on observers' recorded counts of children and staff present at various times during the classroom observation. The ratio variable indicates that the child's setting either did or did not meet the standard for the type and size of the observed setting. The challenge for this analysis was the need to create a standard that covered both classrooms in centers and child care homes and was applicable to children of different ages. Consequently, the ratio variable used is based on standards adopted by the American Academy of Pediatrics (AAP) and American Public Health

[45] Individual subscale scores are presented in Tables in Appendix D of this Volume.

[46] Items indicating punitiveness and detachment were reverse coded to provide a measure of more positive teacher/child relationships.

Association(APHA) (2002).[47] A dichotomous variable of meets/does not meet standards was developed.[48]

- **Parental involvement in the classroom.** This measure is based on the parent interview where parents were asked how often they (1) volunteered or observed in the child's classroom, (2) attended conferences, (3) attended parental meetings, (4) attended activities such as field trips, and (5) were involved in center planning groups. This question reflects the Head Start Program Performance Standard requirement that programs provide parents with opportunities to be involved in their child's classroom and offer a variety of parenting education services. A dichotomous variable was devised—parent participated in at least one activity or did not participate in any activity.

- **Classroom and child care home activities.** The number and frequency of a variety of activities were reported by teachers. The more activities and frequent use of activities were combined to indicate a higher level of availability and variety of learning opportunities provided in the classroom. The measures used are:

  - **Language and literacy activities.** Teachers and care providers were asked how often they used each of 12 reading and language activities with children in their classroom or child care home (e.g., work on learning the names of letters, practice writing or spelling their name, practice sounds, make-up stories). These items are an indicator of the availability and variety of learning opportunities provided in the classrooms and child care homes. Chapter 3 presents the percentage of children who were exposed to at least seven of the 12 activities at least three times a week to differentiate between classrooms providing more emphasis on language and literacy activities from those not placing as much emphasis on them. For other analyses, the percentage of children in classrooms where literacy activities were provided at least three times weekly is presented using three categories—0-5, 6-8, and 9-12 activities done at least three times weekly.

  - **Math activities.** Teachers and care providers were asked how often they used each of eight math activities with children in their classroom or child care home (e.g., count out loud, play with shape blocks, work with rulers or measuring cups). Chapter 3 presents the percentage of children who were exposed to at least five of the eight activities at least three times a week to differentiate between those

[47] The AAP/APHA standards were used because, first, the Head Start requirements do not cover child care home ratios while the AAP/APHA standards do, so it is more consistent to apply the AAP/APHA standards to all the settings. In addition, using the AAP/APHA standards follows the practice of the NICHD study of early child care, which used the AAP/APHA standards because they represent a higher than minimum standard of quality that can promote better child outcomes (National Institute of Child Health and Human Development, 2006). The Head Start Performance Standards specify a maximum of 7.5 to 8.5 children per staff member in classrooms of 3-year-olds and 10 children per staff member in classrooms of 4-year-olds.

[48] The standards used for the child/staff ratio variable are as follows: Classrooms for the 3-year-old cohort: ratio must be at least 7:1; Classrooms for the 4-year-old cohort: ratio must be at least 8:1; Small child care homes (six or fewer children) and: No child under age two: ratio must be at least 6:1; At least one child under age two: ratio must be at least 4:1; Large child care homes (more than six children) with at least one 3-year-old and: No child under age two: ratio must be at least 7:1; At least one child under age two: ratio must be at least 5:1; Large child care homes (more than six children) with no child under age four: ratio must be at least 8:1.

centers that place more emphasis on providing math activities and those that did not. For other analyses, the percentage of children in classrooms where math activities were provided at least three times weekly is presented using three categories—0-3, 4-5, and 6-8 activities at least three times weekly.

- **Other activities.** Teachers and care providers were asked how often they used four other instructional activities (arts and crafts, games, sports, and chores) with the children in their classroom or child care home. Chapter 3 presents the percentage of children who were exposed to at least three or four activities at least three times a week.
- **Overall classroom and child care home quality.** To describe preschool quality, several indicators were combined to create an overall quality score. The composite brings together information from the observation ratings, activities provided in the setting, teacher qualifications and experiences, parent involvement, home visits, and program services.[49] The quality score is a continuous variable that can range from zero to one. Two variables were created: (1) a dichotomous variable to differentiate from higher overall quality (.75 or above) and lower quality, those that scored in the upper quarter of the measure as compared to those who were not in the upper quarter, and (2) children who attended a classroom where the score was above the cohort mean and where it was below the cohort mean.

- **Early elementary school measures** are presented in three categories: school environment and characteristics, teacher and classroom characteristics, and classroom activities. Measures for both kindergarten and 1st grade experiences are provided. More detailed information about children's early school experiences will be provided in a subsequent report that will cover the period through the end of 3rd grade. By the end of 1st grade, the two cohorts of children (excluding those in Puerto Rico) had attended 2,275 schools across the country.
  - **School environment.** Elementary school observations were not conducted, and so several measures were obtained from secondary data sources described earlier in this chapter to characterize the school environment: (1) type of school attended (i.e., public, private, charter, or home-school), (2) extent of poverty as measured by percentage of children eligible for free or reduced-price lunch, (3) racial and ethnic distribution of the children enrolled in the schools, and (4) percentage of students in the study children's schools who scored at the "proficient level" on state math and reading assessments.
  - **Teacher and classroom characteristics.** Teachers of the sampled children were asked about their education, certification, number of years teaching, and their educational beliefs. The teacher qualification variables represent the sampled children's teachers, not all the teachers in a school. The measures are:

[49] The 12 variables incorporated in the quality composite are (1) ECERS-R or FDCRS, (2) the Arnett, (3) literacy activities, (4) math activities, (5) other activities, (6) staff/child ratio, (7) teacher/care provider education, (8) teacher/care provider ECE coursework or CDA, (9) teacher/care provider training, (10) parent involvement, (11) home visits, and (12) program services to children and families.

- Teaching certificate. To cover the different state licensing requirements, teachers were asked whether they had a state teaching certificate, teaching license, or teaching credential. Findings are reported as "yes" if a teacher reported having any of these three credentials.
- Teacher education. Teachers reported their highest level of education attainment, and this information was used to create a dichotomous variable as having a bachelor's degree or higher vs. less than a bachelor's degree. Also reported is the mean number of college courses that teachers completed in elementary education, early childhood education, methods in teaching reading, and methods in teaching math.
- Teaching experience. Teachers were asked the total number of years they had taught and the number of years employed at their current school. Teacher experience is reported as the mean number of years teaching and the mean number of years employed at the current school across sampled children.
- Teacher beliefs and attitudes. A Teacher Belief Scale (Charlesworth, Hart, Burts, Mosley & Fleege, 1993) in the teacher survey measured responses to 14 statements about how children should be taught and managed in the classroom. Teachers responded to each statement using a five-point Likert scale of one to five (strongly disagree, disagree, neither agree nor disagree, agree, or strongly agree). An average score was developed for each child based on his/her teacher's responses. A higher score indicates more positive practices. Teachers were also asked if they enjoyed their present teaching position, whether they would choose teaching again as a career, and if they believed they were making a difference in the lives of children.

In addition to teacher characteristics, four teacher-reported measures are used to provide some indication of the overall classroom environment. These classroom characteristics are:

- Presence of teaching assistants. Teachers were asked whether their classrooms had at least one paid assistant, co-teacher, or team teacher in the class in a typical week. A dichotomous variable was created distinguishing classrooms where any of these people were in the classroom from those where none of these people were in the classroom.
- Use of volunteers in the classroom. Teachers were asked whether their classrooms had at least one adult volunteer assistant in the class in a typical week. A dichotomous variable was created distinguishing those classrooms with at least one volunteer from those who did not have any.
- The percentage of Dual Language Learners. Teachers reported the number of students with limited English proficiency in each sampled child's classroom.
- Behavior of children in the classroom. Teachers reported whether children in the classroom as a group were well-behaved, misbehaved occasionally, or misbehaved frequently.

- **Classroom activities.** Kindergarten teachers were asked how often they did the same 12 reading and language activities and eight math activities as pre-school teachers. First-grade teachers were asked about 25 reading and language activities and 18 math activities. As with preschool, these items are an indicator of the availability and variety of learning opportunities provided in the classrooms.

Analysis Sample

The unit of analysis for all impact analyses is the child. This is true irrespective of the outcome measure or data source considered; even outcomes reported by parents, caregivers, and teachers are weighted and analyzed according to the children they describe. This makes all impact findings representative of all newly entering Head Start children in the nation in 2002 in communities in which there were more potential program participants than funded Federal Head Start slots. The weights applied make each child in this universe count equally, not each parent/caregiver/teacher nor each Head Start center nor each grantee/delegate agency.

This section describes the construction and characteristics of the annual samples of children used to measure Head Start's impact in spring 2003, 2004, 2005, and 2006. Similar information is provided for the longitudinal sample used to examine trends in outcomes and impacts over time. For both purposes, the set of completed questionnaires and assessments was divided into two separate samples, one for children entering Head Start one year before anticipated kindergarten entry—referred to as the 4-year-old cohort—and one for children entering Head Start two years prior to expected kindergarten entry—the 3-year-old cohort.

The annual cross-sectional samples are chosen to maximize the data available for analysis each spring. Thus, they include every completed child assessment, parent interview, or teacher-child report for that year, depending on which of these instruments is the source of the particular outcome measure being examined. Information from the parent interview and child assessment is used even when the other source is missing. Use of teacher-child reports is contingent upon having a parent interview and child assessment. In each instance, the comparability of the treatment and control group samples established at random assignment is maintained to the greatest extent possible by adjusting the initial sampling weights to offset observable differences between respondents and nonrespondents for each cohort (see the Technical Report for the Head Start Impact Study).

For cognitive outcomes, the annual analysis samples are subdivided according to the primary language used to assess the child. Children who initially (in fall 2002, as they entered Head Start) could not complete all the assessment batteries in English, had their assessments administered principally in Spanish or, for a small fraction of the sample, some other non-English language. When examining Head Start's impact on selected cognitive outcomes—those outcomes pertinent to Spanish-speaking children only—we report separately on the set of children assessed initially (in fall 2002) in Spanish and then switched in spring 2003 to English as the primary language of assessment.[50] In addition, all of the children in the study sample from Puerto Rico began with Spanish-language assessments and continued exclusively in that language throughout the study period (since transition to bilingualism through English acquisition does not commonly take place until 4th grade). Because cognitive measures administered in different languages are not directly comparable, Puerto Rican children are analyzed separately from their "mainland" counterparts, and the findings from Puerto Rico are found in Appendix F of this volume.

The characteristics of the children and families in the 2003 analysis sample, using characteristics measured at baseline in fall 2002, are presented in Exhibits 2.9-A for the 4-year-old cohort, and 2.9-B for the 3-year-old cohort.

In these tables, observations are weighted to reflect the share of the national population they represent, i.e., each spring's impact analysis sample seeks to replicate this population as accurately as possible even though not every child can be included every year. For example, only respondents to the spring 2003 child assessments are included in Exhibits 2.9-A and 2.9-B, but the data are weighted to adjust for any observed nonresponse.[51] Any noted differences between the program and control groups in these tables reflect any remaining differences due to nonresponse after the weighting adjustments, chance differences between the treatment and control groups created at random assignment, and possible early impacts of Head Start in fall 2002 before baseline data could be collected. Given the small size of these differences and their

[50] The two cognitive assessments administered to these children in Spanish in fall 2002 and again in spring 2003 (when the children had advanced sufficiently in their English language skills to be assessed primarily in English) were the TVIP (adapted) and the Woodcock-Muñoz Letter-Word Identification Test.

[51] The Technical Report for the Head Start Impact Study reports the characteristics of the analysis sample each spring, both before and after weighting for nonresponse.

almost universal statistical insignificance, it does not appear that important imbalances arose from any of these sources (see Exhibits 2.9-A and 2.9-B).

Exhibit 2.9-A: Description of the Year 1 Analysis Sample: 4-Year-Old Group (Weighted Data)

Characteristic	Head Start Group	Control Group	Difference: Head Start – Control
Child Gender:			
Boy	49.6%	51.2%	-1.6
Girl	50.4%	48.8%	1.6
Child Race/Ethnicity:			
White	27.8%	24.6%	3.2%
Black	25.5%	23.3%	2.2%
Hispanic	42.4%	45.8%	-3.4%
Other	4.3%	6.2%	-1.9%
Child Has a Disability	12.8%	11.4%	1.4%
Fall-Spring Language of Child Assessment:			
English-English	67.2%	64.3%	2.9%
Spanish-English	25.9%	28.3%	-2.5%
Spanish-Spanish	5.9%	5.4%	0.4%
Primary Home Language Is English	63.6%	63.2%	0.0%
Biological Mother Was a Teen Mom	38.6%	35.2%	3.4%
Biological Mother Is a Recent Immigrant	24.1%	23.5%	0.6%
Biological Mother Is Employed	48.5%	52.0%	-3.4%
Both Biological Parents Live with Child	51.3%	51.3%	0.0%
Child's Parents Are:[a]			
Married	45.2%	45.4%	-0.2%
Separated or Divorced	15.9%	14.9%	1.0%
Primary Caregiver's Age as of 9/1/02	29.3 years	29.5 years	-0.2 years
Mother's Education:			
Less Than High School	38.6%	41.6%	-3.0%
High School/GED	31.7%	35.2%	-3.5%
Beyond High School	29.8%	23.3%	6.5%*
Grandparent Lives in Home	2.4%	1.4%	1.0%
Parent's Self-Reported Health Is Excellent or Good	86.6%	86.4%	0.1%
Average Household Income:			
$500/month or less	11.8%	9.1%	2.7%
$501-$1500/month	46.2%	50.8%	-4.6%
Over $1500/month	42.0%	40.0%	2.0%
Household Receives TANF	10.0%	14.4%	-4.5%*

*= $p \leq 0.05$, ** = $p \leq 0.01$, *** = $p \leq 0.001$.

Data source: Roster information collected at the time of random assignment and fall 2002 Parent Survey.

Note: Due to rounding, the sum of the percents within a cell may not equal 100 percent.

[a]Never married and widowed are excluded from this characteristic.

Exhibit 2.9-B: Description of the Year 1 Analysis Sample: 3-Year-Old Group (Weighted Data)

Characteristic	Head Start Group	Control Group	Difference: Head Start – Control
Child Gender:			
Boy	47.9%	49.1%	-1.2%
Girl	52.1%	50.9%	1.2%
Child Race/Ethnicity:			
White	24.3%	26.0%	-1.7%
Black	33.3%	31.4%	1.9%
Hispanic	37.0%	36.4%	0.6%
Other	5.4%	6.3%	-0.8%
Child Has a Disability	13.5%	11.9%	1.6%
Fall-Spring Language of Child Assessment:			
English-English	75.4%	75.9%	-0.5%
Spanish-English	18.9%	18.0%	0.9%
Spanish-Spanish	4.3%	4.6%	-0.3%
Primary Home Language Is English	71.9%	68.5%	3.4%
Biological Mother Was a Teen Mom	36.2%	37.6%	-1.3%
Biological Mother Is a Recent Immigrant	17.0%	17.8%	-0.8%
Biological Mother Is Employed	51.4%	57.4%	-6.0%
Both Biological Parents Live With Child	48.5%	50.7%	-2.2%
Child's Parents Are:[a]			
Married	43.7%	45.3%	-1.6%
Separated or Divorced	11.5%	13.7%	-2.2%
Primary Caregiver's Age as of 9/1/02	29.5 years	28.6 years	0.9 years*
Mother's Education:			
Less Than High School	32.4%	34.8%	-2.3%
High School/GED	34.7%	33.9%	0.8%
Beyond High School	32.9%	31.4%	1.5%
Grandparent Lives in Home	3.6%	1.7%	1.9%**
Parent's Self-Reported Health Is Excellent or Good	85.5%	86.5%	-1.0%
Average Household Income:			
$500/month or less	14.8%	12.0%	2.9%
$501-$1500/month	48.3%	53.4%	-5.1%
Over $1500/month	36.9%	34.6%	2.3%
Household Receives TANF	10.6%	10.5%	0.1%

*= $p \leq 0.05$, ** = $p \leq 0.01$, *** = $p \leq 0.001$.

Data source: Roster information collected at the time of random assignment and fall 2002 Parent Survey.

Note: Due to rounding, the sum of the percents within a cell may not equal 100 percent.

[a]Never married and widowed are excluded from this characteristic.

The comparability of the analysis samples remained stable through all study years (see the Technical Report for the Head Start Impact Study). This is largely because, as shown in response rate tables (Exhibits 2.5 and 2.6), attrition from year to year was low.

Data Analysis

This section describes the procedures used to calculate the impact of Head Start on children's cognitive and social-emotional development, health outcomes, parenting practices, and children's preschool and early elementary school experiences. Findings in the report come from four types of impact analysis:

- Impact of access to Head Start on annual outcomes of all of the children randomly assigned to the Head Start group, in each of the age cohorts studied;
- Impact of participation in Head Start on annual outcomes, adjusting for the facts that some of the children in the Head Start group given access did not participate and some children in the control group who were not given access did participate;
- Impact of access to Head Start on children's growth trajectories—for example, the impact on change in children's language and literacy or math development over time; and
- Impact of access to Head Start on annual outcomes of selected subgroups of children defined by background characteristics of children and families (measured at baseline).

Our methods for obtaining estimates in each case are described below; details of all analysis methods are provided in the Technical Report for the Head Start Impact Study.

General Approach to Impact Estimation

For both the overall annual impacts and impacts on subgroups, the most unbiased estimate of Head Start's effect is the difference in average outcomes between children randomly assigned to the treatment group and children randomly assigned to the control group. Random assignment ensures that no systematic differences other than access to Head Start exist between the program and control groups on average. Therefore, differences in later outcomes can be attributed to Head Start's impact, not other confounding factors. Inclusion of background characteristics in the analysis, using methods described below, increases the precision of the analysis and its ability to detect any true program impacts that do occur.

The most basic version of the analysis contrasts the average outcome level for the treatment group with the average outcome level for the control group, using analysis weights.

The weights account for the different rates at which primary sampling units (PSU's), grantee/delegate agencies, centers, and children within centers were sampled and with which children of different types are absent from the follow-up data. Thus, the study sample can be used to accurately represent the national Head Start population of all children served in non-saturated communities. These weighted difference-in-means impact estimates are reported as the basic estimates for this report, along with statistical tests showing which of the measured outcome differences are unlikely to be the result of chance treatment-control group differences. Thus, they are probably impacts of the Head Start program.[52] A two-tailed test of the null hypothesis of no Head Start impact is used, to allow the possibility of program effects in either direction, up or down. Three different levels of statistical significance—i.e., three different probabilities of rejecting a true null hypothesis—are reported in the tables of results, alpha-values of 0.10, 0.05, and .01.

The analysis of main impacts generated a very large number of statistical tests, and the subgroup analyses discussed below generated even more. Such conditions increase the probability that one or more statistically significant differences will emerge by random chance alone in the absence of a true impact—an event known as a "false discovery." To guard against false discoveries, Benjamini and Hochberg (1995) developed a statistical test designed to screen out marginally significant findings from large sets of impact estimates. This procedure was applied to the complete set of outcomes within each domain (cognitive, social-emotional, health, and parenting outcomes) for the main analysis and for each of the specified subgroups. This was done separately for each of the two study cohorts and for each year for which impact estimates were generated. Because the Benjamini-Hochberg test is conservative (i.e., it limits discovery of true impacts below what conventional test procedures do), we present findings both with and without the test's standards applied. We have established three separate categories of statistically significant results[53] and use this language throughout this report:

---

[52] Differences in mean impact estimates were created by regressing the outcome variable on a constant and a 0/1 indicator variable for membership in the treatment group. The estimated coefficient for the indicator variables equals the treatment group mean outcome minus the control group mean outcome.

[53] If we fail to identify a statistically significant difference, we do not have conclusive evidence that the program "doesn't work." Rather, statistically insignificant impacts mean that the effect is indeterminate—access to Head Start may or may not have had a non-zero impact on a particular outcome, and we cannot with this study sample make a confident conclusion either way. The one thing that will be known with confidence is that a large true impact has not occurred.

- **Strong Evidence*:*** the estimated impact for a particular outcome is statistically significant at the typical level ($p \leq 0.05$), and this result holds up under the test for multiple comparisons.
- **Moderate Evidence*:*** the estimated impact for a particular outcome is statistically significant at the typical level ($p \leq 0.05$), but this result *does not* hold up under the test for multiple comparisons.
- **Suggestive Evidence*:*** the estimated impact for a particular outcome is statistically significant under a relaxed standard ($p \leq 0.10$), and this result *may or may not* hold up under the test for multiple comparisons.

While an intact randomized sample and complete outcome data ensure that no systematic biases enter into the difference-in-mean estimates of Head Start's impact, more sophisticated analysis methods provide further advantages. In addition to assignment to the Head Start treatment group, other factors such as a child's background and family characteristics and the initial starting points for the key outcome measures may influence his/her outcomes in later years. If these factors can be included in models that "explain" child outcomes as the joint result of Head Start access, demographic background characteristics, and pre-intervention values of the outcome, uncertainty about the process used to generate outcomes will decline. As a result, confidence in the role of each measured factor, including assignment to the treatment group, will increase. This effect, known statistically as "reducing variance," will increase the chances of detecting as statistically significant any impact Head Start has on the outcomes of interest.

To add the explanatory power of child and family background factors to the analysis key demographic variables measured in fall 2002 were included as covariates. All of the analyses include the same set of demographic variables (shown in Exhibit 2.10) as covariates, irrespective of the age cohort, outcome, and follow-up year. The same set of covariates is also used for every subgroup analysis described in the next section. The selected variables meet two criteria: (1) they likely correlate with child and family outcomes (and thereby help to increase the explanatory power of the model), and (2) they could not have been influenced by Head Start during the first weeks of participation (i.e., prior to the time they were measured).

Exhibit 2.10: Baseline Demographic Variables Included in the Statistical Models Estimating the Impact of Head Start

Child Covariates
Child Gender
Child Age at Time of Spring Assessment
Child Race/Ethnicity (White/Other, Black, Hispanic)
Child's Primary Language at Baseline (English vs. Spanish/Other)
Number of Weeks Elapsed between 9/1/2002, and Fall Testing (for child assessment outcomes)
Parent Covariates
Primary Language Spoken at Home (English vs. Spanish/Other)
Primary Caregiver's Age as of 9/1/2002
Both Biological Parents Live with Child
Biological Mother Is a Recent Immigrant
Mother's Highest Level of Educational Attainment (less than high school, high school, beyond high school)
Mother's Marital Status (not married, married, separated/divorced/widowed)
Mother Gave Birth to Study Child as a Teenager (i.e., 19 years old or younger)
Number of Weeks Elapsed between 9/1/2002, and Parent Interview (for all other outcomes derived from the parent interview)

The precision of the analyses is further enhanced by adding a pre-intervention measure of the outcome variable to the impact equation as a covariate. When a pre-intervention measure of an outcome was not available, a closely related baseline measure was used instead. Exhibit 2.11 lists the pre-intervention variable used for each outcome measure.

Missing values on the baseline background factors, due to both item and instrument nonresponse in fall 2002, were imputed using hot deck imputation, a procedure in which observations with missing values are filled in with values from the same time point taken from children with similar baseline characteristics. To ensure comparability, the distribution of each variable was compared before and after imputation to check that the imputation procedures had not appreciably changed the dataset.

Exhibit 2.11: Pretest Measures Used in All Impact Analyses, by Outcome Measure

Outcome Measure	Fall 2002 Measure Used as a Covariate
Cognitive Domain	
Peabody Picture Vocabulary Test (PPVT)*	PPVT
Comprehensive Test of Phonological and Print Processing (CTOPPP) Elision Subtest *	PPVT
Letter Naming Task *	PPVT
Color Naming	Color Naming
Counting Bears	Counting Bears
McCarthy Scales of Children's Abilities Draw-a-Design Subtest	McCarthy Scales of Children's Abilities Draw-a-Design Subtest
Woodcock-Johnson III: Letter-Word Identification *	Woodcock-Johnson III: Letter-Word Identification
Woodcock-Johnson III: Spelling*	PPVT
Woodcock-Johnson III: Applied Problems*	*For Children Assessed Primarily in English in Fall 2002* Woodcock-Johnson III: Applied Problems *For Children Assessed Primarily in Spanish in Fall 2002* Woodcock-Muñoz Problemas Aplicados
Woodcock-Johnson III: Oral Comprehension*	PPVT
Woodcock-Johnson III: Pre-Academic Skills Composite*	PPVT
Woodcock-Johnson III: Writing Samples*	PPVT
Woodcock-Johnson III: Passage Comprehension*	PPVT
Woodcock-Johnson III: Calculation*	*For Children Assessed Primarily in English in Fall 2002* Woodcock-Johnson III: Applied Problems *For Children Assessed Primarily in Spanish in Fall 2002* Woodcock-Muñoz Problemas Aplicados
Woodcock-Johnson III: Academic Applications* Composite	PPVT
Woodcock-Johnson III: Academic Skills Composite*	PPVT
Woodcock-Johnson III: Basic Reading Skills*	PPVT
Woodcock-Johnson III: Math Reasoning*	*For Children Assessed Primarily in English in Fall 2002* Woodcock-Johnson III: Applied Problems *For Children Assessed Primarily in Spanish in Fall 2002* Woodcock-Muñoz Problemas Aplicados
Woodcock-Johnson III: Word Attack*	PPVT

*Fall measure was residualized separately by English and Spanish language groups.

Exhibit 2.11: Pretest Measures Used in All Impact Analyses, by Outcome Measure (continued)

Outcome Measure	Fall 2002 Measure Used as a Covariate
Cognitive Domain (continued)	
Woodcock-Johnson III: Quantitative Concepts*	*For Children Assessed Primarily in English in Fall 2002* Woodcock-Johnson III: Applied Problems *For Children Assessed Primarily in Spanish in Fall 2002* Woodcock-Muñoz Problemas Aplicados
Test de Vocabulario en Imágenes Peabody (TVIP)	Test de Vocabulario en Imágenes Peabody (TVIP)
Batería Woodcock-Muñoz Pruebas de aprovechamiento-Revisada: Identificación de letras y palabras	Batería Woodcock-Muñoz Pruebas de aprovechamiento-Revisada: Identificación de letras y palabras
Parent (reported) Emergent Literacy Scale (PELS)	Parent (reported) Emergent Literacy Scale (PELS)
Social-Emotional Domain	
Social Skills and Positive Approaches to Learning	Social Skills and Positive Approaches to Learning
Total Child Behavior Problems Scale	Total Child Behavior Problems Scale
Aggressive Behavior Scale	Aggressive Behavior Scale
Hyperactive Behavior Scale	Hyperactive Behavior Scale
Withdrawn Behavior Scale	Withdrawn Behavior Scale
Pianta Scale: Closeness	None
Pianta Scale: Conflict	None
Pianta Scale: Positive Relationship	None
Social Competencies Checklist	Social Competencies Checklist
Parenting Practices Domain	
Parent used time out in the last week	Parent used time out in the last week
Parent spanked child in the last week	Parent spanked child in the last week
Parental Safety Practices Scale	Parental Safety Practices Scale
Family Cultural Enrichment Scale	Family Cultural Enrichment Scale
Parenting Style Is Authoritarian	Parenting Style is Authoritarian
Parenting Style Is Authoritative	Parenting Style is Authoritative
Parenting Style Is Neglectful	Parenting Style is Neglectful
Parenting Style Is Permissive	Parenting Style is Permissive
Child was read to every day in the last week by parent or other family member	Child was read to every day in the last week by parent or other family member

*Fall measure was residualized separately by English and Spanish language groups.

Exhibit 2.11: Pretest Measures Used in All Impact Analyses, by Outcome Measure (continued)

Outcome Measure	Fall 2002 Measure Used as a Covariate
Health Domain	
Child seen by dentist since last September	Child seen by dentist since last September
Overall child's health status	Overall child's health status
Child has injury in last month requiring medical treatment	Child has injury in last month requiring medical treatment
Child has health insurance	Child has health insurance
Child has condition that requires ongoing medical care	Child has condition that requires ongoing medical care

* Fall measure was residualized separately by English and Spanish language groups.

Most of the fall 2002 data were collected during a three-month period from October 2002 through December 2002 (with most completed by mid-November) at a considerable lag from random assignment (which took place from May to September 2002).[54] As a result, there is a possibility that Head Start had some early impact on these measures (some measures, such as demographic variables could not plausibly have been affected by Head Start (e.g., sex, race/ethnicity, and mother's age at time of study child's birth).

However, the fall 2002 versions of the cognitive, social-emotional, health, and parenting outcomes measures pose a larger problem. That is, if impacts of Head Start occurred quickly that fall, inclusion of the unadjusted fall measures as covariates in the impact equation would attenuate the impact estimate, since the portion of the impact achieved prior to fall 2002 data collection would be removed from the impact estimate. To avoid this problem, all fall 2002 measures of outcome variables used as covariates are "residualized" before inclusion in the impact regressions. The "residualization" procedure described in the Technical Report for the Head Start Impact Study removes any systematic differences between treatment and control group levels in the fall measures, including those potentially due to Head Start's impact. The procedure has the drawback that the covariates produced are no longer capable of controlling for purely chance differences between treatment and control groups on the fall factors involved. This reduces, but does not eliminate, the value of the covariates in increasing the statistical precision of the impact estimates.

[54] It was not feasible to conduct parent interviews and administer child assessments prior to randomization in this study due to a confluence of circumstances explained in the Technical Report for the Head Start Impact Study.

For all statistically significant results, the impact estimates in their initial units are converted into effect sizes by dividing by the measured standard deviation of the outcome in the population. This provides a yardstick for gauging the quantitative importance of the estimated impact in relation to the natural variation of the outcome. Effect sizes tell us how much the improvements produced by Head Start move children upward in the distribution of outcomes that would have prevailed had no Head Start intervention been available.

Many researchers have used Cohen's (1987) guidelines for interpreting the relevance of effect sizes, with an effect size of 0.2-0.5 being considered small, 0.5-0.8 moderate, and over 0.8 large. Within the field of education research, some researchers have argued that an effect size has to be at least 0.25 or 0.33 of a standard deviation to be considered "educationally meaningful" (Slavin, 1990; Wolf, 1986).

In contrast, Glass et al. (1981) and McCartney and Rosenthal (2000) have asserted that the effect sizes derived from a given study always should be interpreted within the context of the empirical literature on comparable interventions designed to produce similar effects. In the NICHD Study of Early Child Care, the quality of child care predicted children's cognitive performance at 54 months (range of effect sizes was 0.04 to 0.08). The Tennessee study examining the benefits of smaller class sizes in the early school grades yielded effect sizes that ranged between 0.13 and 0.27 on several direct assessments of children's reading and math performance (Finn & Achilles, 1990). A meta-analysis of evaluations of family support programs yielded the following weighted mean effect sizes across several key outcome domains: children's cognitive development (0.253), social-emotional development (0.258), physical health and development (0.091), parenting attitudes and knowledge (0.182), parenting behavior (0.246), and family functioning/family resources (0.284) (U.S. Department of Health and Human Services, 2001). Finally, another recent meta-analysis of 33 studies focusing primarily on early childhood education programs for low-income 3- and 4-year-olds revealed a weighted mean effect size of 0.118 across the studies reviewed (Aos, Lieb, Mayfield, Miller, & Pennucci, 2004). For this report, effect sizes are provided but no descriptive categorizations are included due to the lack of clarity in the guidelines for reporting effect sizes.

Estimating Variations in Impact

In addition to looking at Head Start's impact *on average*, it is important to understand how impact *varies* among different types of participants. Congress mandated that the study identify circumstances under which the program achieves its greatest impact, in terms of both child and family circumstances—what works best for which children? There is also interest in determining whether the benefits of Head Start measured for children and families in general are widespread—i.e., whether gains compared to the control group occur for many types of children and families, or whether the overall gain reflects big gains for some participants and little or no gain for others.

There are many examples of variations in program impact that would have important policy or program implications. Analyses can spotlight groups of children who are not advanced by their participation and suggest needs for program improvement. For example, Head Start programs are required to serve children with special needs, making it important to understand the extent to which these children benefit from their participation.

To examine who benefits, and under what circumstances, the study analyzed the impact of Head Start on subgroups of program participants, defined by child and family characteristics at baseline. Some typical questions to be addressed in this realm include: Does Head Start help Dual Language Learners? Does it help children with the lowest academic achievement scores at baseline?

A computationally efficient and statistically powerful way to examine such subgroup impacts uses an extension of the impact regression models discussed above. Interaction terms are added between the subgroup-defining variables and the indicator of Head Start assignment. The coefficients on the interactions show how impacts vary between subgroups. Impacts for the individual subgroups can be obtained from the regression, as can differences in impact between any two subgroups. For example, for subgroups defined by a mother's race/ethnicity, a single regression can provide information on how large an impact Head Start had on children of White mothers, children of Black mothers, and children of Hispanic mothers, as well as how impacts vary across these subgroups. Separate regressions are run for each subgroup-defining factor, referred to in this report as a subgroup "dimension," such as special needs status, household risk index, and urban/rural location. Each analysis tests for variation in impact along a particular

dimension, using the entire sample of children available in a given age cohort—an improvement in terms of statistical precision over separate analysis of each subgroup one at a time.

Exhibit 2.12 lists the dimensions used to define subgroups for analysis. These dimensions were chosen in advance of conducting the analyses, based on their program and policy importance to the Administration for Children and Families, on past Head Start and child development research, and recommendations from the Advisory Committee. Impacts are examined for each subgroup, separately by age cohort, to determine how widespread the benefits of Head Start might be and to identify the child/family types that benefit and the program characteristics associated with impact.[55] Details of subgroup dimensions used, their rationales, and the subgroup impact estimation method appear in Chapter 8, as well as in the Technical Report for the Head Start Impact Study.

Exhibit 2.12: Factors Used To Define Subgroups

Child and Family Characteristics*
Special needs (Yes/No)
Child's Pre-Academic Skills 2002 baseline as assessed on the Woodcock-Johnson III Pre-Academic Composite Measure (Yes/No)
Child's Home Language (English, Dual Language Learner)
Biological Mother/caregiver race/ethnicity (White, Black, Hispanic)
Household Risk Index (Low/No, Medium, High)
Urbanicity (Urban, Non-Urban)
Parent/Caregiver-Reported Depressive Symptoms (No, Mild, Moderate, or Severe Depressive Symptoms)

*These subgroups are defined in Chapter 8.

Longitudinal Growth Analysis

Additional impact analyses were done using Hierarchical Linear Modeling (HLM) [56] to analyze the impact of Head Start on children's growth trajectories, for example, the impact on

[55] Any subgroup estimate (e.g., for a particular outcome for a specified subgroup) based on fewer than 40 observations in either the treatment or control group is not reported.

[56] HLM is an approach to analyzing data that have the capability of estimating individual and group characteristics on individual-level processes. Research questions can be framed in terms of the complex groupings that exist in natural settings and to pose multi-level hypotheses that single-level analyses cannot always directly address.

the change in children's language and literacy or math development over time. Such an analysis uses information from the repeated measures taken at each wave of the study on a particular student. As described in more detail in the Technical Report for the Head Start Impact Study, a three-level model was estimated for each of five cognitive outcome measures. At level one of the regression, the cognitive outcome is modeled as a function of time using the child's age at time of testing (this made it possible to control for the fact that there were different intervals of time between waves for different students). The resulting model defines an individual growth curve for each student. The variability among students in growth was then modeled at level two, including the variability induced by access to Head Start since some students are randomly assigned to the program and others are not. The coefficient on this factor gives the impact estimate. The third level of the model captured variation of mean center-level outcomes.

A separate longitudinal sample and associated analysis weights were created to conduct this analysis, encompassing all children with two or more data collection points between fall 2003 and spring 2005 (spring 2006 for the 3-year-old cohort). The longitudinal weight allows these children to represent the population of children who applied for their first year of Head Start in fall 2002 from that point until the end of 1st grade.

Impacts on Head Start Participants

The main impact estimates in this report measure the effect of Head Start on the average child randomly assigned to the Head Start treatment group—that is, the impact of *granting access to Head Start services* for the population randomized. However, not all of the children given access to Head Start in the study sites actually participated in federally funded Head Start services, the intended treatment. This is not an unexpected phenomenon, in the normal course of events, some children and families accepted into Head Start never participate. As a result, there are two different versions of the key research question of this study:

- How much does Head Start help the typical child and family *admitted to* the program, on average?
- How much does Head Start help the typical child and family who *actually participated* in Head Start, on average?

Answers to both questions matter for policy and program administration purposes. As a comparison, consider the importance of understanding the health benefits of a diet and exercise

program both (1) if followed as prescribed and (2) as imperfectly followed by the average patient. The Head Start program can offer opportunities to participate, but it cannot compel any child to attend. Hence, the impact of admission into the program measures the typical result of what grantees do—provide access—rather than the effect of delivering services to every selected child and family. Yet the question of how much children gain from actually participating in Head Start's services remains an important one. For local programs at full attendance (not simply full enrollment, on paper), impacts per participant correspond to the impacts of offering access. Furthermore, when considering whether to expand or contract a fully attended center, the value of the program slots that might be added depends on the gains provided to the children who actually occupy those slots compared to children completely closed out of the program by capacity constraints. Finally, if impacts per participant are large but impacts per admitted child are comparatively small because of low participation, the evaluation will highlight the value of increasing participation rates.

In addition to children given access to Head Start but not attending (known as no-shows), some of the families of children randomized into the *control group* managed to get their children into Head Start anyway (as happens in many social experiments). This subpopulation is known as "crossovers." The Head Start Impact Study had no way to fully ensure that the children and families randomly assigned to the control group did not participate in federally funded Head Start.[57] A total of 17.6 percent of the weighted sample in the control group are known to have participated in a federally funded Head Start program for some time during the first year of the study. The presence of no-shows and crossovers changes the meaning of the experimental comparison between the full treatment group and the full control group; it becomes the impact of *intent to treat* (ITT). The impact of actual *receipt* of the Head Start intervention (compared to non-receipt)—remains important to policy for the reasons stated. This leads to interest in estimates of the "*impact on the treated*" (IOT), which show how Head Start affects the outcomes

[57] The grantees and delegate agencies whose applicants made up the research sample agreed not to serve those families using Federal Head Start funds during the 2002-03 program year. But other grantees and delegate agencies in nearby communities (or, in the case of several large cities, in overlapping neighborhoods) did not enter into such agreements and, for reasons of privacy, could not be told the identities of the children and families involved in the study, even had agreement been reached not to serve them. Moreover, no mechanisms existed for enforcing the commitments made by the participating grantees and delegate agencies. In light of these limitations and the strong attraction of Head Start to many families, it is not surprising that a number of families from the control group in fact obtained Head Start services for their children during that year.

of a set of children who participate in Head Start compared to what would have happened to those same children had none of them participated.

While methods to adjust for no-shows are fairly straightforward, adjusting for crossovers is more challenging. After assessing multiple research options (as described in the Technical Report for the Head Start Impact Study) we concluded that the best way to provide information on Head Start's IOT impact is through the use of an "instrumental variable" (IV) methodology. This methodology uses assignment to the treatment group as a statistical instrument for participation in the program. This method, long known in the econometric literature and applied in recent years to random assignment evaluations of government social programs,[58] treats crossovers symmetrically with no-shows and adjusts the initial ITT impact estimates to remove the influence of both groups in attenuating the magnitude of the estimates. This is achieved by dividing the ITT estimate by 1 minus the no-show rate minus the crossover rate to get an "impact on the treated" (IOT) impact estimate: IOT = ITT / (1 – n – c), where n is the no-show rate, and c is the crossover rate.

Like the classic "Bloom no-show adjustment" (Bloom, 1984), this methodology postulates that a program's overall impact on the treatment group accrues to just a subset of the sample. The Bloom assumption is that no impact occurs for no-shows since they are never exposed to the intervention; this is widely viewed as an innocuous assumption. The IV methodology further assumes that Head Start's impact on crossovers equals, on average, the program's impact on the corresponding children in the treatment group—i.e., the children who would have crossed over and participated in Head Start had they been assigned to the control group. The combination of these two assumptions makes it appropriate to reallocate the total impact of the program observed in the contrast between the full treatment and control groups to just the remaining set of children whose impacts are neither zero nor offsetting: essentially, the children who comply with the intention of random assignment by participating in Head Start when randomized into the treatment group and not participating when randomized into the control group. This average impact on "compliers" can be inferred from the ITT impact estimate

[58] See, for example, L.A. Gennetian, P.A. Morris, J.M. Bos, and H.S. Bloom (2005). Constructing instrumental variables from experimental data to explore how treatments produce effects. *Learning more from social experiments: Evolving analytic approaches*. H. S. Bloom (Ed.), New York: Russell Sage Foundation.

without further analysis, since the ITT estimate under the assumptions above is just the average impact on the "compliers" and a zero net impact on everyone else.

It should be noted, however, that for the IOT measure defined here to characterize *all* Head Start participants, it must be true that crossover-equivalent children in the treatment group experience the same impact on average as other Head Start participants in the treatment group.[59] Hence, the key assumptions in the crossover adjustment are:

- equal impact for "crossover-like" children in the treatment group and other Head Start participants in the treatment group; and
- equal outcomes on average for "crossovers" in the control group and "crossover-like" children in the treatment group.

These assumptions are discussed in detail in the Technical Report for the Head Start Impact Study which provides the reasoning and empirical evidence from the field to justify them in the Head Start context. As explained there, tests of statistical significance for the original ITT estimates apply as well to the IOT estimates calculated in this manner. That is, adjusting for no-shows and crossovers changes the magnitude of the estimated impacts, but not their statistical significance.

Understanding Impacts on the 3-Year-Old Cohort[60]

As noted earlier, the control group for the 3-year-old cohort was given access to Head Start, by study design, in the second year of the study. Excluding children from Head Start for two years was considered both infeasible and undesirable because it likely would undermine Head Start programs' willingness to participate in the study. The long-term goal for this study was to determine 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, initially and over the longer term, if the program did not enroll them until age four. This is not only important to

[59] This is necessary so that IOT = ITT/ (1 – n – c), which relates directly to just the "compliers" in the participant population, is a good estimate of the effect on all other participants as well—i.e., on the children in the treatment group who would be crossovers if assigned to the control group but who as part of the treatment group comprise part of the total set of all participants in the treatment group, the group of policy interest in IOT analysis.

[60] In a sense, one can think of the 4-year-old cohort as "delay-entry" eligible 3-year-olds, i.e., parents who had the option to enroll their child at age three but opted instead to 'hold them back" until age four. Not surprisingly, as discussed elsewhere, the parents who elected to seek enrollment for their child at age three are different from those who decided to wait until age four, and our data bear this out. For example, the 3-year-old cohort of newly entering children is more African American and the 4-year-old cohort of newly entering children is more Hispanic.

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. Therefore, the research design for the 3-year-old cohort only varied the first year of Head Start participation. Hence, impacts for the 3-year-old cohort reflect the benefits of being provided an *earlier* year of Head Start, rather than the effects of being provided two years of Head Start.

This difference has important implications for how IOT estimates are calculated. Control group members who attended Head Start in that second year are not considered crossovers, since they were not violating random assignment. Therefore, they are not part of the IV adjustment when computing IOT impact estimates in spring 2004 and beyond. It is only members of the 3-year-old control group who attended Head Start during that first year who are included in these IOT adjustments.

Impacts on Children's Experiences

It is also important to understand how Head Start affects the experiences children have during their preschool and early elementary school years. A great deal of data have been collected to characterize these experiences and contrast them between the Head Start and control groups to see how settings, activities, and adult interactions with children are altered by their access to Head Start. These include reports from parents, teachers, and center directors, and independent classroom and family day care home observations conducted by study staff during the preschool years. Information is also available from children's schools during the kindergarten and 1st grade years. The specific measures were discussed earlier in this chapter, and further discussion about children's program experiences is presented in Chapter 3. Chi-square tests were used to identify differences in distributions for each of the children's experience measures described earlier in this chapter.

Descriptive Analyses of Children's Experiences

Children in the control group were much more likely than those in the Head Start group to remain at home in parental care during the year they first applied to Head Start. Therefore, differences in Head Start's impacts on children's experiences may reflect both differences in the proportion of children who remained at home and differences in the characteristics of the care

settings for those in non-parental care (e.g., children at home can't be categorized as attending a high quality classroom because they did not attend a classroom). Consequently, in addition to providing impact estimates, we provide descriptive information on children's preschool experiences for the group of children who were in a non-parental care setting. Because these data represent non-random subsets of children in both study groups, the observed differences do not represent the impact of access to Head Start and are only provided for descriptive purposes.

One or Two Years of Head Start

As discussed earlier in this chapter, the Head Start Impact Study included two separate age cohorts: children who were age three at the time of application to Head Start (typically two years before the start of kindergarten), and children who were age four at the time of application (typically one year before the start of kindergarten). The 3-year-old Head Start cohort had the choice of participating in one or two years of Head Start. Approximately 60 percent of the Head Start group returned for a second year. What factors are related to families' decisions to remain with Head Start? Are there child or family characteristics driving the decision, or is the decision related more to the actual experiences that families had during their first year of Head Start? Information on these factors is derived from parent reports and classroom observations. This information can be useful for increasing Head Start's ability to retain these families. Four categories of measures were examined: (1) child characteristics, (2) parent and household characteristics, (3) parental satisfaction with child's first year of Head Start, and (4) classroom characteristics.

Some of the measures represent information that was collected at baseline and were not affected by Head Start participation (e.g., mother recent immigrant, home language). Other measures were more likely influenced by children's participation in a year of Head Start and are presented as of spring 2003. In Chapter 3, simple bivariate relationships are presented to provide descriptive information on which children remain in Head Start for two years as compared to one year.

Further details of study methodology are presented in the Technical Report for the Head Start Impact Study.

Chapter 3: Head Start, Child Care, and Early School Experiences

Introduction

This chapter, which describes the experiences of children and the services they receive from the start of Head Start through the end of 1st grade, is divided into six sections: (1) a discussion of the data sources and measures used in these analyses; (2) an overview of the variation across Head Start centers on selected children's experiences; (3) estimates of the impact of access to Head Start on the types of child care settings children attended for preschool; (4) a discussion of the experiences children had in those settings, including data on a variety of program and service variables; (5) a discussion of what factors were related to whether 3-year-old children attended Head Start for one or two years; and (6) a discussion of some of the characteristics of the schools and classrooms that children attended for kindergarten and 1st grade. A subsequent report will cover the period through the end of 3rd grade and will provide more detailed information about children's early school experiences during that time.

The information presented in this chapter is important for two reasons. First, it provides details on how random assignment affected the opportunities and choices of families regarding their children's early childhood care and education experiences. Specifically, it examines what the "treatment" was in this study—the difference random assignment made in whether children actually attended Head Start and how the experiences of the Head Start group children differed from those of the children in the control group. Second, the information provides descriptive details on the experiences and services that children and families received for the children who attended non-parental child care settings. (i.e., care outside the home). Although this study cannot isolate whether a particular experience contributed specifically to a child's or parent's outcome(s), the study can describe the nature of children's Head Start experience and highlight how it differed from the experiences of equivalent children for whom access to Head Start was not available (i.e., the control group). These differences in early childhood experiences drive the impacts on child and parent outcomes discussed in the following chapters and provide the reader with a better understanding of how, and why, the observed effects arose.

Data Sources and Measures

The data for these analyses come from a variety of sources, including both information specifically collected as part of this study and secondary data on the schools that the children attended. Exhibit 3.1 provides a summary of the data sources used by applicable year in school. Further discussion about the data sources is provided in Chapter 2 and in the Technical Report for the Head Start Impact Study.

Exhibit 3.1: Data Sources for Experiences by Children's School Year

Data Source	Preschool	Kindergarten/1st Grade
Center Director Interview	✓	
Teacher Survey	✓	✓
Care Provider Interview	✓	
Classroom/Child Care Home Observation	✓	
Parent Interview	✓	✓
NCES, CCD and PSS (secondary analysis)[a,b]		✓
Great Schools (secondary analysis)[c]		✓

[a] U.S. Department of Education. Institute of Educational Sciences. National Center for Education Statistics (NCES). Common Core of Data (CCD). Retrieved at www.nces.ed.gov/ccd.

[b] U.S. Department of Education, Institute of Education Sciences. National Center for Education Statistics. Private School Universe Survey (PSS). Retrieved at www.nces.ed.gov/surveys/pss.

[c] *Great schools: The parent's guide to K-12 success*. Retrieved at www.greatschool.net.

Preschool Measures

The preschool measures were selected both because of their relevance to the Head Start Performance Standards and because of an important role they may play in children's subsequent development and overall preschool experiences. The following measures are presented across six constructs, and further description of individual measures is presented in Chapter 2:

- **Type of preschool setting attended by children.** Information was obtained from parent interviews each spring to identify a focal child care setting for each study child. The focal setting is defined as the child care arrangement that accounted for at least five hours between Monday and Friday between the hours of 8 am and 6 pm.
- **Center environment.** Information on the operation and overall environment of Head Start and non-Head Start center-based preschool programs was collected from interviews with center directors. Measures span an array of center characteristics, including director qualifications, teacher turnover, use of a specific curriculum, center

size, center affiliation, extent of competition from other preschools, and services available for children and families. Where applicable, similar measures were collected on the environment of family day care (e.g., caregiver qualifications, training, use of a curriculum, services available for children and families).

- **Teacher training and qualifications.** The teacher measures in this construct are specific to the teachers of the study children, whereas the teacher training and mentoring measures mentioned above in the center environment reflect the availability of training and mentoring center wide. Teacher education is measured at various levels of education completion (e.g., received BA, AA, early childhood education coursework). Also, the amount of mentoring and training that the teachers of study children received is presented.
- **Classroom environment.** The classroom environment construct is intended to capture overall quality of the preschool classrooms that children attended. Measures include the Early Childhood Environmental Rating Scale-Revised (ECERS-R), the Family Day Care Rating Scale (FDCRS), the Arnett Scale of Lead Teacher Behavior, child/staff ratio, and parental involvement in the classroom. Data were collected on both center and family day care settings.
- **Classroom activities.** Teachers were asked how often (e.g., daily, weekly) they used each of 12 reading and language activities, eight math activities, and four other instructional activities (e.g., arts and crafts, sports). Data were collected on both center and family day care settings.
- **Overall quality of children's experience.** This measure was developed using 12 variables to create an overall composite indicator of classroom quality. Variables include ECERS-R/FDCRS, the Arnett, math activities, language and literacy activities, other activities, staff/child ratio, teacher education, early childhood education coursework, teacher training, parent involvement, home visits, and program services available to children and families.

The same measures were used for Head Start and other center-based programs. As indicated above, when applicable, similar measures were used for family day care homes.

Early Elementary School Measures

School measures are presented in three categories covering both kindergarten and 1st grade: (1) school environment, (2) teacher and classroom characteristics, and (3) classroom activities. Early elementary school measures included information from secondary data sources on the overall school environment and teacher input on the teacher and classroom characteristics and activities. The measures are discussed in Chapter 2 and include the following:

- **School environment.** Measures to characterize the school environment include type of school, percentage of children eligible for free or reduced-price lunch, racial/ethnic composition of students, and percentage of students in the study children's schools who scored at the "proficient" level on state and math reading assessments.
- **Teacher and classroom characteristics.** This category comprises both teacher characteristics (education, experience, and beliefs) and classroom characteristics. Classroom characteristics include four teacher-reported measures: teaching assistants, use of volunteers, percent of Dual Language Learners, and behavior of children in the classroom.
- **Classroom activities.** Frequency of language and literacy and math activities is reported. Also, reported is the percentage of time spent daily in child or teacher-directed activities as well as small-group or whole-class activities.

More detailed information about children's early school experiences will be provided in a subsequent report that will cover the period through the end of 3rd grade. The result of the analyses of children's experiences through 1st grade is presented below. The section begins with a discussion of the variation in Head Start centers across a subset of these measures. The chapter then provides estimates of the impact of access to Head Start on the types of preschool settings children attended. This is followed by a discussion of the programs and services children experienced in their preschool settings and then a description of the classrooms and schools children attended for kindergarten and 1st grade. Findings are presented separately for the 4-year-old and 3-year-old cohorts.

Variation in Characteristics of Head Start Settings

As discussed previously, the Head Start children in this study were randomly assigned at 383 nationally representative Head Start centers. The experiences and services available to children at these centers varied. Before addressing the differences in experiences between the Head Start and control group children, we present highlights of the variation in Head Start centers themselves (Exhibits 3.2 and 3.3). Measures highlighted below are a subset of the overall set of measures presented in this chapter and are discussed in detail in Chapter 2. As shown in the exhibits, both the 3- and 4-year-old cohorts had access to Head Start programs that were, on average, of high quality, but that varied widely in terms of environments, teacher or caregiver qualifications, and classroom or program characteristics.

Exhibit 3.2: Percentage of 4-Year-Old Head Start Group Children in Head Start Centers by Center Characteristics (Center Environment, Teacher Qualifications and Training, Classroom Environment and Activities)

Center Characteristics	4-Year-Old Cohort
Center Environment	
Center is always filled to capacity	
Yes	50.2
No	49.8
Center is affiliated with a school	
Yes	18.6
No	81.4
Low proportion of new lead teachers (low turnover)	
Yes	44.3
No	55.7
Center size	
Greater than 50 children	54.4
Less than or equal to 50 children	45.6
Teacher Qualifications and Training	
Teacher's highest degree was AA	
Yes	31.2
No	68.8
Teacher had BA or higher	
Yes	28.6
No	71.4
Teacher received 25 hours of training in past year	
Yes	40.5
No	59.5
Teacher received mentoring at least once a month	
Yes	42.3
No	57.7
Classroom Environment and Activities	
ECERS-R mean total score	
5 or greater	71.5
Less than 5	28.5
Number of language and literacy activities at least 3 times a week	
Provide at least 7 of 12	61.5
Provide fewer than 7	38.5
Number of math activities at least 3 times a week	
Provide at least 5 of 8	60.7
Provide fewer than 5	38.3
Number of other activities at least 3 times a week	
Provide at least 3 of 4	84.9
Provide fewer than 3	15.1
Child/teacher ratio	
Meets APHA standard	64.7
Does not meet standard	35.3

Exhibit 3.3: Percentage of 3-Year-Old Head Start Group Children in Head Start Centers by Center Characteristics (Center Environment, Teacher Qualifications and Training, Classroom Environment and Activities)

Center Characteristics	3-Year-Old Cohort
Center Environment	
Center is always filled to capacity	
Yes	37.4
No	62.6
Center is affiliated with a school	
Yes	6.2
No	93.8
Low proportion of new lead teachers (low turnover)	
Yes	44.0
No	56.0
Center size	
Greater than 50 children	62.8
Less than or equal to 50 children	37.2
Teacher Qualifications and Training	
Teacher's highest degree was AA	
Yes	29.7
No	70.3
Teacher had BA or higher	
Yes	30.1
No	69.9
Teacher received 25 hours of training in past year	
Yes	33.2
No	66.8
Teacher received mentoring at least once a month	
Yes	39.8
No	60.2
Classroom Environment and Activities	
ECERS-R mean total score	
5 or greater	69.0
Less than 5	31.0
Number of language and literacy activities at least 3 times a week	
Provide at least 7 of 12	55.7
Provide fewer than 7	44.3
Number of math activities at least 3 times a week	
Provide at least 5 of 8	66.5
Provide fewer than 5	33.5
Number of other activities at least 3 times a week	
Provide at least 3 of 4	85.4
Provide fewer than 3	14.6
Child/teacher ratio	
Meets APHA standard	58.2
Does not meet standard	41.8

Center environment. Approximately one-half of the 4-year-old Head Start group attended centers that were always filled to capacity, had high turnover, and had the capacity to serve 50 or more children. Approximately 20 percent of this cohort attended centers that were affiliated with a school.

Children in the 3-year-old cohort also attended centers with high turnover. They were more likely to attend large centers (serve 50 or more children) that were not always filled to capacity. Only six percent of the 3-year-old children attended centers affiliated with a school.

Teacher qualifications and training. About 30 percent of children in both cohorts had teachers with BA degrees, and another 30 percent had teachers with AA degrees. Approximately 40 percent of the children were with teachers who did not have either a BA or AA degree. There was also variation in teacher training and mentoring. Less than half the children in both cohorts had lead teachers who had received 25 hours of training in the last year or received mentoring at least once a month.

Classroom environment. The majority of children attended classrooms that had ECERS-R scores of at least a five (good), emphasized language and literacy and math activities, and met ratio standards. However, there was variation in the experiences of both cohorts. About 40 percent of the children in both cohorts were in classrooms that did not emphasize language and literacy or math activities (teachers reported providing less than half of a list of 12 language and literacy activities and eight math activities three times per week).

As noted, although quality was good in many areas, including environmental ratings and use of language and literacy and math activities, there was variation in the services and experiences available to children who attended Head Start for both cohorts. In reviewing the outcomes for children and families, the reader needs to consider this variation as well as the extent to which Head Start and control group children had differences in experiences as described below.

The Impact of Head Start on Children's Preschool Settings

As would be expected, providing access to Head Start led to a change in where children spent their time during the preschool years. Children assigned to the Head Start group were expected to enroll in the program. Families assigned to the control group, on the other hand,

could participate in other types of child care or preschool programs, but could also opt to keep their child at home or under the care of another adult. In some cases, the non-Head Start programs may have looked very much like Head Start, while others may have been quite different. How access to Head Start affected parental decisions regarding their child's preschool setting was the most important determinant of the program's impact as it set the stage for differences in the children's preschool experiences.

To examine children's preschool settings, we used information obtained from the parent interview, conducted each spring, to define a "focal setting" for each child in the study. The focal setting represented the child's care arrangement that accounted for at least five hours a week between 8 am and 6 pm Monday through Friday. Head Start was always defined as the focal setting for children enrolled in the program whether in the Head Start or control group. For children not in Head Start, but in multiple arrangements, each of which lasted at least five hours per week, the following hierarchy was used to prioritize and select the focal setting: (1) center-based program, (2) non-relative's home, (3) relative's home, (4) non-parental care in the child's own home by a non-relative, and (5) non-parental care in the child's own home by a relative. In the absence of non-parental care that met the time criteria, the child's focal setting was parent care. The seven types of focal settings are defined and presented in hierarchical order in Exhibit 3.4.

Exhibit 3.4: Definitions of Children's Focal Settings

1. *Head Start*: center-based, home-based, and combination programs funded with Federal Head Start dollars.
2. *Non-Head Start Center*: center-based program as differentiated from child care that takes place in someone's home or in federally funded Head Start classrooms.
3. *Non-Relative's Home*: non-parental care that takes place in a non-relative's home that is not the child's own home. This category includes regulated family child care providers as well as home-based child care providers who are exempt from licensing requirements.
4. *Relative's Home:* non-parental care that takes place in a relative's home that is not the child's own home. This category includes regulated family child care providers who are relatives of the child, as well as home-based relative care providers who are exempt from licensing requirements.
5. *Child's Own Home with a Non-Relative*: non-parental care that takes place in the child's own home by a non-relative of the child. Providers in this category generally are exempt from licensing requirements.
6. *Child's Own Home with a Relative*: non-parental care that takes place in the child's own home by a relative of the child. Providers in this category generally are exempt from licensing requirements.
7. *Parent Care*: care by the child's parent or guardian, typically in the child's own home.

Impact on Preschool Setting: 4-Year-Old Cohort

As shown in Exhibit 3.5, over three-quarters of children in the Head Start group were in Head Start during the spring of their preschool year in 2003. Among children in the control group, the largest proportion (almost 40%) was in parent care, followed by non-Head Start centers (about 35%). Since all parents applied to Head Start in the fall, we know that the parents of the study population were interested in having their children attend Head Start. Yet, when they did not have access to Head Start due to random assignment, two out of five control group families kept their children at home with a parent. About the same fraction of these families enrolled their child in a non-Head Start center-based program.

Exhibit 3.5: Percentage of Children by Focal Setting, 4-Year-Old Cohort, Head Start Year, Spring 2003

Focal Setting	Head Start Group	Control Group	Magnitude of Impact
Head Start	76.7	13.8	62.9***
Non-Head Start center	11.1	35.3	-24.2***
Home of non-relative	1.5	5.8	-4.3**
Home of relative	0.9	3.0	-2.1
Own home with relative	0.4	2.3	-1.9
Own home with non-relative	0.3	0.1	0.2
Parent care	9.1	39.7	-30.6***
Total	100%	100%	

*p≤0.10, **p≤0.05, ***p≤0.01

Not all children and families adhered to the status given to them through the random assignment process. Nearly one-quarter (23%) of Head Start group children were in non-Head Start settings or at home with their parents in spring 2003, although some of these children may have attended Head Start earlier in the year. Further, nearly 14 percent of control group children were enrolled in a Head Start Program.[61] These deviations from random assignment are not unexpected, as perfect conditions in an experimental design can rarely be achieved. During recruitment, Head Start programs reported it is typical for some families to enroll in a Head Start program but subsequently opt for a different care setting—sometimes another nearby Head Start

[61] Note that these children are not exactly the same as those defined as "no-shows and cross-overs" in Chapter 2, where children's settings are examined over the entire school year. Chapter 3 presents only the spring 2003 settings, in order to help interpret the other tables presented in this chapter.

program. Similarly, "crossovers," children who were assigned to the control group but attended Head Start, also could have applied to another nearby program. This is particularly likely to occur in densely populated areas with Head Start programs operating in proximity.

The parent interviews also obtained information on how long 4-year-olds assigned to the treatment group had been in Head Start by spring 2003. As shown in Exhibit 3.6, about 90 percent of the Head Start group children who were enrolled in Head Start attended the program for at least 8 months, indicating that most of these children had participated for the entire school year by the time the parents were interviewed.[62]

Exhibit 3.6: Percentage of Head Start Group Children Who Enrolled in Head Start by Months of Participation, 4-Year-Old Cohort, Head Start Year, Spring 2003

Months in Head Start as of Spring 2003	Percentage
≤4	0.0
5	0.1
6	0.9
7	9.0
8	47.5
9	35.5
10	7.0
Total	100%
Mean Length of Participation for Enrollees	8.43 months

Additionally, 4-year-olds who had access to Head Start attended the program, on average, 24 hours a week. Head Start group children, who did not attend Head Start but went instead to other center-based programs or family day care homes, attended for an average of 31 hours a week. It is possible that in the Head Start group, parents who chose not to send their children to Head Start may have done so because they needed more extensive day care coverage. Similarly, children in the control group who attended Head Start attended for an average of 25 hours a week and when they attended other center-based care, they attended for 29 hours a week.

[62] This exhibit presents information only on the Head Start group children who actually enrolled in Head Start.

Exhibit 3.7 presents the average per week that the 4-year-old cohort spent in non-parental child care settings. Non-parental care settings include, Head Start, other center-based care and home-based care.

Exhibit 3.7: Average Hours Per Week That the 4-Year-Old Cohort Spent in Non-Parental Child Care Settings, Head Start Year, Spring 2003

	Average Hours Per Week	
	Head Start Group	Control Group
Head Start	24	25
Other center-based care	31	29
Home-based care	37	35
Overall average across the three non-parental care settings	25	29

Impact on Preschool Setting: The Head Start Year, 3-Year-Old Cohort

Children in the 3-year-old cohort had two years in which to attend Head Start or child care before they entered kindergarten. As shown in Exhibit 3.8, like the 4-year-olds, the majority of 3-year-old Head Start group children (84.2%) were in Head Start during the first year (2002-03). Among children in the control group, the largest proportion was in parent care (38.4%), followed by non-Head Start centers (25.2%). Again, children did not always end up where expected; about 14 percent of the Head Start group children enrolled in non-Head Start settings or stayed home with their parents, and about 17 percent of the control group children managed to enroll in Head Start during their first year in the study.

Exhibit 3.8: Percentage of Children by Focal Setting, 3-Year-Old Cohort, Head Start Year, Spring 2003

Focal Setting	Head Start Group	Control Group	Magnitude of Impact
Head Start	84.2	17.8	66.4***
Non-Head Start center	6.7	25.2	-31.9***
Home of non-relative	0.8	6.8	-6.0***
Home of relative	1.0	8.0	-7.0***
Own home with relative	0.6	3.6	-3.0***
Own home with non-relative	0.0	0.2	-0.2
Parent care	6.7	38.4	-31.7***
Total	100%	100%	

*p≤0.10, **p≤0.05, ***p≤0.01

As shown in Exhibit 3.9, like the 4-year-old cohort, about 90 percent of the 3-year-old Head Start group children who enrolled in Head Start attended the program for at least 8 months.

Exhibit 3.9: Percentage of Head Start Group Children Who Enrolled in Head Start by Months of Participation in Head Start, 3-Year-Old Cohort, Head Start Year, Spring 2003

Months in Head Start as of Spring 2003	Percentage
≤3	0.0
4	0.3
5	0.4
6	0.7
7	8.4
8	48.8
9	36.8
10	4.6
Total	100%
Mean Length of Participation for Enrollees	8.34 months

There was variation in the number of hours per week spent in Head Start as compared to other non-parental care settings for the Head Start and control group children (Exhibit 3-10). Fewer hours were spent in Head Start than other setting options for both the Head Start and control groups. Because control children were more likely to be in non-parental settings that were not Head Start, on average they spent five more hours a week in non-parental care.

Exhibit 3.10: Average Hours Per Week That the 3-Year-Old Cohort Spent in Non-Parental Child Care Settings, Head Start Year, Spring 2003

	Average Hours Per Week	
	Head Start Group	Control Group
Head Start	27	27
Other center-based care	36	34
Home-based care	37	38
Overall average across the three non-parental care settings	28	33

Impact on Preschool Setting: The Age 4 Year, 3-Year-Old Cohort

The 3-year-old cohort's second year (2003-04) experiences were very different from those of their first year. As noted earlier, control group children were no longer prohibited from enrolling in Head Start during this second year, and indeed nearly half of them did. As a result, as shown in Exhibit 3.11, there were no longer any significant differences between the Head Start and control group children in terms of their enrollment in early childhood care and education environments in this second year. Moreover, similar percentages of children in the Head Start and control groups were in parent care, and the proportion of control group children in Head Start was not significantly different from the proportion of Head Start group children in Head Start in this year. Most of the change in the control group early care and education settings from the first to the second year of the study was a reduction in parent care and an increase in Head Start enrollment.

Exhibit 3.11: Percentage of Children by Focal Setting, 3-Year-Old Cohort, Age 4 Year, Spring 2004

Focal Setting	Head Start Group	Control Group	Magnitude of Impact
Head Start	63.4	49.6	13.8
Non-Head Start center	26.3	36.4	-10.1
Home of non-relative	0.9	2.5	-1.6
Home of relative	0.8	1.7	-0.9
Own home with relative	0.2	1.0	-0.8
Own home with non-relative	0.1	0.0	0.1
Parent care	7.0	7.6	-0.6
Missing/not ascertained	1.2	1.0	0.2
Total	100%	100%	

*p≤0.10, **p≤0.05, ***p≤0.01

Note: Due to rounding, the sum of the percents may not equal 100 percent.

As in the first year, those children who attended Head Start, attended for 8 months or more on average. The Head Start group averaged 8.4 months of Head Start, and the control group averaged 8.9 months of Head Start.

In their second year of Head Start, the Head Start and control group children spent the same amount of time in non-parental care. On average, both groups still spent less time when in Head Start than in other non-parental care settings (see Exhibit 3-12).

Exhibit 3.12: Average Hours Per Week That the 3-Year-Old Cohort Spent in Non-Parental Child Care Settings, Age 4 Year, Spring 2004

	Average Hours Per Week	
	Head Start Group	Control Group
Head Start	28	26
Other center-based care	32	31
Home-based care	37	39
Overall average across the three non-parental care settings	29	29

Characteristics of Early Childhood Settings

This section describes the early childhood caregiving and education environments of children during their preschool years; the next and final section of this chapter focuses on a description of the experiences of children in kindergarten and 1st grade. Both sections begin with a description of the experiences of children in the 4-year-old cohort (who had a single preschool or Head Start year) and then describe the experiences of children in the 3-year-old cohort (who typically had two preschool or Head Start years).

The main analysis of children's early education experiences is based on two tables (Exhibit 3.13 for the 4-year-old cohort and Exhibit 3.14 for the 3-year-old cohort) that present the data in two different ways: first, including data on all randomly assigned children[63] and second, including data on only those children who were in non-parental care settings. The two approaches are necessary because many measures of setting characteristics may not be relevant, may have different interpretations, or may have dramatically different levels based on whether children were in non-parental care outside of their home. Since children in the control group were much more likely than those in the Head Start group to remain at home in parental care during the year they first applied to Head Start, differences in average setting characteristics between the Head Start and control group may reflect both differences in the proportion who remained at home and differences in the characteristics of care settings for those in non-parental care (e.g., children at home cannot be categorized as having a teacher with a bachelor's degree

[63] For the analysis on all randomly assigned children, data on center characteristics are coded as 0's when children were in parent care or in their own homes with relatives or non-relatives. If a comparable family day care variable is not available, then children in family day care homes also get a 0 for that variable.

because they have no non-parental teacher). Consequently, we have opted to display the information on children's preschool experiences in two ways.

- **Impact.** The columns marked "impact" reflect data on all of the children randomly assigned regardless of their focal setting. These data represent the impact of having access to Head Start on the experiences of the entire cohort of children (noted by the column label, magnitude of impact).
- **Description.** The columns marked "description" include only those children who were in a non-parental setting (i.e., excluding children who were cared for in their own homes).[64] These comparisons show differences in the experiences of Head Start and control group children who were in a non-parental caregiving environment for at least five hours a week. Because these data represent non-random subsets of children in both study groups, the observed differences do not represent the impact of access to Head Start and have been provided only for descriptive purposes.

The first set of comparisons—the impact findings—are important for showing how providing access to Head Start affects children's care settings and environments. This information is critical to understanding the impacts on child and parenting outcomes presented in later chapters, since it highlights how the overall experience of the Head Start group differs from that of the control group on average. However, because this analysis includes all children in both groups—regardless of whether they are in any out-of-home care environment—it does not allow readers to understand how the caregiving environments look for those children receiving non-parental care outside the home. Providing access to Head Start dramatically affected whether children were in out-of-home care, as previously noted. Thus, it is of interest to know how the quality of care received by control group children in non-parental care compares to the quality of care received for children in the Head Start group. This latter type of information is gleaned from the information provided in the last two, shaded columns on Exhibit 3.13 and Exhibit 3.14.

For both age cohorts, the tables present findings for each of the constructs and measures described in Chapter 2. For each measure, the numbers shown in the respective table cells represent the percentage of children who were in a setting with the indicated experiential characteristic (i.e., the center DOES use a curriculum). Since the converse is not presented (i.e., the percentage in settings where a curriculum is NOT used), the columns do not total to 100

[64] Approximately one percent of the sample's focal setting was care by relatives or non-relatives in the child's own home. These children are excluded from the descriptive analysis.

percent.[65] In addition, the percentage of children being cared for by their parents or in home-based care (in their own homes by relatives or non-relatives), is shown separately in the first row. As mentioned, these children are included in the calculations under the "Impact" heading, but excluded from the calculations under the "Description" heading.

Characteristics of Early Childhood Settings: 4-Year-Old Cohort

As shown in the "Impact" columns of Exhibit 3.13, there are statistically significant differences between the Head Start and control group on every measure of children's preschool experiences, and in most cases the magnitude of the differences is quite large (e.g., 20-40 percentage points). These large differences are not unexpected because of the inclusion of children in parental care (most of whom were in the control group). This general shift of Head Start group children from parental care to non-parental care is what drives the size of the observed differences in children's preschool experiences.

Some of the key differences in preschool experiences between the Head Start program and control group are highlighted below:

- Having access to Head Start significantly increased the percentages of children who attended centers where structured curricula were used,[66] home visits were conducted, transportation was provided, and training and/or mentoring were available for teachers.
- Head Start group children were more likely than children in the control group to be in settings that offered a variety of child support services including health services, hearing and vision screening, mental health services, and nutrition.
- Children who had access to Head Start were more likely to have a teacher with a bachelor's degree than children in the control group.
- Head Start group children were more likely to have had a variety of language/literacy and math instructional activities at least three times a week than children in the control group.

---

[65] Complete tables are available on the Administration for Children and Families, Office of Planning, Research, and Evaluation Website.

[66] This largely reflects compliance with the Head Start Performance Standard mandating use of a curriculum, consistent with developmentally appropriate early childhood care and education. The Performance Standard does not specify which curriculum programs must use, but data collected from the Center Director interviews, as well as data from the Head Start Family and Child Experiences Survey (FACES), have shown that the large majority of Head Start programs use either Creative Curriculum or High/Scope (U.S. Department of Health and Human Services, 2005; Zill et al., 2003).

Exhibit 3.13: Percentage of Children by Their Settings' Characteristics: 4-Year-Old Cohort, Head Start Year, Spring 2003

Characteristic	Impact (All children in each group)			Description (Only children in non-parental care)	
	Head Start Group (% YES)	Control Group (% YES)	Magnitude of Impact	Head Start Group (% YES)	Control Group (% YES)
Parent or own home care	9.8	42.1	-32.3***		
Center Environment and Characteristics					
Qualifications: Director has at least a bachelor's degree	57.6	30.2	27.4***	66.4	61.7
Qualifications: Director has been in current position at least 4 years	43.8	29.5	14.3***	50.6	60.4
Training: Center provides teacher training at least monthly	61.6	26.5	35.1***	71.0	54.2
Training: Center provides teacher mentoring	75.8	33.4	42.2***	87.4	68.4
Turnover: Center has low proportion of new lead teachers (≤ 20%)	44.3	27.3	17.0***	51.1	55.8
Center size > 50	54.4	26.0	28.4***	62.8	53.1
Competition from other preschools:					
Lots	22.9	14.6	8.3***	26.4	29.8
Some	30.7	16.7	14.0***	35.4	34.2
Little or none	32.7	17.3	15.4***	37.7	35.5
Center always is filled to capacity	50.2	17.8	32.4***	57.9	36.5
Center is affiliated with a school	18.6	13.5	5.1***	21.4	27.6
Center uses curriculum	86.3	42.8	43.5***	99.5	87.6
Services Available for Children:					
Hearing/vision screening/ referrals	84.6	38.3	46.3***	94.0***	65.4
Mental health services	76.5	31.2	45.3***	85.0***	53.3
Health services	80.3	36.0	44.3***	89.2***	61.4
Nutrition services	82.5	40.2	42.3***	91.6***	68.6
Center provides transportation	52.4	18.8	33.6***	60.5***	38.4
Services Available for Families:					
Job training/employment assistance	67.4	22.0	45.4***	74.9***	37.4
Adult education/literacy	77.1	32.7	44.4***	85.6***	55.8
Family counseling or mental health services	71.0	27.3	43.7***	78.9***	46.6
Help with dealing with family violence	68.6	25.3	43.3***	76.2***	43.1
Help with housing	62.4	21.1	41.3***	69.4***	35.9
Help with medical care	63.7	24.7	39.0***	70.7***	42.1

Exhibit 3.13: Percentage of Children by Their Settings' Characteristics: 4-Year-Old Cohort, Head Start Year, Spring 2003 (continued)

Characteristic	Impact (All children in each group)			Description (Only children in non-parental care)	
	Head Start Group (% YES)	Control Group (% YES)	Magnitude of Impact	Head Start Group (% YES)	Control Group (% YES)
Services Available for Families: (cont'd)					
Family received home visit from focal setting	54.3	10.1	44.2***	60.1***	17.5
Alcohol/drug abuse treatment or counseling	55.9	18.1	37.8***	62.1***	30.8
Food and nutrition assistance	60.5	23.6	36.9***	67.2***	40.2
Income assistance	48.3	17.7	30.6***	53.7***	30.1
Foster care program	25.1	7.2	17.9***	27.9***	12.3
Other	8.2	5.4	2.8	9.1***	9.2
Teacher/Care Provider Qualifications and Training					
Had college ECE courses or obtained CDA	85.7	48.5	37.2***	95.1**	78.0
Obtained CDA (with/without college ECE courses)	47.9	21.1	26.8***	53.2**	33.9
Highest educational attainment was associate's degree	31.2	14.4	16.8***	34.6	23.2
Highest educational attainment was bachelor's degree or higher	30.8	19.7	11.1***	34.2	31.6
Attained bachelor's degree or higher in ECE *(subset of previous row)*	28.6	15.9	12.7***	31.7	25.6
Received at least 25 hours of training in past year	40.5	17.4	23.1***	44.9	28.0
Received mentoring at least once/ month	42.3	20.8	21.5***	47.0	33.4
Classroom Environment					
Had average ECERS-R/FDCRS rating of at least 5 out of 7	71.5	38.3	33.2***	79.9***	64.3
Had average ECERS-R/FDCRS rating of at least 6 out of 7	40.2	15.8	24.4***	44.9***	26.6
Had highest average ECERS-R/ FDCRS rating (7)	4.0	1.4	2.6***	4.5	2.4
Had average Arnett rating of at least 3 out of 4	83.9	52.9	31.0***	93.9	88.8
Had highest average Arnett rating (4)	3.5	1.4	2.1***	4.0	2.4
Met child/staff ratio standard	64.7	31.5	33.2***	72.4**	52.9
Parent participated in setting activity at least once	79.7	44.8	34.9***	88.2	76.7

Exhibit 3.13: Percentage of Children by Their Settings' Characteristics: 4-Year-Old Cohort, Head Start Year, Spring 2003 (continued)

Characteristic	Impact (All children in each group)			Description (Only children in non-parental care)	
	Head Start Group (% YES)	Control Group (% YES)	Magnitude of Impact	Head Start Group (% YES)	Control Group (% YES)
Classroom Activities					
Provides at least 3 of 4 activities other than literacy and math at least 3 times/week	84.9	49.1	35.8***	94.1***	78.8
Provides at least 7 of 12 literacy activities at least 3 times/week	61.5	28.1	33.4***	68.2***	45.1
Provides at least 5 of 8 math activities at least 3 times/week	60.7	29.3	31.4***	67.4***	47.2
Overall Quality Composite					
At or above overall mean	66.6	21.4	45.2***	74.5***	34.9

*p≤0.10, **p≤0.05, ***p≤0.01

- Children in the Head Start group were more likely to be in higher quality settings than children in the control group, and this finding is consistent across both the standardized ECERS-R/FDCRS and Arnett ratings and using the "quality composite" created for this study.[67]
- Nearly two-thirds of the children in the Head Start group were in settings that achieved the AAP/APHA ratio standard, compared to less than one-third of children in the control group. Providing access to Head Start increased the likelihood that children were in classrooms or family care settings that met the standards.

Comparing only those children in the Head Start and control groups who were in non-parental care (the last two columns of Exhibit 3.13), we see the same pattern of differences. Like the differences described above, these comparisons suggest that Head Start group children are in higher quality care settings than control group children. However, the magnitudes of the differences are, in most cases, smaller than when all children were included in the analysis. In a few instances, these differences were in the opposite direction when children exclusively in parental care were left out of the analysis.

[67] Readers interested in more details about each of the subscale findings for the ECERS-R, FDCRS, and Arnett are referred to Exhibits D.1 through D.3 in Appendix D of this volume, which presents details of subscale scores for both age cohorts. Findings are consistent with previous studies that found that Head Start classrooms rarely score in the ECERS-R/FDCRS minimal range and tend to have higher ratings than other types of child care centers and preschools (Zill et al., 2003).

The statistically significant differences in preschool experiences between the Head Start and control groups of this second analysis are highlighted below. However, it is important to note that these analyses do not represent the impact of access to Head Start; rather, they are purely descriptive:

- Children in the Head Start group were more likely to be in centers where services were available to children and families.
- Children in the Head Start group were more likely to be with teachers who had obtained a Child Development Associate Credential (CDA) or had college courses in early childhood education.
- Children in the Head Start group were more likely to be in classrooms with ECERS-R scores of five or higher, that met the child/staff ratio standard, and provided more math and literacy activities.

Characteristics of Early Childhood Settings: 3-Year-Old Cohort

The chapter now turns to the experiences of the 3-year-old cohort, who had two potential years of Head Start (2002-03 and 2003-04). As discussed earlier, the impacts estimated for the 3-year-old cohort represent the effects of access to Head Start two years prior to entering kindergarten. The children randomized into the control group (and therefore embargoed from Head Start the first preschool year) were allowed to enroll in Head Start the following year.

The Head Start Year: 3-Year-Old Cohort

As shown in Exhibit 3.14, there are statistically significant differences between the Head Start and control group on every measure of children's preschool experiences, and in most cases, the magnitude of the differences is quite large (e.g., 20-50 percentage points). As stated previously, these large differences were not unexpected because of the dramatic difference between the Head Start and control groups in the percentages of children in exclusively parental care. Like the 4-year-old cohort, about 40 percent of the 3-year-old control group stayed home with parents rather than attended a preschool program during this Head Start year. With slight changes in percentages and means, the main findings for the 3-year-old cohort mirror those for the 4-year-old cohort.

- Having access to Head Start significantly increased the percentages of children who attended centers where structured curricula were used, home visits were conducted, transportation was provided, and training and/or mentoring were available for teachers.

Exhibit 3.14: Percentage of Children by Their Settings' Characteristics: 3-Year-Old Cohort, Head Start Year, Spring 2003

Characteristic	Impact (All children in each group)			Description (Only children in non-parental care)	
	Head Start Group (% YES)	Control Group (% YES)	Magnitude of Impact	Head Start Group (% YES)	Control Group (% YES)
Parent or own home care	7.3	42.2	-34.9***		
Center Environment and Characteristics					
Qualifications: Director has at least a bachelor's degree	62.3	23.1	39.2***	68.6	52.5
Qualifications: Director has been in current position at least 4 years	53.5	26.9	26.6***	58.9	61.0
Training: Center provides teacher training at least monthly	56.6	24.4	32.1***	62.1	55.4
Training: Center provides teacher mentoring	78.1	36.6	41.8***	85.9	82.4
Turnover: Center has low proportion of new lead teachers (≤ 20%)	44.0	24.4	19.6***	48.4	55.5
Center size > 50	62.8	26.8	36.0***	69.0	60.9
Competition from other preschools:					
Lots	33.5	18.5	15.0***	36.8	41.9
Some	32.4	15.6	16.8***	35.6	35.5
Little or none	24.8	9.8	15.0***	27.3	22.3
Center always is filled to capacity	37.4	15.5	21.9***	41.1	35.3
Center is affiliated with a school	6.2	6.1	0.1***	6.9	13.8
Center uses curriculum	90.0	41.6	48.4***	99.0	94.6
Services Available for Children:					
Hearing/vision screening/referrals	85.8	37.1	48.7***	91.9***	61.1
Mental health services	79.6	27.3	52.3***	85.3***	44.9
Health services	83.3	33.8	49.5***	89.3***	55.8
Nutrition services	85.6	44.6	41.0***	91.7***	73.6
Center provides transportation	56.9	16.8	40.1***	62.6***	38.1
Services Available for Families:					
Job training/employment assistance	64.8	26.6	38.2***	69.4***	43.8
Adult education/literacy	75.9	30.2	45.7***	81.3***	49.8
Family counseling or mental health services	69.6	27.5	42.1***	74.5***	45.2
Help with dealing with family violence	62.6	27.3	35.3***	67.0***	45.0
Help with housing	57.0	23.2	33.8***	61.0**	38.3
Help with utilities	55.8	23.1	32.7***	59.8***	38.0

Exhibit 3.14: Percentage of Children by Their Settings' Characteristics: 3-Year-Old Cohort, Head Start Year, Spring 2003 (continued)

Characteristic	Impact (All children in each group)			Description (Only children in non-parental care)	
	Head Start Group (% YES)	Control Group (% YES)	Magnitude of Impact	Head Start Group (% YES)	Control Group (% YES)
Services Available for Families: (cont'd)					
Help with medical care	62.4	24.4	38.0***	66.8***	40.2
Family received home visit from focal setting	53.2	11.3	41.9***	57.3***	19.5
Alcohol/drug abuse treatment or counseling	51.0	20.2	30.8***	54.6***	33.2
Food and nutrition assistance	56.2	26.0	30.2***	60.3***	42.8
Income assistance	43.3	22.1	21.2***	46.4	36.5
Foster care program	20.0	7.3	12.7***	21.5**	12.0
Other	7.2	3.3	3.9***	7.7	5.5
Teacher/Care Provider Qualifications and Training					
Had college ECE courses or obtained CDA	87.7	42.4	45.3***	94.0***	69.3
Obtained CDA (with/without college ECE courses)	54.2	20.1	34.1***	58.1***	33.0
Highest educational attainment was associate's degree	29.7	10.1	19.3***	31.9***	16.5
Highest educational attainment was bachelor's degree or higher	30.1	15.4	14.7***	32.3	25.2
Attained bachelor's degree or higher in ECE *(subset of previous row)*	24.5	11.3	13.2***	26.2	18.4
Received at least 25 hours of training in past year	33.2	12.0	21.2***	35.6	19.7
Receives mentoring at least once/ month	39.8	21.2	18.6***	42.6	34.7
Classroom Environment					
Had average ECERS-R/FDCRS rating of at least 5 out of 7	69.0	31.9	37.1***	74.4***	56.3
Had average ECERS-R/FDCRS rating of at least 6 out of 7	39.5	13.7	25.8***	42.7***	24.1
Had highest average ECERS-R/ FDCRS rating (7)	5.4	1.4	4.0***	5.8	2.4
Had average Arnett rating of at least 3 out of 4	83.3	49.1	34.2***	89.9	86.4
Had highest average Arnett rating (4)	3.2	0.5	2.7***	3.4	0.9
Met child/staff ratio standard	58.2	28.6	29.6***	62.8*	50.4
Parent participated in setting activity at least once	79.3	38.8	40.5***	85.6*	66.7

Exhibit 3.14: Percentage of Children by Their Settings' Characteristic: 3-Year-Old Cohort, Head Start Year, Spring 2003 (continued)

Characteristic	Impact (All children in each group)			Description (Only children in non-parental care)	
	Head Start Group (% YES)	Control Group (% YES)	Magnitude of Impact	Head Start Group (% YES)	Control Group (% YES)
Classroom Activities					
Provides at least 3 of 4 activities other than literacy and math at least 3 times/week	85.4	40.9	44.5***	91.5***	66.9
Provides at least 7 of 12 literacy activities at least 3 times/week	55.7	22.9	32.8***	59.7***	37.4
Provides at least 5 of 8 math activities at least 3 times/week	66.5	29.9	36.6***	71.2***	48.9
Overall Quality Composite					
At or above overall mean	70.8	20.3	50.5***	76.0***	34.7

*p≤0.10, **p≤0.05, ***p≤0.01

The columns marked "description" include only those children who were in a non-parental setting (i.e., excluding children who were cared for in their own homes).[68] These comparisons show differences in the experiences of Head Start and control group children who were in a non-parental caregiving environment for at least 5 hours a week. Because these data represent non-random subsets of children in both study groups, the observed differences do not represent the impact of access to Head Start and have been provided only for descriptive purposes.

- Head Start children were more likely than control group children to be in settings that offered a variety of services, including health services, hearing and vision screening, mental health services, and nutrition. The magnitude of differences in service availability was as large as 52.3 percentage points for mental health services.
- Children who had access to Head Start were more likely to have a teacher who had a bachelor's degree, had college-level early childhood education (ECE) course(s) or had obtained a CDA.
- Head Start children were more likely to have had a variety of language/literacy and math instructional activities at least three times a week than children in the control group.
- Children in the Head Start group were more likely to be in higher quality settings than children in the control group, and this finding is consistent across both the standardized ECERS-R/FDCRS and Arnett ratings and using the "quality composite" created for this study.[69]

[68] Approximately one percent of the sample's focal setting was care by relatives or non-relatives in the child's own home. These children are excluded from the descriptive analysis.

[69] Readers interested in more details about each of the sub-scale findings for the ECER-R, FDRCS, and Arnett are referred to Exhibits D.4 through D.9 in Appendix D of this volume, which presents details of subscale scores.

- Approximately, 58 percent of the children in the Head Start group were in settings that achieved the AAP/APHA ratio standard, compared to nearly 30 percent of children in the control group. Access to Head Start had an impact on children being in classrooms that met the standards.

As with the 4-year-old cohort, when focusing only on children in non-parental settings the same pattern of differences favoring the Head Start group is found. In general, statistically significant experiences include:

- Children in the Head Start group were more likely to be in centers where services were available to children and families.
- Children in the Head Start group were more likely to be with teachers who had obtained a CDA or had college courses in early childhood education. They were also more likely to have a teacher whose highest educational attainment was an associate's degree.
- Children in the Head Start group were more likely to be in classrooms with ECERS-R scores of five or higher and provided more math and literacy activities. There was also suggestive evidence that children in the Head Start group were more likely to be in classrooms that met the child/staff ratio standard and have parents who were more involved in classroom activities.

The Age 4 Year: 3-Year-Old Cohort

The Head Start experiences of children in the 3-year-old cohort were very different in the age 4 year. As discussed earlier, most of these children were in some type of center-based care by the second year (90 percent for children in the Head Start group and 86 percent for children in the control group), and in fact, there was much less difference between the Head Start and control group children on measures of the characteristics of care. Control group children were allowed to go to Head Start for their second preschool year, and approximately 50 percent did so. This rate of Head Start enrollment is not significantly different from the approximately 60 percent rate for the Head Start group. Consequently, measures of the quality and structure of care received by the Head Start and control groups also show fewer differences in this second year. As shown in Exhibit 3.15, there were only three statistically significant differences between the Head Start and control groups.

- Children in the control group (32.1%) were more likely to be in a center that was affiliated with a school than children in the Head Start group (25.6%).
- Children in the Head Start group were more likely to have a teacher with a CDA: 40.2 percent of the Head Start group vs. 27.5 percent of the control group had a teacher with a CDA.

Exhibit 3.15: Percentage of Children by Their Settings' Characteristics: 3-Year-Old Cohort, Age 4 Year, Spring 2004

Characteristic	Impact (All children in each group)			Description (Only children in non-parental care)	
	Head Start Group (% YES)	Control Group (% YES)	Magnitude of Impact	Head Start Group (% YES)	Control Group (% YES)
Parent care or own home	7.3	8.6	-1.3		
Center Environment and Characteristics					
Qualifications: Director has at least bachelor's degree	65.8	63.1	2.7	72.8	72.2
Qualifications: Director has been in current position at least 4 years	58.6	54.1	4.5	64.9	61.9
Training: Center provides teacher training at least monthly	64.1	56.8	7.3	71.0	65.0
Turnover: Center has low proportion of new lead teachers ($\leq$ 20%)	23.0	23.3	-0.3	25.5	26.7
Center size > 50	30.7	31.2	-0.5	34.0	35.7
Competition from other preschools:					
Lots	27.9	22.5	5.4	30.9	25.7
Some	33.5	38.9	-5.4	37.1	44.5
Little or none	27.9	25.5	2.4	30.9	29.1
Center always is filled to capacity	39.5	42.2	-2.7	43.7	48.3
Center is affiliated with a school	25.6	32.1	-6.5*	28.3	36.7
Center uses curriculum	88.6	86.8	1.8	98.2	99.3
Center provides teacher mentoring	74.9	70.8	4.1	83.0	81.0
Services Available for Children:					
Hearing/vision screening/referrals	86.8	80.6	6.2*	93.1	87.5
Mental health services	75.6	67.2	8.4	81.0	73.0
Health services	76.9	69.1	7.8	82.4	75.2
Nutrition services	83.9	78.9	5.0	90.0	85.8
Center provides transportation	56.2	51.6	4.6	62.2	59.0
Services Available for Families:					
Job training/employment assistance	52.2	42.3	9.9	55.9	46.0
Adult education/literacy	69.8	60.9	8.9	74.9	66.2
Family counseling or mental health services	60.6	54.9	5.7	65.0	59.7
Help dealing with family violence	60.6	54.2	6.4	65.0	58.9
Help with housing	42.7	38.4	4.3	45.8	41.7
Help with utilities	40.8	37.2	3.6	43.8	40.4
Help with medical care	53.7	45.6	8.1	57.6	49.6

Exhibit 3.15: Percentage of Children by Their Settings' Characteristics: 3-Year-Old Cohort, Age 4 Year, Spring 2004 (continued)

Characteristic	Impact (All children in each group)			Description (Only children in non-parental care)	
	Head Start Group (% YES)	Control Group (% YES)	Magnitude of Impact	Head Start Group (% YES)	Control Group (% YES)
Services Available for Families: (cont'd)					
Alcohol/drug abuse treatment or counseling	43.3	42.5	0.8	46.4	46.2
Food and nutrition assistance	44.4	40.3	4.1	47.6	43.9
Income assistance	33.8	31.7	2.1	36.2	34.5
Foster care payments	17.2	15.6	1.6	18.5	16.9
Other	12.3	11.9	0.4	13.2	13.0
Teacher/Care Provider Qualifications and Training					
Had college ECE courses or obtained CDA	87.6	84.3	3.3	93.7	91.5
Obtained CDA (with or without college ECE courses)	40.2	27.5	12.7**	43.0	29.8
Received at least 25 hours of training in past year	25.6	24.7	0.9	27.4	26.8
Receives mentoring at least once/ month	37.3	34.0	3.3	39.9	36.9
Highest educational attainment was bachelor's degree or higher	18.0	19.7	-1.7	19.3	21.4
Attained bachelor's degree or higher in ECE *(subset of previous row)*	16.1	18.0	-1.9	17.3	19.5
Classroom Environment					
Had average ECERS-R/FDCRS rating of at least 5 out of 7	63.4	60.9	2.5	69.0	66.9
Had average ECERS-R/FDCRS rating of at least 6 out of 7	28.3	28.5	-0.2	30.7	31.3
Had highest average ECERS-R/ FDCRS rating (7)	4.4	3.3	1.1	4.8	3.6
Had average Arnett rating of at least 3 out of 4	83.2	82.5	0.7	90.5	90.5
Had highest average Arnett rating (4)	2.7	2.7	0.0	2.9	2.9
Met child/staff ratio standard	43.1	36.4	6.7	46.9	40.0
Parent participated in setting activity at least once	82.2	82.6	-0.4	88.6	89.4

Exhibit 3.15: Percentage of Children by Their Settings' Characteristics: 3-Year-Old Cohort, Age 4 Year, Spring 2004 (continued)

Characteristic	Impact (All children in each group)			Description (Only children in non-parental care)	
	Head Start Group (% YES)	Control Group (% YES)	Magnitude of Impact	Head Start Group (% YES)	Control Group (% YES)
Classroom Activities					
Provides at least 3 of 4 activities other than literacy and math at least 3 times/week	84.8	83.7	1.1	90.7	90.8
Provides at least 5 of 8 math activities at least 3 times/week	65.7	62.1	3.6	70.3	67.4
Provides at least 7 of 12 literacy activities at least 3 times/week	65.6	64.9	0.7	70.2	70.4
Overall Quality Composite:					
At or above overall mean	49.9	52.1	-2.2	53.9	57.1

*p≤0.10, **p≤0.05, ***p≤0.01

- Children in the Head Start group were more likely to have hearing and vision screening referral services available (86.8%) compared to control group children (80.6%).

Comparing only those children in the Head Start and control groups who were in non-parental care, in the last two columns of Exhibit 3.15, we see that there was no statistically significant difference in the characteristics of non-parental settings for children in the Head Start group and the control group in the age 4 year.

Characteristics Related to Attending One or Two Years of Head Start

As presented in Chapter 2, the interpretation of the estimated impacts on the 4-year-old cohort is very straightforward. Impacts observed at the end of Head Start (spring 2003) represent the effect of access to Head Start when the children were in preschool prior to entry into elementary school. The two subsequent waves, spring 2004 and spring 2005, can be thought of as the consequences at the end of kindergarten and 1st grade of any benefits provided by Head Start. In other words, the later impacts represent the subsequent effect of the Head Start intervention received the year before the children entered school.

The situation is quite a bit different for the 3-year-old cohort because of the nature of the study design. As noted earlier, the children who were new applicants to Head Start when they were three (fall 2002) were randomized into two groups. One group was allowed entry into Head Start right away (the Head Start group), while the other group was not allowed to enter Head Start at age three (the control group). Consequently, impacts estimated at the end of that first year (spring 2003), represent the effect of access to Head Start when the children were in preschool two years prior to entry into elementary school. So far, this is similar to the interpretation of first year impacts for the 4-year-old children, but represents the effect when services are provided a year earlier.

Yet, after receiving one year of Head Start access (or being denied access that first year), the 3-year-old cohort had another year to go before they started kindergarten. For reasons of feasibility,[70] children in both the randomly assigned Head Start group and the randomly assigned control group were allowed to re-enroll in Head Start the following year when the children turned four (fall 2003). Not all of those children who were assigned to have access to Head Start in the first year went on for a second year at age four. Only about 60 percent of the Head Start group participated during two years.[71] Of the children randomly assigned to the control group, about 46 percent participated in Head Start for the first time when they turned four, and about 11 percent enrolled at that age for a second year (having found their way into Head Start in the first year despite the intended embargo).

What factors were related to families' decisions to remain with Head Start? Were there child or family characteristics driving the decision, or was the decision related more to the actual experiences that families had during their Head Start year (e.g., classroom characteristics or parental satisfaction)? Information on these characteristics is derived from parent reports and classroom observations. This information can be useful for increasing Head Start's ability to retain these families.

Ideally, this study would also address the question of how participation in Head Start for two years or one year affects children's outcomes. The HSIS study design does not allow for

---

[70] Primarily, Head Start programs were unwilling to embargo children assigned to the control group for two years.

[71] Since some of the children in the Head Start group never attended Head Start in the first year, this number is lower than the percentage of children who "return" for a second year. Of those Head Start group children who actually attended during the Head Start year, 72 percent returned for a second year.

differences in impacts for children who receive one year or two years of Head Start to be directly assessed, since children were not randomly assigned to these conditions.

Measurement

A variety of measures were chosen to explore what might predict parents' decisions to keep children in Head Start for a second year. Four broad categories have been analyzed: (1) child characteristics, (2) parent/household characteristics, (3) parental satisfaction with child's early Head Start year, and (4) center/program characteristics. Measures were selected to capture an array of issues that could potentially predict a parent's decision to continue a child in Head Start or make a change after one year. For example, were there parental concerns about a child's academic needs (e.g., child has an individualized education program (IEP), in lowest academic quartile) that are related to a child being in one or two years of Head Start? Were there parental characteristics that may be related to choices or opportunities (e.g., race/ethnicity, mother's education, household income)?

Parents' satisfaction with the services children received (e.g., safety, respect for cultural differences, amount of individual attention) is also considered. Finally classroom environment characteristics such as program quality indicators are also explored (e.g., were children in centers with poor ECERS-R scores less likely to return to Head Start for a second year?) Also considered were circumstances that go across categories. For example, what might be related to the availability of other child care options (e.g., either other family members to care for the child or competition from other options in the community).

Note that the analyses do not address causality. We do not attempt to distinguish whether the measures presented cause children to return only whether they are associated with the rate of re-enrollment.

The descriptive findings are presented in four tables, one for each of the four broad categories (Exhibits 3.16 through 3-19). Each table presents individual measures in the related category and depicts the percentage of children who attended only the early year of Head Start as compared to those who attended two years.

Exhibit 3.16: Child Characteristics by Percent of 3-Year-Old Head Start Group Children Who Participated in an Early Year of Head Start or 2 Years of Head Start

Child Characteristics	Percent of Children		
	Early Year of Head Start	2 Years of Head Start	p-value
Child's race/ethnicity			0.0727*
White	29.49	31.85	
Black	46.45	30.74	
Hispanic	24.06	37.41	
Child's gender			0.3418
Boy	43.43	48.18	
Girl	56.57	51.82	
Child has an IEP			0.7724
No	92.62	91.97	
Yes	7.38	8.03	
Change in child's IEP			0.5818
No	94.52	93.15	
Yes	5.48	6.85	
Child had low academic ability at baseline			0.7313
No	63.48	61.71	
Yes	36.52	38.29	

*p≤0.10, **p≤0.05, ***p≤0.01

Exhibit 3.17: Parent and Household Characteristics by Percent of 3-Year-Old Head Start Group Children Who Participated in an Early Year of Head Start or 2 Years of Head Start

Parent/Household Characteristics	Percent of Children		
	Early Year of Head Start	2 Years of Head Start	p-value
Biological father lives in household			0.0783*
No	57.67	48.54	
Yes	42.33	51.46	
Grandparent lives in the household			0.6771
No	96.08	96.79	
Yes	3.92	3.21	
Number of adults 18 and over in the household			0.9005
1	26.52	25.36	
2	55.30	54.36	
3 or more	18.18	20.27	
Other children under age 6 in the household			0.7541
Yes	48.42	47.03	
No	51.58	52.97	
Home language			0.0387**
Not English	17.53	27.42	
English	82.47	72.58	
Number of moves in the past 12 month			0.8887
No moves	67.48	70.21	
1 or more moves	32.52	29.79	
Family monthly income range			0.2337
Less than $250	3.15	3.10	
$251 - $500	11.43	8.70	
$501 - $1000	19.21	28.21	
$1001 - $1500	28.64	21.57	
$1501 - $2000	16.39	15.85	
$2001 - $2500	10.42	12.00	
Over $ 2500	10.76	10.57	
Economic difficulty			0.0805*
No	75.32	69.28	
Yes	24.68	30.72	
Father's employment status			0.6853
Full time	77.42	81.87	
Part time	7.94	6.30	
Not working	14.64	11.82	
Mother's employment status			0.7487
Full time	36.80	33.35	
Part time	14.95	14.94	
Not working	48.26	51.71	
Biological mother recent immigration status			0.0318**
No	89.54	82.55	
Yes	10.46	17.45	
Mother's age , fall 2002			0.2410
Less than 20 years old	4.80	1.90	
20-30 years old	61.67	67.49	
31-40 years old	29.16	26.87	
41 and over years old	4.37	3.75	

Exhibit 3.17: Parent and Household Characteristics by Percent of 3-Year-Old Head Start Group Children Who Participated in an Early Year of Head Start or 2 Years of Head Start (continued)

Parent/Household Characteristics	Percent of Children		
	Early Year of Head Start	2 Years of Head Start	p-value
Mother teenager at birth of study child			0.9342
No	87.72	87.93	
Yes	12.28	12.07	
Mother's marital status			0.7268
Never married	44.19	40.85	
Married	41.91	46.02	
Separated/divorced/widowed	13.90	13.13	
Mother's highest level of education attained			0.6461
Less than High School	27.63	31.40	
High School diploma or GED	37.88	35.96	
Beyond high school	34.48	32.64	
Mother reported depressive symptoms			0.7989
None	51.03	56.18	
Mild	27.25	23.90	
Moderate	10.58	9.99	
Severe	11.14	9.93	

*p≤0.10, **p≤0.05, ***p≤0.01

Note: Due to rounding, the sum of the percents may not equal 100 percent.

Exhibit 3.18: Parent Satisfaction and Involvement with Head Start Center by Percent of 3-Year-Old Head Start Group Children Who Participated in an Early Year of Head Start or 2 Years of Head Start

Satisfaction and Involvement with Center	Percent of Children		
	Early Year of Head Start	2 Years of Head Start	p-value
Center respect of family's culture			0.0347**
Very dissatisfied	1.59	0.36	
Somewhat dissatisfied	0.41	1.08	
Somewhat satisfied	13.54	6.16	
Very satisfied	83.80	92.17	
Center helped child grow & develop			0.0977*
Very dissatisfied	2.40	0.22	
Somewhat dissatisfied	2.80	0.58	
Somewhat satisfied	13.53	10.47	
Very satisfied	81.11	88.72	
Center was open to parent's ideas and participation			0.1689
Very dissatisfied	1.83	0.51	
Somewhat dissatisfied	0.51	0.76	
Somewhat satisfied	23.02	14.60	
Very satisfied	74.17	83.89	
Stability in child care relationships			0.1161
Never	0.24	0.91	
Sometimes	6.16	3.54	
Often	7.31	11.66	
Always	86.29	83.88	
Too much turnover in care providers			0.0497**
Never	86.91	86.35	
Sometimes	10.30	5.79	
Often	1.33	2.72	
Always	1.46	5.13	
Child feels safe/secure in care			0.5226
Never	0.00	0.16	
Sometimes	4.89	3.81	
Often	5.52	6.11	
Always	89.43	89.65	
Child receives individual attention			0.4352
Never	2.36	0.51	
Sometimes	14.06	13.79	
Often	19.08	22.76	
Always	64.00	61.64	
Teacher is open to new information and learning			0.9610
Never	0.10	0.06	
Sometimes	2.88	3.27	
Often	8.58	9.16	
Always	88.16	85.60	
Child in familiar place			0.1743
Never	1.79	2.10	
Sometimes	3.90	6.60	
Often	11.92	12.31	
Always	82.39	78.99	

Exhibit 3.18: Parent Satisfaction and Involvement with Head Start Center by Percent of 3-Year-Old Head Start Group Children Who Participated in an Early Year of Head Start or 2 Years of Head Start (continued)

Satisfaction and Involvement with Center	Percent of Children		
	Early Year of Head Start	2 Years of Head Start	p-value
Number of activities parent has done at least once with child care setting			0.2143
Has never done any of the activities	16.78	14.45	
Has done one of the activities	12.68	15.96	
Has done two of the activities	19.03	20.04	
Has done three of the activities	30.41	23.43	
Has done four of the activities	20.98	24.80	
Has done all five of the activities	0.11	1.31	

*p≤0.10, **p≤0.05, ***p≤0.01

Note: Due to rounding, the sum of the percents may not equal 100 percent.

Exhibit 3.19: Classroom Characteristics by Percent of 3-Year-Old Head Start Group Children Who Participated in an Early Year of Head Start or 2 Years of Head Start

Classroom Characteristics	Percent of Children		
	Early Year of Head Start	2 Years of Head Start	p-value
Classroom in Census designated urbanized area			0.6194
Rural	22.76	19.99	
Urban	77.24	80.01	
ECERS-R mean total score			0.3766
1-3	7.02	5.26	
4-7	92.98	94.74	
Arnett lead teacher mean score			0.55436
1-2	39.33	35.15	
3	60.67	64.85	
Number of classroom literacy activities at least 3 times a week			0.1302
0-5	20.30	30.43	
6-8	28.15	26.30	
9-12	51.55	43.27	
Number of classroom math activities at least 3 times a week			0.3389
0-3	15.61	12.30	
4-5	21.64	27.50	
6-8	62.75	60.20	
Comprehensive services provided by center			0.6649
0-8	27.85	33.24	
9-14	53.99	48.79	
15 or more	18.16	17.97	
Meets child/staff ratio standard			0.7015
No	35.37	37.89	
Yes	64.63	62.11	
Teacher education level			0.9918
Less than a B.A.	69.02	69.91	
B.A. or higher	30.98	30.09	
Overall quality score—At mean or above			0.2387
No	30.92	39.07	
Yes	69.08	60.93	
Competition from other pre-schools/pre-k			0.0106***
No competition	58.54	69.41	
Competition	41.46	30.59	
Center affiliation			0.2293
School	5.56	8.24	
Not school	94.44	91.76	
Center provides part-day, full-day, or both types of programs			0.0126***
Part-day only	45.41	24.03	
Full-day only	28.20	45.31	
Both	26.05	28.37	

*p≤0.10, **p≤0.05, ***p≤0.01

Note: Due to rounding, the sum of the percents may not equal 100 percent.

The sample for this analysis includes only those Head Start group children in the 3-year-old cohort who attended Head Start the first year (i.e., in the 2002-03 school year), which includes 1,083 children (309 with an early year only and 774 with two years of Head Start). Weighted percents are provided.[72]

Findings

The likelihood that children return for a second year of Head Start was significantly related to both the services provided by the program and the availability of other options in the community. There is clear evidence that when there is less child care competition in the area and local Head Start centers offer full-day services, families are more likely to return their children to Head Start for a second year. Likewise, a few key elements of parents' satisfaction with the program were related to retention. Parents who were more satisfied with the center's sensitivity to cultural issues were more likely to return children to Head Start for a second year than those in the other satisfaction categories. Further, parents who were more satisfied with how the center helped their child to grow and develop were more likely to return the child for a second year.

There was also an association between racial, ethnic, and immigration characteristics and the likelihood that children returned for a second year. There is suggestive evidence that children from Hispanic families are more likely than children from Black or White families to return for a second year of Head Start and that children from Black families were less likely to return for a second year than children in the other two race/ethnicity groups. Likewise, children from families in which mothers were recent immigrants, and Spanish was the household language, were significantly more likely to return for a second year of Head Start.

Few of the household characteristics were related to the child's returning for a second year. Households in which the study child's biological father was living as well as households reporting economic difficulties within the last three months of the interview were more likely to return the study child to a second year of Head Start.

Finally, while most parents reported there was not too much turnover in their children's care settings (86 percent reported that there was never too much turnover in care providers),

[72] Weighting is discussed in detail in the Technical Report for the Head Start Impact Study.

those who reported that there was "sometimes" too much turnover were less likely to return for a second year.

Children's Experiences in Kindergarten and 1st Grade

The 4-Year-Old Cohort: Overall School Characteristics

Head Start aims to make parents better advocates for their children and to support them in their children's transition to elementary school. Thus, it was hypothesized that access to Head Start could change the type of schools that children attended and the nature of the experiences they had in kindergarten and 1st grade (recognizing the limited options often available to low-income families in 2004 through 2006). However, as shown in Exhibit 3.20, there was no statistically significant impact on the 4-year-old cohort's attendance at public, charter, or private schools for either kindergarten or 1st grade. About 80 percent of the children attended traditional public or charter schools for kindergarten, and about 90 percent did so for 1st grade.

Not surprisingly, the 4-year-old cohort attended schools with much higher levels of poverty than the average nationwide (as indicated by the proportions of students eligible for free and reduced-price meals). Further, these children were more likely to go to schools with higher proportions of Black and Hispanic students than the average nationwide. Nationally, approximately 43 percent of students are eligible for free and reduced-price meals as compared to about 66 percent of students at the schools the study children attended. The study children attended schools in which about 60 percent of the enrollment was Hispanic and Black children, compared to about 40 percent nationally.[73]

[73] This represents 133,910 schools and excludes schools without K and 1st grade for the 2003-04 school year.

Exhibit 3.20: School Type Characteristics for Schools Attended by Treatment and Control Groups by Year, 4-Year-Old Cohort

Characteristics of Children	Average Percent of Students in Schools Attended By: Head Start Group (%)	Average Percent of Students in Schools Attended By: Control Group (%)	Magnitude of Impact (%)
2004 (Kindergarten)			
School type			
Public	80.0	77.8	2.2
Charter	1.2	1.2	0.0
Private	2.2	1.2	1.0
Home schooled	0.3	0.2	0.1
Missing	16.3	19.7	3.4
Poverty indicator			
Students eligible for free and reduced-price meals	66.5	65.9	0.6
Race/ethnicity:			
White, not Hispanic	38.6	39.7	-1.1
Black, not Hispanic	21.6	21.1	0.5
Hispanic	35.5	35.3	0.2
American Indian/Alaska Native	0.7	0.6	0.1
Asian/Pacific Islander	3.0	2.8	0.2
Other/missing	0.6	0.5	0.1
2005 (1st Grade)			
School type			
Public	89.1	90.3	-1.2
Charter	1.2	1.3	-0.1
Private	2.2	2.3	-0.1
Home schooled	0.1	0.2	-0.1
Missing	7.4	5.9	1.5
Poverty indicator			
Students eligible for free and reduced-price meals	66.3	66.5	-0.02
Race/ethnicity:			
White, not Hispanic	39.6	38.3	1.3
Black, not Hispanic	23.4	21.9	1.5
Hispanic	32.8	35.3	-2.5
American Indian/Alaska Native	0.7	0.7	0.0
Asian/Pacific Islander	3.1	3.3	-0.2
Other/missing	0.4	0.5	-0.1

*p≤0.10, **p≤0.05, ***p≤0.01

The Head Start and control groups did not differ on any of these measures. There were no significant differences in the types of schools or demographic characteristics of the students in the schools attended by children in the two groups.[74]

Among the states where the study children went to school, the average percentages of students achieving at least a proficient rating ranged from 30 to 94 percent in math and 31 to 92 percent in reading. As shown in Exhibit 3.21, the proficiency levels at the schools attended by the study children averaged between about 55 percent and 64 percent.[75] These levels were in the middle of the state averages, indicating that most of the schools attended by the study children were not among the worst or best schools in their respective states. There were no differences on these measures in the schools attended by the Head Start and control groups.

Exhibit 3.21: Average Percentage of Students At or Above Proficient in Math and Reading in Schools Attended by 4-Year-Old Cohort, by Year

		Average Percent At or Above Proficient in Schools Attended By:		
Year	Area	Head Start Group	Control Group	Magnitude of Impact
2004	Math	61.3	63.5	-2.2
(K)	Reading	54.9	56.5	-1.6
2005	Math	65.7	64.1	1.6
(1st)	Reading	59.4	57.0	2.4

*p≤0.10, **p≤0.05, ***p≤0.01

[74] As an indication of the overall academic ability within the schools attended by the study children, the percentages of students in the schools who scored at least at the "proficient level" on state math and reading assessments were obtained from the GreatSchools database. The exams and the requirements for proficient ratings are set by each state, and students' ratings depend on their performance on the exams. Proficiency data were not available for most private schools, schools that do not administer the tests (generally schools serving only lower grades such as K-2 do not require testing in those grades); a few schools that were not in the GreatSchools database; and do not include children who were home schooled. Missing data were comparable for Head Start and control children, but varied by cohort and year. In kindergarten, for the 4-year-old cohort, there was 23 percent missing data for reading and math proficiency levels; in 2005 (1st grade), it was reduced to 12 percent for the 4-year-old cohort and was 20 percent for the 3-year-old cohort in kindergarten. For the 3-year-old cohort in 1st grade, there was missing information on 12 percent of the students.

[75] In the GreatSchools database, the percentages of children achieving proficient ratings usually were presented for several grade levels within a school. Here the percentages refer to the lowest grade in the school for which ratings were available – generally 3rd or 4th grade – to correspond to the students closest in age to the study children.

The 4-Year-Old Cohort: Teacher and Classroom Characteristics

Exhibit 3.22 presents information on the characteristics of the classroom teacher, classroom environment, and classroom activities, for both the Head Start and control group children in kindergarten and 1st grade. The key findings presented below highlight differences between the experiences of the Head Start group and control group that might contribute to an understanding of long-term program impacts.

- There are statistically significant differences between the Head Start and control groups on the extent to which there was an adult assistant or volunteer in the classroom. In kindergarten, the control group was more likely to have an adult volunteer in the classroom than the Head Start group, while in 1st grade, the control group was more likely to have a paid assistant, co-teacher, or volunteer.
- No significant differences were found between the teachers of the Head Start and control children for any of the teacher qualifications measures (certification, educational attainment, educational preparation, and tenure), or on measures of job satisfaction, in either kindergarten or 1st grade.
- No significant differences were found in either year on a measure of teacher beliefs based on how children ought to be taught or on any other measures of classroom activities.

While there were very few statistically significant differences in experiences for the Head Start and control group children, the overall findings for both groups can also contribute to an understanding of the school environment experienced by both groups of children.

- There were, on average, about five Dual Language Learner's (DLL) in the children's kindergarten classrooms and approximately four in the 1st grade classrooms. Assuming an average elementary school class size of 20 children (U.S. Department of Education, National Center for Education Statistics, 2007), about 25 percent of the students in each kindergarten class were DLL and 20 percent in the 1st grade classes were DLL. This is a higher proportion than children nationally.[76]
- Nearly 50 percent of the children were in classrooms where the teachers reported well-behaved students, with slightly smaller percentages reporting occasional misbehavior and much smaller percentages reporting frequent misbehavior. These percentages held steady for both kindergarten and 1st grade years.

[76] Nationally about 13 percent of children in elementary school classrooms are DLL as reported in the 2003-04 NCES Schools and Staffing Survey (Strizek et al., 2006).

Exhibit 3.22: Percentage of Children by Kindergarten and 1st Grade School Experience, 4-Year-Old Cohort

Characteristic	Head Start Group	Control Group	Magnitude of Impact	p-value
2004 – Kindergarten Year				
Classroom Activities				
Children in classrooms where 7 or more of the 12 reading and language activities are done at least 3 times weekly	84.2%	88.2%	-4.0	0.43
Children in classrooms where 5 or more of the 8 math activities are done at least 3 times weekly	38.3%	33.2%	5.1	0.54
Average number of reading and language activities done daily	6.44	6.46	-0.02	0.90
Average number of math activities done daily	2.70	2.59	0.11	0.42
Over half of the daily class time is spent in whole-class or large group activities directed by an adult	31.4%	37.0%	-5.6	0.13
Over half of the daily class time is spent in individual or small group activities directed by an adult	18.7%	17.6%	1.1	0.69
Over half of the daily class time is spent in activities chosen by the child	2.7%	1.8%	0.9	0.35
Teacher and Classroom Characteristics				
Teacher had a state teaching certificate, teaching license, or teaching credential	91.7%	93.2%	-1.50	0.26
Teacher had a bachelor's degree or higher	93.1%	95.2%	-2.10	0.40
Mean number of college courses completed in:				
Elementary education	5.38	5.50	-0.12	0.18
Early childhood education	4.45	4.52	-0.07	0.68
Methods in teaching reading	3.65	3.64	0.01	0.95
Methods in teaching math	3.02	3.13	-0.11	0.55
Mean number of years teaching	13.7 years	13.5 years	0.32	0.75
Mean number of years employed at current school	8.8 years	8.1 years	0.70	0.26
Mean score based on teachers' beliefs on how children should be taught and managed in the classroom	3.66	3.66	0.00	1.00
Mean number of students who were Dual Language Learners	4.80	5.30	-0.50	0.22
Teacher enjoyed present teaching position	92.2%	94.6%	-2.40	0.32
Teacher believed he/she was making a difference in the lives of children	96.7%	97.2%	-0.50	0.20
Teacher would choose teaching again as a career	85.9%	87.8%	-1.90	0.26
Children in classrooms with at least one paid assistant, co-teacher, or team teacher in the class in a typical week	78.7%	79.2%	-0.50	0.46

Exhibit 3.22: Percentage of Children by Kindergarten and 1st Grade School Experience, 4-Year-Old Cohort (continued)

Characteristic	Head Start Group	Control Group	Magnitude of Impact	p-value
Teacher and Classroom Characteristics (cont'd)				
Children in classrooms with at least one adult volunteer assistant in the class in a typical week	47.2%	48.1%	-0.90	0.00***
Behavior of children in classroom as a group:				
Children are well-behaved	46.4%	45.4%	1.00	
Children misbehave occasionally	34.2%	39.3%	-5.10	1.48
Children misbehave frequently	11.9%	8.9%	3.00	
2005 – 1st Grade Year				
Classroom Activities				
Children in classrooms where 16 or more of the 25 reading and language activities are done at least 3 times weekly	50.0%	48.5%	1.5	0.74
Children in classrooms where 9 or more of the 18 math activities are done at least 3 times weekly	45.4%	44.7%	0.7	0.88
Average number of reading and language activities done daily	10.44	10.45	-0.1	0.98
Average number of math activities done daily	4.78	5.06	0.28	0.25
Over half of the daily class time is spent in whole-class or large-group activities directed by an adult	59.0%	59.1%	-0.1	0.99
Over half of the daily class time is spent in individual or small-group activities directed by an adult	15.2%	16.0%	0.8	0.71
Over half of the daily class time is spent in activities chosen by the child	0.8%	1.6%	0.8	0.38
Teacher and Classroom Characteristics				
Teacher had a state teaching certificate, teaching license, or teaching credential	98.1%	95.7%	2.4	0.20
Teacher had a bachelor's degree or higher	98.4%	97.1%	1.3	0.46
Mean number of college courses completed in:				
Elementary education	5.33	5.35	-0.02	0.88
Early childhood education	3.65	3.47	0.18	0.35
Methods in teaching reading	3.80	3.81	-0.01	0.96
Methods in teaching math	3.01	2.99	0.02	0.88
Mean number of years teaching	13.9 years	13.0 years	0.94	0.24
Mean number of years employed at current school	9.0 years	8.3 years	0.70	0.21
Mean score based on teachers' beliefs on how children should be taught and managed in the classroom	3.45	3.48	-0.03	0.15

Exhibit 3.22: Percentage of Children by Kindergarten and 1st Grade School Experience, 4-Year-Old Cohort (continued)

Characteristic	Head Start Group	Control Group	Magnitude of Impact	p-value
Teacher and Classroom Characteristics (cont'd)				
Mean number of students who were Dual Language Learners	3.9	4.6	-0.70	0.12
Teacher enjoyed present teaching position	92.0%	91.5%	0.5	0.84
Teacher believed he/she was making a difference in the lives of children	95.7%	95.8%	-0.1	0.65
Teacher would choose teaching again as a career	86.9%	87.7%	-0.8	0.44
Children in classrooms with at least one paid assistant, co-teacher, or team teacher in the class in a typical week	51.8%	59.2%	-7.4	0.00***
Behavior of children in classroom as a group:				
Children are well-behaved	45.5%	48.5%	-3.0	
Children misbehave occasionally	39.8%	41.9%	-2.1	0.18
Children misbehave frequently	13.9%	8.8%	5.9	

*p≤0.10, **p≤0.05, ***p≤0.01

- About a third of the children in kindergarten and nearly 60 percent in 1st grade spent over half of the school day in whole-class or large-group activities directed by an adult. Very few children spent over half of the school day in child-chosen activities. On average, kindergarten children participated in about six of the 12 reading and language activities and about three of the eight math activities that were included in the teacher survey on a daily basis. In the 1stgrade year, the averages were about 10 out of 25 reading and language activities and about five out of 18 math activities provided daily. (Exhibit 3.23 provides more detail on the learning activities showing the activities that were most frequently provided to children on a daily basis.)
- Overall, kindergarten and 1st grade teachers appeared to be credentialed, experienced, and well-rooted in their current schools. They reported being content in their jobs and believed they were making a difference in the lives of children.

Exhibit 3.23: Most Frequent Daily Activities, by Year, 4-Year-Old Cohort

Year	Type of Activities:	
	Reading and Language	**Math**
2004 – Kindergarten Year	▪ Have the children practice writing or spelling their names ▪ Practice the sounds that letters make (phonics) ▪ Work on learning the names of letters ▪ Listen to the teacher read stories where children see the print ▪ Practice writing the letters of the alphabet	▪ Talk about the calendar or days of the week ▪ Count out loud ▪ Counting things such as small toys and chips to learn math ▪ Play math games ▪ Work with shape blocks
2005 – 1st Grade Year	▪ Work on phonics ▪ Read aloud ▪ Discuss new or difficult vocabulary ▪ Read silently ▪ Read books children have chosen for themselves	▪ Engage in calendar-related activities ▪ Do math worksheets ▪ Count out loud ▪ Do math problems from their textbooks ▪ Explain how a math problem is solved

The 3-Year-Old Cohort: Overall School Characteristics

Like the 4-year-old cohort, the 3-year-old cohort attended schools with much higher levels of poverty than schools nationwide (as indicated by the proportions of students eligible for free and reduced-price meals) and were in schools with higher proportions of Black and Hispanic students than schools nationwide (see Exhibit 3.24). Nationwide, approximately 43 percent of students were eligible for free and reduced-price meals as compared to 68 percent of students at the schools the 3-year-old cohort attended. The study children attended schools in which about 60 percent of the enrollment was Black and Hispanic children compared to about 40 percent nationally.[77]

However, as with the 4-year-old cohort, there were no significant differences in the types of schools or demographic characteristics of the students in the schools attended by children in the Head Start and control groups. There were, however, cohort differences in the percentage of Black and Hispanic students attending schools. The percentage of Hispanic children in schools

[77] This represents 133,910 schools and excludes schools without K and 1st grade for the 2003-04 school year.

Exhibit 3.24: School Type and Student Characteristics for School Attended by Treatment and Control Groups by Year, 3-Year-Old Cohort

Characteristics of Children	Average Percent of Students in Schools Attended By: Head Start Group (%)	Average Percent of Students in Schools Attended By: Control Group (%)	Magnitude of Impact (%)
2005 (Kindergarten)			
School type			
Public	86.3	84.5	1.8
Charter	1.5	1.2	0.3
Private	5.8	6.7	-0.9
Home schooled	0.2	0.2	0
Missing	6.2	7.4	-1.2
Poverty indicator			
Students eligible for free and reduced-price meals	68.3	68.4	-0.1
Race/ethnicity			
White, not Hispanic	38.4	39.1	-0.7
Black, not Hispanic	30.3	29.3	1.0
Hispanic	27.0	26.5	0.5
American Indian/Alaskan Native	1.5	1.6	-0.1
Asian/Pacific Islander	2.1	2.6	-0.5
2006 (1st Grade)			
School type			
Public	91.9	91.6	0.3
Charter	1.3	0.8	0.5
Private	5.2	6.3	-1.1
Home schooled	0.0	0.1	0.1
Missing	1.6	1.3	0.3
Poverty indicator			
Students eligible for free and reduced-price meals	67.2	67.0	0.2
Race/ethnicity			
White, not Hispanic	40.1	41.3	-1.2
Black, not Hispanic	28.8	28.6	0.2
Hispanic	27.1	25.6	1.5
American Indian/Alaskan Native	1.5	1.5	0.0
Asian/Pacific Islander	2.3	2.5	-0.2

*p≤0.10, **p≤0.05, ***p≤0.01

attended by the 4-year-old cohort (Exhibit 3.20) was greater than the 3-year-old cohort (35% compared to 27%), and the percentage of Black children in schools attended by the 3-year-old cohort was greater than that of the 4-year-old cohort (21% compared to 30%).

As shown in Exhibit 3.25, the math and reading proficiency levels at the schools attended by the 3-year-old cohort ranged from 62.7 percent to 67.6 percent. These levels were in the middle of the state averages (30% to 94%), indicating that most of the schools attended by the study children were not among the worst or best schools in their respective states.

Exhibit 3.25: Average Percentage of Students At or Above Proficient in Math and Reading for Schools Attended by 3-Year-Old Cohort, by Year

Year	Average Percent At or Above Proficient in Schools Attended By			Magnitude of Impact (%)
	Area	Head Start Group	Control Group	
2005 (Kindergarten)	Math	64.6	67.0	-2.4**
	Reading	62.7	64.4	-1.7
2006 (1st Grade)	Math	66.2	67.6	-1.4
	Reading	63.8	64.2	-0.4

*p≤0.10, **p≤0.05, ***p≤0.01

Whereas no differences were found in the proficiency levels of the schools attended by the Head Start and control groups for the 4-year-old cohort, one difference was found for the 3-year-old cohort. In kindergarten, control children in the 3-year-old cohort attended schools with higher math proficiency levels. No other statistically significant differences were found in the proficiency levels of the Head Start and control group children.

The 3-Year-Old Cohort: Teacher and Classroom Characteristics

Exhibit 3.26 presents information on the characteristics of the classroom teacher, classroom environment, and classroom activities, for both the Head Start and control group children in kindergarten and 1st grade. The key findings, presented below, highlight Head Start and control group differences that might contribute to an understanding of long-term program impacts.

- In kindergarten, the Head Start group had teachers with more coursework in methods for teaching reading, and in 1st grade the Head Start group had teachers with slightly more coursework in early childhood education and methods for teaching reading.
- In 1st grade, a slightly higher proportion of Head Start children attended classrooms where nine or more of the 18 math activities were done at least three times weekly.

Exhibit 3.26: Percentage of Children by Kindergarten and 1st Grade School Experience, 3-Year-Old Cohort

Characteristic	Head Start Group	Control Group	Magnitude of Impact	p-value
2005 – Kindergarten Year				
Classroom Activities				
Children in classrooms where 7 or more of the 12 reading and language activities are done at least 3 times weekly	88.8%	88.1%	0.70	0.78
Children in classrooms where 5 or more of the 8 math activities are done at least 3 times weekly	38.7%	40.9%	-2.2	0.43
Mean number of reading and language activities done daily	6.79	6.86	-0.07	0.70
Mean number of math activities done daily	2.74	2.73	0.01	0.88
Over half of the daily class time is spent in whole-class or large-group activities directed by an adult	40.5%	39.9%	0.6	0.89
Over half of the daily class time is spent in individual or small-group activities directed by an adult	18.0%	16.5%	1.5	0.56
Over half of the daily class time is spent in activities chosen by the child	3.1%	4.0%	-0.9	0.48
Teacher and Classroom Characteristics				
Teacher had a state teaching certificate, teaching license, or teaching credential	92.7%	90.4%	2.3	0.19
Teacher had a bachelor's degree or higher	94.1%	93.1%	1.0	0.55
Mean number of college courses completed in:				
Elementary education	5.36	5.31	0.05	0.62
Early childhood education	4.43	4.50	-0.07	0.60
Methods in teaching reading	3.63	3.38	0.25	0.02**
Methods in teaching math	2.98	2.86	0.12	0.30
Mean number of years teaching	13.88	13.23	0.65	0.35
Mean number of years employed at current school	8.3	8.7	-0.40	0.46
Mean score based on teachers' beliefs on how children should be taught and managed in the classroom	3.64	3.66	-0.02	0.53
Mean number of students who were Dual Language Learners	3.5	3.7	-0.20	0.50
Teacher enjoyed present teaching position	95.0%	93.4%	1.6	0.38
Teacher believed he/she was making a difference in the lives of children	97.1%	95.8%	1.3	0.68
Teacher would choose teaching again as a career	85.3%	85.6%	-0.3	0.99
Children in classrooms with at least one paid assistant, co-teacher, or team teacher in the class in a typical week	70.2%	74.8%	-4.6	0.20

Exhibit 3.26: Percentage of Children by Kindergarten and 1st Grade School Experience, 3-Year-Old Cohort (continued)

Characteristic	Head Start Group	Control Group	Magnitude of Impact	p-value
Children in classrooms with at least one adult volunteer assistant in the class in a typical week	34.9%	33.5%	1.4	0.33
Behavior of children in classroom as a group:				
Children are well-behaved	43.4%	47.2%	-3.8	
Children misbehave occasionally	44.2%	40.7%	3.5	0.18
Children misbehave frequently	11.7%	10.3%	1.4	
2006 – 1st Grade Year				
Classroom Activities				
Children in classrooms where 16 or more of the 25 reading and language activities are done at least 3 times weekly	49.7%	45.2%	4.5%	0.19
Children in classrooms where 9 or more of the 18 math activities are done at least 3 times weekly	53.5%	46.1%	7.4%	0.04**
Mean number of reading and language activities done daily	10.33	9.86	0.47	0.16
Mean number of math activities done daily	5.07	4.90	0.17	0.37
Over half of the daily class time is spent in whole-class or large-group activities directed by an adult	48.1%	51.5%	-3.4	0.44
Over half of the daily class time is spent in individual or small-group activities directed by an adult	18.3%	18.7%	-0.4	0.98
Over half of the daily class time is spent in activities chosen by the child	1.1%	0.6%	0.5	0.28
Teacher and Classroom Characteristics				
Teacher had a state teaching certificate, teaching license, or teaching credential	95.6%	94.5%	1.1	0.17
Teacher had a bachelor's degree or higher	96.9%	95.5%	1.4	0.10
Mean number of completed college courses in:				
Elementary education	5.33	5.35	-0.02	0.88
Early childhood education	3.03	2.72	0.31	0.04**
Methods in teaching reading	3.67	3.37	0.30	0.07*
Methods in teaching math	2.93	2.74	0.19	0.18
Mean number of years teaching	13.07	12.89	0.18	0.79
Mean number of years employed at present school	8.20	7.90	0.3	0.60
Mean score based on teachers' beliefs on how children should be taught and managed in the classroom	3.42	3.37	0.05	0.04**
Mean number of students who were Dual Language Learners	2.9	3.5	-0.60	0.13
Teacher enjoys present teaching position	90.4%	90.7%	-0.3	0.45

Exhibit 3.26: Percentage of Children by Kindergarten and 1st Grade School Experience, 3-Year-Old Cohort (continued)

Characteristic	Head Start Group	Control Group	Magnitude of Impact	p-value
Teacher and Classroom Characteristics (cont'd)				
Teacher believes he/she is making a difference in the lives of children	95.4%	97.9%	-2.5	0.31
Teacher would choose teaching again as a career	82.6%	81.4%	1.2	0.54
Children in classrooms with at least one paid assistant, co-teacher, or team teacher in the class in a typical week	52.2%	51.8%	0.40	0.11
Children in classrooms with at least one adult volunteer assistant in the class in a typical week	31.1%	32.0%	-0.9	0.09*
Behavior of children in classroom as a group:				
Children are well-behaved	41.2%	44.8%	-3.6	
Children misbehave occasionally	43.9%	39.5%	4.4	
Children misbehave frequently	13.0%	14.9%	-1.9	0.38

*p≤0.10, **p≤0.05, ***p≤0.01

- In 1st grade, children in the Head Start group experienced teachers with slightly higher mean scores on the Teacher Belief Scale than the control group (a mean of 3.42, compared to 3.37). Also, there is suggestive evidence that children in the control group were more likely to be in classrooms with at least one adult volunteer assistant.

Like the 4-year-old cohort, there were very few statistically significant differences in experiences for the Head Start and control group children. However, the overall findings for both groups can also contribute to an understanding of the school environment that children experienced.

- Nearly seven out of 10 kindergarten children were in classrooms with at least one paid assistant, co-teacher, or team teacher in the class in a typical week, and about one-third had at least one adult volunteer assistant in their classroom in a typical week. In 1st grade, the percentages dropped to about 52 percent of classrooms with at least one paid assistant and 31 percent with an adult volunteer in a typical week.
- Overall, kindergarten and 1st grade teachers appeared to be credentialed, experienced, and well-rooted in their current schools. They reported being content in their jobs and believed they were making a difference in the lives of children. In rating the children's behavior as a class, approximately 40 percent of the children were in classrooms where the teachers reported well-behaved students, with almost equal percentages reporting occasional misbehavior, and much smaller percentages reporting frequent misbehavior. These percentages held steady for both kindergarten and 1st grade years.

- More than a third (40 percent) of the children in kindergarten and about 50 percent in 1st grade spent over half of the school day in whole-class or large-group activities directed by an adult. Very few children spent over half of the school day in child-chosen activities.
- In kindergarten, children participated in about seven of the 12 assessed reading and language activities and about three of the eight assessed math activities on a daily basis, on average. In the 1st grade, the averages were about 10 out of 25 reading and language activities and five out of 18 math activities provided daily. (Exhibit 3.27 provides more detail on the learning activities showing the activities that were most frequently provided to children on a daily basis.)

Exhibit 3.27: Most Frequent Daily Activities, by Year, 3-Year-Old Cohort

Year	Type of Activities	
	Reading and Language	**Math**
2004 – Kindergarten Year	▪ Practice the sounds that letters make (phonics) ▪ Have the children practice writing or spelling their names ▪ Work on learning the names of letters ▪ Listen to teacher read stories where children see the print ▪ Practice writing the letters of the alphabet	▪ Talk about calendar or days of the week ▪ Count out loud ▪ Counting things such as small toys and chips to learn math ▪ Play math games ▪ Work with shape blocks
2005 – 1st Grade Year	▪ Work on phonics ▪ Read silently ▪ Read aloud ▪ Discuss new or difficult vocabulary ▪ Read books children have chosen for themselves	▪ Engage in calendar-related activities ▪ Do math worksheets ▪ Count out loud ▪ Do math problems from their textbooks ▪ Explain how a math problem is solved

Summary

This chapter focused on the experiences of children and the services they received during their preschool years (when they may have been in Head Start or other child care environments), as well as during their kindergarten and 1st grade years. It provided a detailed longitudinal look at what access to Head Start means for children, including what, if any, difference it makes in the type of care-giving arrangement; whether and where the children attend preschool, Head Start, or child care; characteristics of their early childhood care and education; characteristics of their early elementary education settings; qualifications of their caregivers and teachers; and quality of

the services received in these settings. In addition, highlights of the variation in services that Head Start children received at Head Start centers and the characteristics related to the 3-year-old cohort's attending one or two years of Head Start were presented. The key findings are:

4-Year-Old Cohort

Head Start Year

- As depicted by the characteristics of the centers and classrooms attended by the Head Start group children, there was variation across the types of services and experiences that children receive in Head Start centers. The majority of Head Start children were in centers with ECERS-R scores of at least a five (approximately 70%) and in classrooms that emphasized language and literacy and math activities (approximately 60%). Yet there was also variety in children's experiences. The remaining children did not experience centers with these high ECERS-R scores or an emphasis on language and literacy or math activities.
- Head Start increased the likelihood that low-income children spent a significant amount of time in nonparental care and, specifically, in center-based care settings. The Head Start group (children who had access to Head Start) was twice as likely as the control group to enroll in a center-based program. Most commonly these children were enrolled in Head Start. Conversely, control group children were substantially more likely than Head Start group children to receive care exclusively from a parent at home and not to attend a center-based care setting.
- There are statistically significant differences between the Head Start and control groups on every measure of children's preschool experiences measured in this study. These measures included but were not limited to teacher training and education, classroom activities, classroom teacher to child ratio, and classroom observations of overall quality and child-teacher interactions. Children with access to Head Start had more positive experiences on these measures than children in the control group, and these differences persist whether or not children in exclusively parent care are included.
- Some of the largest differences between the Head Start and control groups were on measures governed by the Head Start Program Performance Standards, i.e., services available for children and families, use of curriculum, and instructional practices.

Early Elementary School Years

- Access to Head Start did not appear to have an overall impact on the schools that children attended in kindergarten and 1st grade or on their early elementary education experiences. With only a few exceptions, teacher, classroom, and school characteristics did not differ significantly between children in the program group and those in the control group. For the few noted differences, some favored the Head Start group, and others favored the control group.

- In general, the children in this study went to schools with larger populations of low-income children and racial, ethnic, and linguistic minorities than children nationwide. These schools tended to score in the middle of their states' ranges on statewide performance tests.

3-Year-Old Cohort

Head Start Years

- As with the 4-year-old cohort, there was variation across the types of services and experiences that the 3-year-old cohort children received in Head Start centers.
- Access to Head Start increased the likelihood that low-income children were enrolled in center-based care in the first year of the study. Approximately 90 percent of the Head Start group children were in center-based care compared to 43 percent of the control group children. Like the 4-year-old cohort, the 3-year-old Head Start group children were predominantly in Head Start. Conversely, approximately 40 percent of the control group children were exclusively in parent care compared to seven percent of the program group children.
- For the 3-year-old cohort's first year of preschool, every measure of children's preschool experiences favored the Head Start group as compared to the control group, and in most cases the magnitude of difference is quite large. These measures included but were not limited to teacher training and education, classroom activities, classroom teacher to child ratio, and observations of classroom quality and child-teacher interactions. The differences persist whether or not children in exclusively parental care are included.
- The impact on the 3-year-old cohort's second year of preschool was very different. By design, the control group of the 3-year-old cohort was allowed to enter Head Start during the second year. In fact, nearly half of the control group children did enroll in Head Start that second year. The early childhood care and education experiences for children in the program group and children in the control group were far more similar, with very few differences on measures of the types or quality of programs children attended that year.
- Of those 3-year-olds that attended a Head Start year, about 72 percent returned to Head Start for a second year. Characteristics related to an increased likelihood of returning for a second year included less competition from other early childhood programs in the area, only full-day classrooms, parental satisfaction with how the center supported and respected their family's culture and background, coming from a household in which the home language was Spanish, or having a mother who was a recent immigrant.

Early Elementary School Years

- Access to Head Start did not appear to have an overall impact on the quality of schools that 3-year-old cohort children attended in kindergarten and 1st grade or on children's early elementary education experiences with the exception of the average

math proficiency levels of the schools the Head Start and control group children attended in kindergarten. In kindergarten, control group children attended schools with higher math proficiency levels. With only minor other exceptions, characteristics of the teachers, classrooms, and schools did not differ significantly between children in the Head Start group and children in the control group. Among the few exceptions found, significant impacts were split. Some favored the Head Start group, while others favored the control group.

The findings in this chapter provide a context for understanding impacts in the four domains described in upcoming chapters. However, the measurement of quality in early childhood care and education settings is an issue of great debate. Thus, the measures collected as part of this study provide an incomplete picture of what constitutes the total early childhood experiences of the study children. A host of factors could not be measured, and some of the variables that are discussed in this chapter are only proxies for children's educational and developmental experiences. Nevertheless, these results show that for the characteristics measured in this study, having access to Head Start tended to result in more positive experiences for children in the Head Start group during their Head Start year. Access to Head Start did not lead to longer term differences in their schooling environments, at least based on the limited measures available for this wave of the study. This suggests that the impacts discussed in subsequent chapters come from those differences in children's experiences during that year that they were in Head Start and not necessarily from longer term differences in children's early care and education.

Chapter 4: The Impact of Head Start on Children's Cognitive Development

Introduction

This chapter focuses on the impact of access to and participation in Head Start on children's development of language and literacy, math, and pre-writing skills. Information on children's development in these domains is derived from direct assessments, as well as reports from parents and from children's teachers once they began kindergarten. Results are presented separately for the 3- and 4-year-old cohorts and include both annual impacts and (for selected outcomes) longitudinal growth analyses, covering the period from application to Head Start through the end of the 1st grade. For those children who were identified as Dual Language Learners, separate results are presented for two tests that were administered to them in Spanish each year.

Measures

The development of early language, pre-writing, math, and literacy skills is important for children's later success in reading, writing, and math in the elementary grades. Thus, the instruments for this study were chosen to represent strong measures of school readiness, particularly in the areas of reading and mathematics. As discussed in more detail below and in Chapter 2, the selected instruments had been used in national studies and evaluations and demonstrated adequate psychometric properties.

Language and Literacy: Research has documented that vocabulary, letter recognition, and phonological awareness are necessary skills in the process of learning to read. Biemiller (2006) has noted that vocabulary is a significant predictor of reading comprehension and a study by Cunningham and Stanovich (1997) showed "a substantial relationship between oral receptive vocabulary in 1st grade and reading comprehension in 11th grade." Rock and Pollack (2002) identified five clusters of test items (letter recognition, beginning sounds, ending sounds, sight words, and comprehension of words in context) for the Early Childhood Longitudinal Study-Kindergarten Cohort (ECLS-K) that measured language and literacy skills in the transition from prereading to reading. Whitehurst and Lonigan (1998) and Ehri and Roberts (2006) reported

letter recognition and phonemic awareness as essential skills to becoming a proficient reader. Consequently, the direct assessments selected to measure language and literacy skills include measures of vocabulary, letter recognition, phonological awareness, and comprehension. A summary of the direct child assessment measures by year and sub-domain (e.g., language and literacy, Spanish language and literacy, pre-writing, and math) is provided in Exhibit 4.1.[78]

Spanish Language and Literacy Skills: The emergence of greater cultural diversity in the population and the rise in the number of children served by Head Start who are Dual Language Learners have combined to increase the challenge and responsibility of Head Start to be responsive to each child's and each family's unique needs and ethnic, cultural, and linguistic heritage. Supporting the home language and culture of Dual Language Learners is a goal of the Head Start program. Understanding the proficiency of Dual Language Learners in their home language can be helpful in understanding their progress or lack of progress in their acquisition of English language skills. To measure baseline skills, as well as the growth in Spanish language and literacy skills, all children with a home language of Spanish were administered two Spanish subtests during each data collection period.

Pre-Writing Skills: A school readiness battery usually includes a measure of fine motor skills or pre-writing. Fine motor skills are addressed in the Head Start Performance Measures, and pre-writing skills are measured in the child assessment battery used for this study.

Math Skills: Duncan et al. (2007) reported that the strongest predictors of later achievement are school-entry math, reading, and attention skills, with early math skills showing the greatest predictive power. Measures of counting, one-to-one correspondence, numeral identification, and solving simple arithmetic problems are included in the math skills domain of the child assessment battery. These measures are addressed in the Curriculum and Evaluation Standards (1989) developed by the National Council of Teachers of Mathematics and incorporated into the objectives of the Head Start Performance Measures.

[78] See Chapter 2 and the Technical Report for the Head Start Impact Study for detailed information on the cognitive measures.

Exhibit 4.1: Cognitive Outcomes by Year

Outcomes
Baseline Year
Language and Literacy Measures
Peabody Picture Vocabulary Test (PPVT) (Adapted)
Woodcock-Johnson III Letter-Word Identification
Woodcock-Johnson III Spelling
Woodcock-Johnson III Oral Comprehension
Preschool Comprehensive Test of Phonological and Print Processing (CTOPPP) Elision
Color Identification+
Woodcock-Johnson III Pre-Academic Skills
Spanish Language and Literacy Measures++
Test de Vocabulario en Imágenes Peabody (TVIP) (Adapted)
Batería Woodcock-Muñoz Identificación de letras y palabras
Pre-Writing Measure
McCarthy Draw-a-Design+
Parent-Reported Literacy Measure
Parent-Reported Emergent Literacy Scale (PELS)
Math Skills Measures
Woodcock-Johnson III Applied Problems
Counting Bears+
Head Start Year(s)
Language and Literacy Measures
Peabody Picture Vocabulary Test (PPVT) (Adapted)
Woodcock-Johnson III Letter-Word Identification
Woodcock-Johnson III Spelling
Woodcock-Johnson III Oral Comprehension
Preschool Comprehensive Test of Phonological and Print Processing (CTOPPP) Elision
Color Identification
Letter Naming
Woodcock-Johnson III Pre-Academic Skills
Spanish Language and Literacy Measures
Test de Vocabulario en Imágenes Peabody (TVIP) (Adapted)
Batería Woodcock-Muñoz Identificación de letras y palabras
Pre-Writing Measure
McCarthy Draw-a-Design
Parent-Reported Literacy Measure
Parent-Reported Emergent Literacy Scale (PELS)
Math Skills Measures
Woodcock-Johnson III Applied Problems
Counting Bears

Exhibit 4.1: Cognitive Outcomes by Year (continued)

Kindergarten Year
Language and Literacy Measures
Peabody Picture Vocabulary Test (PPVT) (Adapted)
Woodcock-Johnson III Letter-Word Identification
Woodcock-Johnson III Spelling
Woodcock-Johnson III Oral Comprehension
Preschool Comprehensive Test of Phonological and Print Processing (CTOPPP) Elision
Letter Naming
Woodcock-Johnson III Pre-Academic Skills
Woodcock-Johnson III Word Attack
Woodcock-Johnson III Basic Reading Skills
Spanish Language and Literacy Measures
Test de Vocabulario en Imágenes Peabody (TVIP) (Adapted)
Batería Woodcock-Muñoz Identificación de letras y palabras
Math Skills Measures
Woodcock-Johnson III Applied Problems
Woodcock-Johnson III Quantitative Concepts
Woodcock-Johnson III Math Reasoning
School Performance Measures
School Accomplishments
Promotion (Parent-reported)
Language and Literacy Ability
Math Ability
Social Studies and Science Ability
1st Grade Year
Language and Literacy Measures
Peabody Picture Vocabulary Test (PPVT) (Adapted)
Woodcock-Johnson III Letter-Word Identification
Woodcock-Johnson III Spelling
Woodcock-Johnson III Oral Comprehension
Woodcock-Johnson III Pre-Academic Skills
Woodcock-Johnson III Word Attack
Woodcock-Johnson III Basic Reading Skills
Woodcock-Johnson III Academic Applications
Woodcock-Johnson III Academic Skills
Woodcock-Johnson III Passage Comprehension
Woodcock-Johnson III Writing Samples
Spanish Language and Literacy Measures
Test de Vocabulario en Imágenes Peabody (TVIP) (Adapted)
Batería Woodcock-Muñoz Identificación de letras y palabras
Math Skills Measures
Woodcock-Johnson III Applied Problems
Woodcock-Johnson III Quantitative Concepts
Woodcock-Johnson III Math Reasoning
Woodcock-Johnson III Calculation

Exhibit 4.1: Cognitive Outcomes by Year (continued)

School Performance Measures
School Accomplishments
Promotion (Parent-reported)
Language and Literacy Ability
Math Ability
Social Studies and Science Ability

+ Indicates tests administered to children in fall 2002 who spoke neither English nor Spanish.

++ In fall 2002, Dual Language Learners on the mainland were administered the following tests: PPVT, Woodcock-Johnson III Letter-Word Identification, CTOPPP Elision (Spanish), McCarthy Draw-A-Design (Spanish), Color Identification (Spanish), Counting Bears (Spanish), Test de Vocabulario en Imágenes Peabody (TVIP), Batería Woodcock-Muñoz Identificación de letras y palabras, Batería Woodcock-Muñoz Problemas aplicados, and Batería Woodcock-Muñoz Dictado.

Parenting Measures and School Performance Measures: In addition to the direct child assessment measures, parents and teachers provided information on the child's educational development. During the Head Start years, parents reported on their child's emergent literacy skills, and during the early school years (kindergarten and 1st grade) parents reported on their child's promotion to the next grade. In kindergarten and 1st grade, teachers reported on children's school performance measures. The parent-reported emergent literacy scale is based on a series of questions about how many letters of the alphabet the child knows, how many colors he or she can identify, how high he or she can count, whether the child can write his or her first name, etc. The teacher reported measures included ratings for each child's overall school accomplishments and the child's ability in language and literacy, math, and science and social studies. The parent and teacher reported cognitive outcomes also are presented in Exhibit 4.1.

Presentation of Impact Estimates

This chapter focuses on the estimated impact of access to Head Start on children's development of language and literacy, math, and pre-writing skills. Information on children's development in these domains is derived from researcher-administered direct child assessments as well as from reports from parents about their child's ability and grade promotion, and reports from teachers once the children enter kindergarten about subject-specific ability evidenced in the classroom, as well as overall school accomplishments.

The estimated impact of Head Start on these outcome measures is first presented for the older 4-year-old cohort and then for the 3-year-old cohort. As discussed in Chapter 2, the

impacts for the 3-year-old cohort are somewhat more complicated due to their having an opportunity to experience two years of Head Start prior to entering kindergarten. For each cohort, annual impacts are first presented for each outcome by year (e.g., Head Start, kindergarten, and 1st grade), and then, for a subset of the outcome measures, a separate longitudinal analysis examines the impact of Head Start on children's cognitive growth across the years.

The annual and longitudinal impact analyses both measure the effect of Head Start on the average child randomly assigned to the Head Start group. These analyses (referred to as "intent to treat" (ITT) impact estimates) include all of the children randomly assigned to the Head Start group and all of the children assigned to the control group in fall 2002 (the methods used to generate these estimates are discussed in Chapter 2 and the Technical Report for the Head Start Impact Study). The resulting impact estimates represent the effect of providing Head Start **access** to program-eligible children.

However, as mentioned in previous chapters, some children who were assigned to the Head Start group never went to Head Start, and some children assigned to the control group found their way into a Head Start classroom. The presence of children who never went to Head Start in the Head Start group—and those who did go to Head Start in the control group—means that the impacts of having access to Head Start will differ from the impacts of actually receiving Head Start services. Consequently, a separate analysis of the impact of Head Start on actual participants—referred to as the "impact on the treated" (IOT)—is also provided at the end of this chapter (the detailed IOT tables are presented in Appendix E). The methods used to generate this second set of impact estimates, and the underlying assumptions, are discussed in Chapter 2 and the Technical Report for the Head Start Impact Study.

The annual impact estimates, which inform most of the discussion in this chapter, are presented in two tables that provide means for both the Head Start group and the control group, differences between the Head Start group and control group means, regression-adjusted estimates

of program impact, and their statistical significance and effect sizes[79] for each outcome by year.[80] As discussed in Chapter 2, because of the large number of statistical tests reflected in these tables we have established three separate categories of statistically significant results[81] and use this language throughout this report:

- ***Strong Evidence***: the estimated impact for a particular outcome is statistically significant at the typical level ($p \leq 0.05$), and this result holds up under the test for multiple comparisons. (See Chapter 2 for a discussion of the Benjamini-Hochberg multiple comparison procedure.)
- ***Moderate Evidence***: the estimated impact for a particular outcome is statistically significant at the typical level ($p \leq 0.05$), but this result *does not* hold up under the test for multiple comparisons.
- ***Suggestive Evidence***: the estimated impact for a particular outcome is statistically significant under a relaxed standard ($p \leq 0.10$), and this result *may or may not* hold up under the test for multiple comparisons.

When reading the impact tables, the regression-adjusted impact is bolded if the impact is statistically significant using the Benjamini-Hochberg guidelines for multiple comparisons discussed in Chapter 2. The level of significance for the regression adjusted impact is indicated by asterisks as follows:

- Three asterisks (***), indicate that the p-value is less than or equal to 0.01 (i.e., the impact is statistically different from zero at the 99 percent confidence level);
- Two asterisks (**), indicate that the p-value is less than or equal to 0.05 (i.e., the impact is statistically different from zero at the 95 percent confidence level); and
- One asterisk (*), indicates that the p-value is less than or equal to 0.10 (i.e., the impact is statistically different from zero at the 90 percent confidence level).

The first set of impact estimates discussed in this chapter measures the effect of access to Head Start on the average child randomly assigned to the Head Start group (the intent to treat or

[79] 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.

[80] Normed percentiles for the Woodcock-Johnson III are found in the Technical Report for the Head Start Impact Study.

[81] If we fail to identify a statistically significant difference, we do not have conclusive evidence that the program "doesn't work." Rather, statistically insignificant impacts mean that the effect is indeterminate-access to Head Start may or may not have had a non-zero impact on a particular outcome, and we cannot with this study sample make a confident conclusion either way. The one thing that will be known with confidence is that a large true impact has not occurred.

ITT estimates), while the end of the chapter provides a discussion of the impact of Head Start on the children who actually participated in the program (i.e., the impact on the treated or the IOT estimates). (See Chapter 2 and the Technical Report for the Head Start Impact Study for a discussion of the methodology used for the impact estimates.)

As a final note, it is important for the reader to remember that, as discussed in Chapters 2 and 3, the control group in this study does not represent a no-service comparison group. Most of these children received some form of non-parental care, many of them in formal child care or preschool centers. Thus, the analyses presented here provide the answer to the policy question, "What is the impact of Head Start compared to other services available to income eligible children?"

Impacts for the 4-Year-Old Cohort

Annual Impacts on Language and Literacy Measures

Research demonstrates that children acquire the basic skills that help them learn to read at a very young age. There is consensus that the key elements to learning to read are oral language, phonological processing, and print awareness (Whitehurst and Lonigan, 2001). Mastering these skills for young children is necessary for school readiness and later success in school. As noted by the Carnegie Corporation, *"...if today's youngsters cannot read with understanding, think about and analyze what they've read, and then write clearly and effectively about what they've learned and what they think, then they may never be able to do justice to their talents and their potential"* (Graham and Perin, 2007). Although Head Start children are not at the age where they are expected to read, mastering language and pre-literacy skills is critical for young children since limited early language and pre-literacy skills tend to translate into persistent deficits in school and later life.

Given the importance of language development, the Head Start Impact Study conducted direct assessments of children's skills and ability at each wave of data collection. As shown in Exhibit 4.2,[82] at the end of the Head Start year for 4-year-olds (i.e., spring 2003), strong evidence

[82] Detailed tables for the impact estimates, including sample sizes, standard errors, and confidence intervals, can be found on the Administration for Children and Families, Office of Planning, Research and Evaluation Website at: http://www.acf.hhs.gov/programs/opre/hs/impact_study/index.html.

of program impacts was found on six child assessment outcomes related to children's language and literacy development:

- PPVT (Adapted), which measures receptive vocabulary, i.e., listening comprehension for the spoken word in standard English (estimated size of the impact, called the effect size= 0.09);
- Woodcock-Johnson III Letter-Word Identification, which measures symbol, letter, and word identification skills (effect size=0.22);
- Woodcock-Johnson III Spelling, which measures the child's ability to correctly write orally presented letters and words (effect size=0.15);
- Color Identification, which assesses the child's ability to correctly name presented colors (effect size=0.16);
- Letter Naming, which measures the child's ability to recognize the letters of the alphabet (effect size=0.25); and,
- Woodcock-Johnson III Pre-Academic Skills, which provides an overall academic measure of children's academic development including their pre-reading and letter and word identification skills, developing mathematics skills, and early writing and spelling skills (effect size=0.19).

Although these effects were found at the end of Head Start, subsequent impacts on the battery of direct child assessments, either at the end of kindergarten or at the end of 1st grade, are limited to a single suggestive finding of an impact on the PPVT (Adapted) at the end of 1st grade (effect size=0.09). Additionally, no statistically significant impacts were identified on children's language and literacy as reported by teachers during the children's kindergarten and 1st grade years.

Annual Impact on Parent-Reported Literacy Measure

In addition to the direct child assessments, parents were asked to report on their child's language and literacy skills (e.g., knowledge of the letters of the alphabet, reading books, recognizing own name, early writing, etc.) during the child's Head Start year. As shown in Exhibit 4.2, there is strong evidence of an impact on the Parent Emergent Literacy Scale at the end of Head Start (effect size=0.31), mirroring the results of the standardized language and literacy assessments. This measure was not administered in subsequent years.

Exhibit 4.2: Estimated Impacts on Cognitive Outcomes by Year: 4-Year-Old Cohort

Outcomes	Mean Estimates				Regression-Adjusted Impact		Effect Size
	Head Start Group	Control Group	Head Start - Control	p-value	Impact	p-value	
Head Start Baseline (Fall 2002)							
Language and Literacy Measures++							
PPVT (Adapted)	270.54	271.73					
WJ III Letter-Word Identification	309.63	306.73					
WJ III Spelling	359.73	356.41					
WJ III Oral Comprehension	442.92	445.43					
CTOPPP Elision	272.24	271.76					
Color Identification	0.58	0.51					
WJ III Pre-Academic Skills	357.81	357.34					
Spanish Language and Literacy Measures+++							
TVIP (Adapted)	263.08	270.47					
WM Letter-Word Identification	351.90	356.14					
Pre-writing Measure++							
McCarthy Draw-a-Design	3.88	3.77					
Parent-Reported Literacy Measure							
Emergent Literacy Scale	2.94	2.58					
Math Skills Measures++							
WJ III Applied Problems	391.58	389.55					
Counting Bears	0.41	0.39					
Head Start Year (Spring 2003)							
Language and Literacy Measures							
PPVT (Adapted)	294.35	290.25	4.10	0.060	**3.55****	0.028	0.09
WJ III Letter-Word Identification	325.46	319.22	6.24	0.034	**5.98****	0.017	0.22
WJ III Spelling	371.56	367.67	3.89	0.046	**3.77****	0.029	0.15
WJ III Oral Comprehension	443.40	443.65	-0.24	0.818	-0.94	0.395	-0.05
CTOPPP Elision	273.85	271.41	2.45	0.463	2.45	0.444	0.05
Color Identification	0.73	0.66	0.07	0.019	**0.08****	0.010	0.16
Letter Naming	11.53	9.21	2.33	0.008	**2.36*****	0.002	0.25
WJ III Pre-Academic Skills	365.00	360.56	4.45	0.041	**4.23****	0.022	0.19

Exhibit 4.2: Estimated Impacts on Cognitive Outcomes by Year: 4-Year-Old Cohort (continued)

Outcomes	Mean Estimates				Regression-Adjusted Impact		Effect Size
	Head Start Group	Control Group	Head Start - Control	p-value	Impact	p-value	
Spanish Language and Literacy Measures+++							
TVIP (Adapted)	298.54	290.77	7.77	0.380	9.04	0.106	0.21
WM Letter-Word Identification	360.70	359.23	1.47	0.328	1.91	0.180	0.14
Pre-writing Measure							
McCarthy Draw-a-Design	4.58	4.40	0.19	0.135	0.20	0.110	0.10
Parent-Reported Literacy Measure							
Emergent Literacy Scale	3.76	3.35	0.42	0.000	**0.43*****	0.000	0.31
Math Skills Measures							
WJ III Applied Problems	397.47	394.42	3.05	0.178	3.17	0.139	0.12
Counting Bears	0.59	0.55	0.04	0.185	0.04	0.181	0.08
Kindergarten Year (Spring 2004)							
Language and Literacy Measures							
PPVT (Adapted)	334.21	331.85	2.37	0.398	1.78	0.328	0.04
WJ III Letter-Word Identification	378.08	378.15	-0.08	0.970	-0.19	0.918	-0.01
WJ III Spelling	413.91	414.12	-0.21	0.899	-0.52	0.764	-0.02
WJ III Oral Comprehension	456.52	457.29	-0.77	0.551	-0.91	0.327	-0.05
CTOPPP Elision	321.89	323.91	-2.02	0.586	-2.85	0.374	-0.06
Letter Naming	22.99	22.65	0.34	0.351	0.40	0.274	0.06
WJ III Pre-Academic Skills	406.23	406.48	-0.26	0.868	-0.47	0.745	-0.02
WJ III Word Attack	431.60	432.68	-1.09	0.628	-1.13	0.639	-0.03
WJ III Basic Reading Skills	404.79	405.39	-0.60	0.765	-0.71	0.728	-0.02
Spanish Language and Literacy Measures+++							
TVIP (Adapted)	326.48	327.18	-0.70	0.927	-1.03	0.868	-0.02
WM Letter-Word Identification	390.55	396.10	-5.55	0.062	-4.28	0.130	-0.16

Exhibit 4.2: Estimated Impacts on Cognitive Outcomes by Year: 4-Year-Old Cohort (continued)

Outcomes	Mean Estimates				Regression-Adjusted Impact		Effect Size
	Head Start Group	Control Group	Head Start - Control	p-value	Impact	p-value	
Math Skills Measures							
WJ III Applied Problems	426.59	426.32	0.27	0.872	0.12	0.936	0.01
WJ III Quantitative Concepts	441.83	441.88	-0.05	0.968	-0.13	0.920	-0.01
WJ III Math Reasoning	434.15	434.12	0.03	0.981	-0.07	0.951	0.00
School Performance Assessment Measures							
School Accomplishments	28.13	28.16	-0.03	0.969	0.00	0.997	0.00
Promotion	0.94	0.92	0.01	0.569	0.00	0.888	0.01
Language and Literacy Ability	0.75	0.71	0.04	0.381	0.04	0.424	0.08
Math Ability	0.80	0.75	0.05	0.164	0.05	0.191	0.11
Social Studies and Science Ability	0.83	0.80	0.03	0.433	0.03	0.501	0.07
1st Grade Year (Spring 2005)							
Language and Literacy Measures							
PPVT (Adapted)	363.07	358.74	4.34	0.075	2.95*	0.072	0.09
WJ III Letter-Word Identification	433.01	432.26	0.75	0.730	0.83	0.705	0.02
WJ III Spelling	451.88	450.13	1.76	0.312	1.55	0.347	0.06
WJ III Oral Comprehension	473.42	472.36	1.06	0.438	0.34	0.717	0.02
WJ III Pre-Academic Skills	446.66	445.44	1.22	0.412	0.95	0.510	0.04
WJ III Word Attack	469.10	467.41	1.69	0.344	1.71	0.324	0.05
WJ III Basic Reading Skills	451.04	449.81	1.22	0.521	1.08	0.550	0.03
WJ III Academic Applications	461.77	461.22	0.55	0.606	0.38	0.730	0.02
WJ III Academic Skills	449.02	447.71	1.30	0.380	1.11	0.446	0.05
WJ III Passage Comprehension	450.28	449.86	0.42	0.814	0.17	0.922	0.01
WJ III Writing Sample	479.87	479.75	0.12	0.863	0.15	0.824	0.01

Exhibit 4.2: Estimated Impacts on Cognitive Outcomes by Year: 4-Year-Old Cohort (continued)

Outcomes	Mean Estimates				Regression-Adjusted Impact		Effect Size
	Head Start Group	Control Group	Head Start - Control	p-value	Impact	p-value	
Spanish Language and Literacy Measures+++							
TVIP (Adapted)	376.86	372.20	4.65	0.361	5.25	0.240	0.13
WM Letter-Word Identification	407.84	415.07	-7.23	0.164	-4.30	0.397	-0.09
Math Skills Measures							
WJ III Applied Problems	455.16	454.13	1.03	0.405	0.81	0.523	0.04
WJ III Quantitative Concepts	461.79	461.28	0.51	0.714	0.31	0.819	0.02
WJ III Math Reasoning	458.36	457.67	0.68	0.580	0.47	0.705	0.03
WJ III Calculation	461.76	460.46	1.30	0.245	1.41	0.255	0.07
School Performance Assessment Measures							
School Accomplishments	43.25	43.79	-0.54	0.481	-0.59	0.500	-0.06
Promotion	0.93	0.92	0.01	0.323	0.01	0.376	0.05
Language and Literacy Ability	0.69	0.70	-0.01	0.705	-0.02	0.433	-0.05
Math Ability	0.77	0.81	-0.04	0.235	-0.05	0.148	-0.12
Social Studies and Science Ability	0.83	0.85	-0.02	0.449	-0.02	0.362	-0.06

Key:
******* $p \leq 0.01$
** $p \leq 0.05$
* $p \leq 0.10$

++ Indicates baseline scores for English-speaking children only except for the PPVT and Woodcock-Johnson III Letter-Word test.

+++ Indicates scores for only the Dual Language Learners on the mainland.

Bold regression-adjusted impact value indicates the outcome passes the Benjamini-Hochberg test for multiple comparisons with a 10 percent false discovery rate.

NOTE: Some cognitive findings in this report may differ from the findings reported in the *Head Start Impact Study: First Year Findings* (U.S. Department of Health and Human Services, 2005) due to redefining of variables (i.e., Color Identification and Counting Bears) and residualization. See Chapter 2 and the Technical Report for the Head Start Impact Study for details of these changes.

Annual Impacts on Spanish Language and Literacy Measures

As shown in Exhibit 4.2, no statistically significant impacts were found in any year on either of the two Spanish-language tests that were administered annually to children who were identified as English language learners at the time of random assignment. These findings suggest that providing access to Head Start does not appear to affect the home language skills of English language learners. It should be kept in mind, however, that this is a subgroup of children in the 4-year-old cohort and, as a result, the smaller sample size provides less statistical power to detect any true impacts.

Annual Impacts on Math Skills Measures

While language and literacy skills are important for a child's development, there is growing acknowledgment of the need for children to develop early math skills as well. As noted by the National Council of Teachers of Mathematics, *"The need to understand and be able to use mathematics in everyday life and in the workplace has never been greater and will continue to increase"* (National Council of Teachers of Mathematics, 2000). Thus, the Head Start Impact Study included three direct math assessments, as well one math composite.

As shown in Exhibit 4.2, no statistically significant impacts were found on any of these direct child assessment outcomes in any year of the study. Additionally, no statistically significant impacts were identified on children's math ability as reported by teachers during the children's kindergarten and 1st grade years.

Annual Impacts on School Performance Measures

Additional parent and teacher assessments of children's skills and achievement were obtained for each child when he/she was in kindergarten and 1st grade. These measures include the teacher-reported measures of language, literacy, and math skills discussed above, as well as teachers' assessments of the child's overall school accomplishments and ability in social studies and science, and parents' report on their child's promotion to the next grade. These ratings are an important addition to the standardized assessments, as both parents and teachers see these children in a broader context over the course of the year and can presumably assess their performance as it translates into school success.

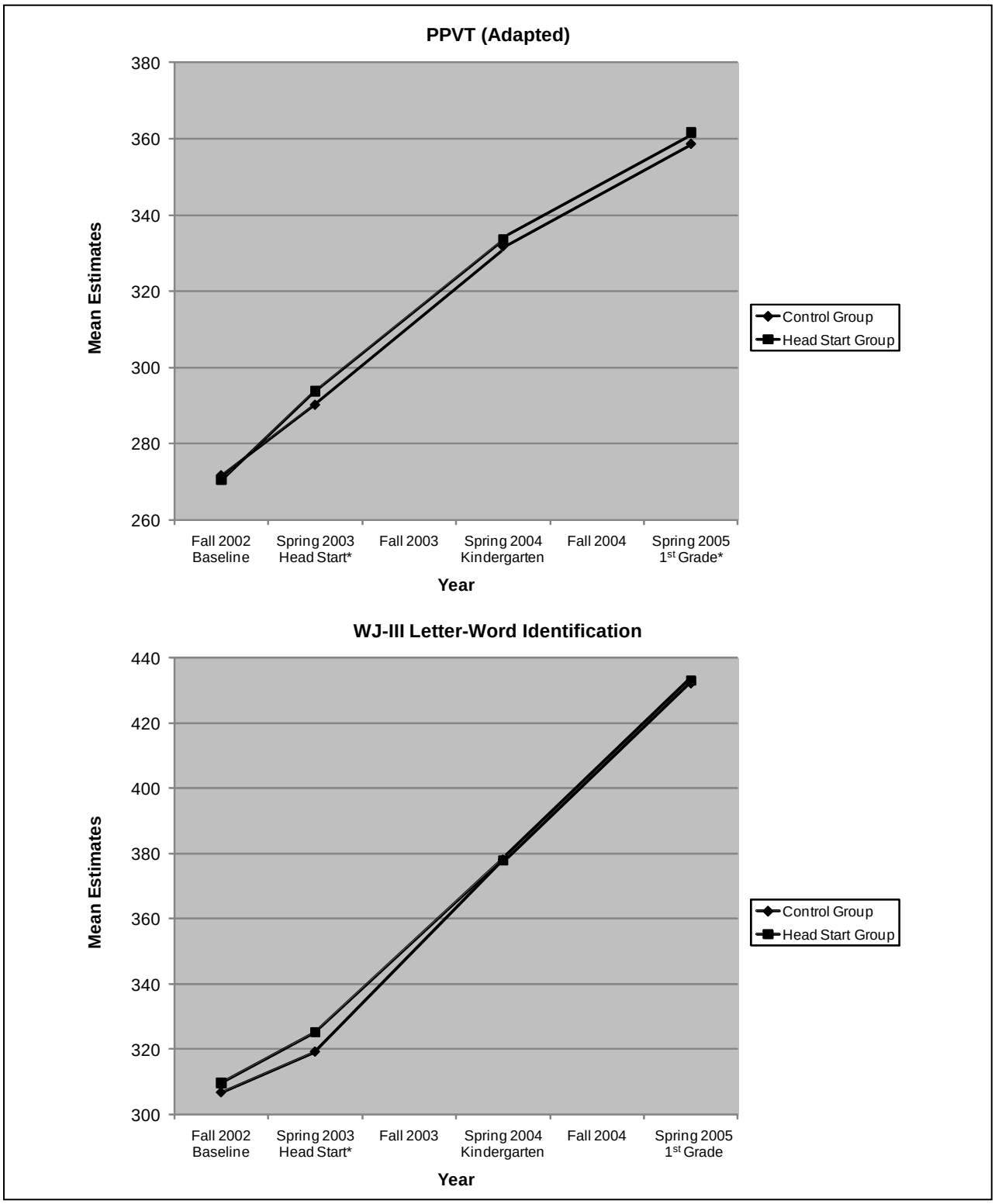
As shown in Exhibit 4.2, no statistically significant impacts were found on any of these measures for children in the 4-year-old cohort. This is true in all domains, through the school years; thus, it is congruent with the lack of statistically significant impacts from the direct child assessments for the 4-year-old cohort in the school years.

Summary of Annual Impacts for the 4-Year-Old Cohort

Based on these annual impact estimates, it appears that access to Head Start has an impact on 4-year-olds' language and literacy skills while they are in Head Start, but these early gains are not sustained as the children develop and move into the early school years. Furthermore, there is no evidence of impacts on children's math ability, pre-writing skills, or teacher assessments at the end of Head Start, at the end of kindergarten, or at the end of 1st grade. In other words, the children in the Head Start group ended their Head Start year with moderately higher skills than their counterparts in the control group, but this advantage did not lead to longer term gains when they were in school. At the end of 1st grade, they end up at the same point as the children who were not given access to the program. Although both groups of children are making progress over time, in most instances, the Head Start group scores are not statistically different from the control group scores in kindergarten and 1st grade.

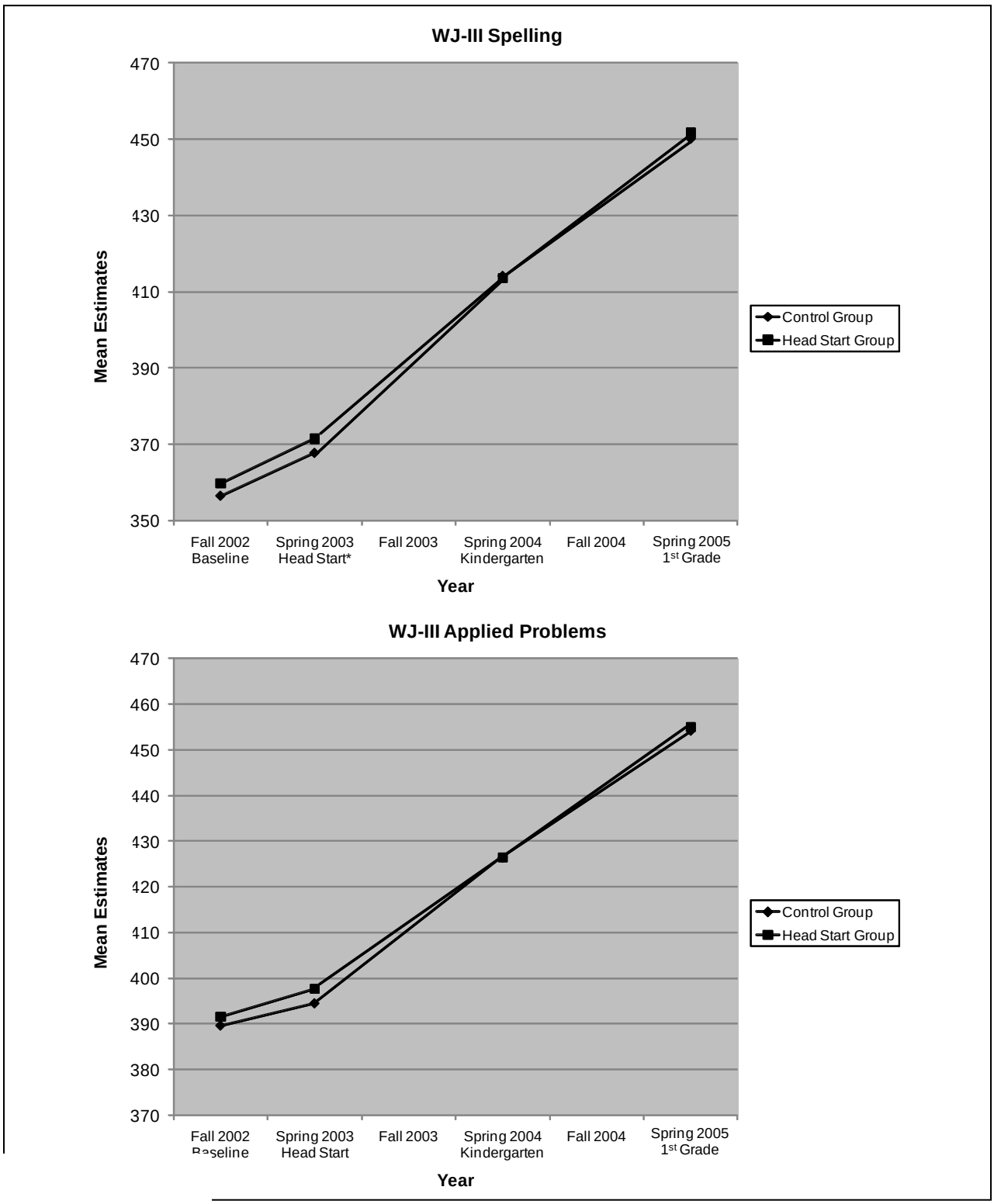
This pattern over time is illustrated in Exhibit 4.3, which graphs the control group means and the regression-adjusted Head Start group means by year for four child assessments: the PPVT (Adapted), and three tests from the Woodcock-Johnson III, Letter-Word Identification, Spelling, and Applied Problems (all four were administered at all waves of data collection, and the first three showed impacts at the end of Head Start). The graphs demonstrate little difference for the outcome means at baseline (i.e., fall 2002). For two literacy measures (i.e., Letter-Word Identification and Spelling), there is a consistent pattern of initial program-related differences at the end of the Head Start year, followed by a closing of the early gains during early elementary school so that the two groups of children are essentially indistinguishable in their measured outcomes by the end of kindergarten and again at the end of 1st grade. For the PPVT (Adapted), the graph illustrates a difference at the end of the Head Start year and again at the end of 1st grade. For the Woodcock-Johnson III Applied Problems, for which no impacts were found, the two groups are essentially indistinguishable in their measured outcomes from the end of

Exhibit 4.3: Estimated Annual Impacts on Cognitive Outcomes, by Year, 4-Year-Old Cohort



Key:
*Indicates a significant difference.

Exhibit 4.3: Estimated Annual Impacts on Cognitive Outcomes, by Year, 4-Year-Old Cohort (continued)



ifference.

kindergarten through the end of 1st grade. Although the graph displays a slight difference at the end of the Head Start year, the difference is not statistically significant.

Longitudinal Analyses for the 4-Year-Old Cohort

The estimated annual impacts presented thus far are based on all of the children for whom data were obtained at a particular wave of data collection. To get a better understanding of the developmental impact of Head Start, separate longitudinal or "growth" analyses were conducted using Hierarchical Linear Models (HLM) for the sample of children for whom two or more data points were available. This technique, described in Chapter 2 and the Technical Report for the Head Start Impact Study, captures individual children's growth trajectories over time[83] and differences in growth among children with different characteristics (including assignment to either the Head Start group or control group).

A linear growth model was selected to allow for the inclusion of as many children as possible in the longitudinal analysis. The 4-year-old cohort has four potential data collection points, a sufficient number of points for an accurate estimate of quadratic growth. However, the sample size would be greatly reduced if the sample was limited to children who have all four data collection points.

The results of these analyses (see Exhibit 4.4), show the estimates of the effect of access to Head Start on 4-year-olds' growth for the five cognitive outcomes that were consistently measured at all waves of data collection—the PPVT (Adapted), and four assessments from the Woodcock-Johnson III, i.e., Letter-Word Identification, Spelling, Applied Problems, and Pre-Academic Skills.

For the four Woodcock-Johnson III assessments, there were no significant differences between the Head Start and control groups in terms of their annual linear growth rates over the four waves of data collection, indicating that providing access to Head Start does not reliably alter children's growth trajectories from the Head Start year through the end of 1st grade. For the PPVT, there is a small, but statistically significant difference in annual growth rates over the same time period, suggesting there was more rapid growth for children in the Head Start group than for children in the control group, on average, over the three years. It is helpful to look back

[83] These analyses use child's age as the measure of time.

Exhibit 4.4: Estimated Impacts on Baseline English-Speaking Children's Growth for Longitudinally Measured Cognitive Outcomes, 4-Year-Old Cohort

Regression-Adjusted Estimated Effect on Linear Growth Between Fall 2002 and Spring 2005: 4-Year-Old Cohort (Four Waves of Data Collection)				
Outcome	Average Annual Growth: Head Start Group	Average Annual Growth: Control Group	Estimated Effect	p-value
PPVT (Adapted)	30.79	29.47	1.32**	0.01
WJ III Pre-Academic Skills	32.81	33.43	-0.62	0.33
WJ III Spelling	32.96	33.51	-0.55	0.40
WJ III Letter-Word Identification	40.49	40.99	-0.50	0.63
WJ III Applied Problems	23.62	23.53	0.09	0.84

Key:
*** $p \leq 0.01$
** $p \leq 0.05$
* $p \leq 0.10$

Note: Time measure is wave of data collection.

at the cross-sectional findings to better understand these impacts on growth trajectories. As the cross-sectional annual analyses show, the difference in growth was found primarily at the end of the Head Start year. However, there was a significant difference ($p \leq 0.10$) on the PPVT (Adapted) at the end of 1st grade.

Impacts for the 3-Year-Old Cohort

The interpretation of the estimated impacts on the 4-year-old cohort is very straightforward. Impacts observed at the end of Head Start (spring 2003) represent the effect of access to Head Start when the children were in Head Start in the year prior to entry into elementary school. The two subsequent waves—spring 2004 and spring 2005—can be thought of as the consequences of any Head Start benefits provided by access to Head Start, observed at the end of kindergarten and 1st grade, respectively. In other words, the later impacts represent the subsequent effect of a Head Start intervention received the year before the children entered school.

The situation is quite a bit different for the 3-year-old cohort because of the nature of the study design. As discussed in Chapter 2, children who were new applicants to Head Start when they were three (fall 2002) were randomized into two groups in a manner similar to that of the 4-year-old cohort; one group was allowed entry into Head Start (the Head Start group), while the

other group was not allowed to enter Head Start (the control group). Consequently, impacts estimated at the end of that first year (spring 2003) represent the effect of access to Head Start when the children were in Head Start two years prior to entry into elementary school. This is similar to the interpretation of the Head Start year impacts for the 4-year-old children, but represents the effect when services are provided a year earlier.

Where the study design differs for the 3-year-old cohort is that these children had another year to go before they started kindergarten.[84] Because Head Start programs had concerns about denying children access to Head Start for two years, children in both the randomly assigned Head Start group and the control group were allowed to re-apply to Head Start the following year when they turned four (fall 2003), and about 60 percent opted to do so. Because some in the Head Start group never attended Head Start, about 57 percent of children randomly assigned to the Head Start group participated for two years. Of the control group, about 46 percent enrolled in Head Start when they turned four. Because some control group children were able to find their way into Head Start in the first year, about 11 percent of the control group was enrolled in Head Start for two years.

As a result, impacts after the first year for children in the 3-year-old group represent the effect of gaining access to Head Start at age three versus not gaining access until age four, regardless of what these children did at age four, (e.g., enroll in another year of Head Start, go to another Head Start or preschool program, or remain at home or with a child care provider) or any subsequent year.

Annual Impacts on Language and Literacy Measures

As shown in Exhibit 4.5,[85] at the end of their Head Start year (spring 2003), strong evidence of program impacts was found on five child assessment outcomes related to children's

---

[84] In a sense, one can think of the 4-year-old cohort as "delayed-entry" eligible 3-year-olds, i.e., parents who had the option to enroll their child at age three but opted instead to "hold them back" until age four. Not surprisingly, as discussed elsewhere, the parents who elected to seek enrollment for their child at age three are different from those who decided to wait until age four, and our data bear this out. For example, the 3-year-old cohort of newly entering children is more African American, and the 4-year-old cohort of newly entering children is more Hispanic.

[85] Detailed tables for the impact estimates, including sample sizes, standard errors, and confidence intervals, can be found on the Administration for Children and Families, Office of Planning, Research and Evaluation Website at: http://www.acf.hhs.gov/programs/opre/hs/impact_study/index.html.

Exhibit 4.5: Estimated Impacts on Cognitive Outcomes by Year: 3-Year-Old Cohort

Outcomes	Mean Estimates				Regression-Adjusted Impact		Effect Size
	Head Start Group	Control Group	Head Start - Control	p-value	Impact	p-value	
Head Start Baseline (Fall 2002)							
Language and Literacy Measures++							
PPVT (Adapted)	230.01	230.49					
WJ III Letter-Word Identification	294.70	293.38					
WJ III Spelling	334.81	333.54					
WJ III Oral Comprehension	433.32	433.72					
CTOPPP Elision	235.14	230.24					
Color Identification	0.29	0.27					
WJ III Pre-Academic Skills	336.81	335.88					
Spanish Language and Literacy Measures+++							
TVIP (Adapted)	223.02	236.01					
WM Letter-Word Identification	351.17	347.06					
Pre-writing Measure++							
McCarthy Draw-a-Design	2.72	2.70					
Parent-Reported Literacy Measure							
Emergent Literacy Scale+	2.05	1.93					
Math Skills Measures++							
WJ III Applied Problems	367.31	365.05					
Counting Bears	0.15	0.17					
Head Start Year (Spring 2003)							
Language and Literacy Measures							
PPVT (Adapted)	257.50	251.43	6.07	0.003	**6.53*** **	0.000	0.18
WJ III Letter-Word Identification	307.00	300.51	6.49	0.001	**6.14*** **	0.000	0.26
WJ III Spelling	346.57	343.64	2.93	0.061	2.28	0.130	0.10
WJ III Oral Comprehension	435.52	435.44	0.09	0.924	0.28	0.698	0.02
CTOPPP Elision	241.44	235.03	6.41	0.078	**5.01***	0.061	0.10
Color Identification	0.49	0.46	0.03	0.385	0.04	0.179	0.07
Letter Naming	5.49	3.92	1.57	0.010	**1.56*** **	0.005	0.24
WJ III Pre-Academic Skills	343.67	339.41	4.26	0.013	**4.25*** **	0.004	0.22

Exhibit 4.5: Estimated Impacts on Cognitive Outcomes by Year: 3-Year-Old Cohort (continued)

Outcomes	Mean Estimates				Regression-Adjusted Impact		Effect Size
	Head Start Group	Control Group	Head Start - Control	p-value	Impact	p-value	
Spanish Language and Literacy Measures+++							
TVIP (Adapted)	256.83	247.05	9.79	0.069	5.21	0.365	0.13
WM Letter-Word Identification	353.78	351.56	2.23	0.264	1.59	0.380	0.13
Pre-writing Measure							
McCarthy Draw-a-Design	3.23	3.05	0.18	0.005	**0.16*** **	0.007	0.14
Parent-Reported Literacy Measure							
Emergent Literacy Scale+	2.86	2.35	0.51	0.000	**0.48*** **	0.000	0.35
Math Skills Measures							
WJ III Applied Problems	377.27	373.57	3.69	0.144	4.31**	0.012	0.15
Counting Bears	0.30	0.28	0.02	0.362	0.03	0.241	0.06
Age 4 Year (Spring 2004)							
Language and Literacy Measures							
PPVT (Adapted)	300.99	298.28	2.72	0.177	2.03	0.251	0.05
WJ III Letter-Word Identification	333.08	330.13	2.95	0.112	2.56	0.112	0.09
WJ III Spelling	376.74	376.26	0.47	0.794	0.28	0.875	0.01
WJ III Oral Comprehension	446.19	445.80	0.38	0.662	0.25	0.743	0.02
CTOPPP Elision	281.07	271.90	9.17	0.004	**8.26*** **	0.002	0.15
Color Identification	0.81	0.80	0.02	0.432	0.01	0.466	0.03
Letter Naming	13.73	12.84	0.89	0.144	0.85	0.155	0.09
WJ III Pre-Academic Skills	370.24	368.72	1.53	0.300	1.24	0.378	0.06
Spanish Language and Literacy Measures+++							
TVIP (Adapted)	292.51	292.45	0.07	0.990	-1.33	0.803	-0.03
WM Letter-Word Identification	365.63	361.99	3.63	0.231	3.05	0.334	0.16
Pre-writing Measure							
McCarthy Draw-a-Design	4.83	4.94	-0.11	0.425	-0.09	0.482	-0.04
Parent-Reported Literacy Measure							
Emergent Literacy Scale	4.01	3.82	0.19	0.005	**0.20*** **	0.002	0.16

Exhibit 4.5: Estimated Impacts on Cognitive Outcomes by Year: 3-Year-Old Cohort (continued)

Outcomes	Mean Estimates				Regression-Adjusted Impact		Effect Size
	Head Start Group	Control Group	Head Start - Control	p-value	Impact	p-value	
Math Skills Measures							
WJ III Applied Problems	401.06	399.86	1.20	0.381	0.75	0.551	0.03
Counting Bears	0.56	0.55	0.01	0.879	0.01	0.777	0.02
Kindergarten Year (Spring 2005)							
Language and Literacy Measures							
PPVT (Adapted)	340.31	339.91	0.40	0.783	0.26	0.851	0.01
WJ III Letter-Word Identification	384.02	383.42	0.61	0.794	0.24	0.899	0.01
WJ III Spelling	420.20	419.43	0.77	0.670	0.45	0.774	0.02
WJ III Oral Comprehension	457.91	457.36	0.55	0.648	0.50	0.633	0.03
CTOPPP Elision	331.33	335.07	-3.74	0.220	-3.52	0.241	-0.08
Letter Naming	23.46	23.67	-0.21	0.580	-0.32	0.340	-0.06
WJ III Pre-Academic Skills	411.62	411.39	0.22	0.898	-0.02	0.988	0.00
WJ III Word Attack	436.00	437.37	-1.38	0.566	-1.37	0.563	-0.04
WJ III Basic Reading Skills	410.05	410.38	-0.33	0.883	-0.54	0.801	-0.02
Spanish Language and Literacy Measures+++							
TVIP (Adapted)	353.10	358.74	-5.64	0.210	-7.51	0.117	-0.19
WM Letter-Word Identification	393.41	382.28	11.13	0.009	8.73*	0.053	0.26
Math Skills Measures							
WJ III Applied Problems	430.39	431.29	-0.90	0.595	-0.94	0.519	-0.04
WJ III Quantitative Concepts	442.98	443.71	-0.73	0.497	-0.88	0.310	-0.05
WJ III Math Reasoning	436.69	437.50	-0.81	0.548	-0.91	0.408	-0.05
School Performance Assessment Measures							
School Accomplishments	27.57	28.32	-0.75	0.172	-0.65	0.203	-0.09
Promotion	0.89	0.90	-0.01	0.772	-0.01	0.709	-0.03
Language and Literacy Ability	0.75	0.79	-0.04	0.103	-0.04	0.127	-0.09
Math Ability	0.79	0.86	-0.07	0.004	**-0.07*****	0.003	-0.19
Social Studies and Science Ability	0.84	0.87	-0.03	0.171	-0.03	0.121	-0.10

Exhibit 4.5: Estimated Impacts on Cognitive Outcomes by Year: 3-Year-Old Cohort (continued)

Outcomes	Mean Estimates				Regression-Adjusted Impact		Effect Size
	Head Start Group	Control Group	Head Start - Control	p-value	Impact	p-value	
1st Grade Year (Spring 2006)							
Language and Literacy Measures							
PPVT (Adapted)	360.41	357.91	2.50	0.280	2.32	0.151	0.08
WJ III Letter-Word Identification	433.69	432.92	0.78	0.729	0.37	0.848	0.01
WJ III Spelling	453.89	454.94	-1.04	0.548	-1.20	0.438	-0.05
WJ III Oral Comprehension	472.60	471.25	1.36	0.118	1.35*	0.051	0.08
WJ III Pre-Academic Skills	447.53	447.17	0.36	0.830	0.24	0.869	0.01
WJ III Word Attack	468.84	469.12	-0.28	0.890	-0.60	0.759	-0.02
WJ III Basic Reading Skills	451.29	450.97	0.32	0.877	-0.08	0.966	0.00
WJ III Academic Applications	463.05	462.29	0.76	0.524	0.73	0.489	0.04
WJ III Academic Skills	449.89	450.19	-0.30	0.844	-0.60	0.633	-0.03
WJ III Passage Comprehension	451.10	450.18	0.92	0.549	0.76	0.580	0.03
WJ III Writing Sample+	483.03	483.04	-0.01	0.990	-0.09	0.928	-0.01
Spanish Language and Literacy Measures+++							
TVIP (Adapted)	376.07	374.41	1.66	0.735	0.04	0.993	0.00
WM Letter-Word Identification	416.96	418.67	-1.71	0.739	-0.54	0.910	-0.01
Math Skills Measures							
WJ III Applied Problems	455.07	453.67	1.40	0.292	1.59	0.161	0.08
WJ III Quantitative Concepts	462.01	461.36	0.64	0.615	0.79	0.450	0.05
WJ III Math Reasoning	458.49	457.45	1.04	0.405	1.20	0.231	0.07
WJ III Calculation	461.81	461.65	0.15	0.872	-0.02	0.977	0.00

Exhibit 4.5: Estimated Impacts on Cognitive Outcomes by Year: 3-Year-Old Cohort (continued)

Outcomes	Mean Estimates				Regression-Adjusted Impact		Effect Size
	Head Start Group	Control Group	Head Start - Control	p-value	Impact	p-value	
School Performance Assessment Measures							
School Accomplishments	42.46	42.74	-0.28	0.728	-0.29	0.705	-0.03
Promotion	0.92	0.93	-0.01	0.480	-0.02	0.248	-0.07
Language and Literacy Ability	0.72	0.71	0.00	0.920	0.00	0.888	0.01
Math Ability	0.78	0.80	-0.02	0.450	-0.02	0.448	-0.05
Social Studies and Science Ability	0.83	0.86	-0.03	0.198	-0.03	0.286	-0.07

Key:
******* $p \leq 0.01$
** $p \leq 0.05$
* $p \leq 0.10$

+ Indicates the reliability (Cronbach alpha) for the measure is <0.6.

++ Indicates baseline scores for English-speaking children only except for the PPVT and Woodcock-Johnson III Letter-Word test.

+++ Indicates scores for only the Dual Language Learners on the mainland.

Bold regression-adjusted impact value indicates the outcome passes the Benjamini-Hochberg test for multiple comparisons with a 10 percent false discovery rate.

NOTE: Some cognitive findings in this report may differ from the findings reported in the *Head Start Impact Study: First Year Findings* (U.S. Department of Health and Human Services, 2005) due to redefining of variables (i.e., Color Identification and Counting Bears) and residualization. See Chapter 2 and the Technical Report for the Head Start Impact Study for details of these changes.

language and literacy development: PPVT (Adapted) (effect size=0.18); Woodcock-Johnson III Letter-Word Identification (effect size=0.26); CTOPPP Elision (effect size=0.10); Letter Naming (effect size=0.24); and the Woodcock-Johnson III Pre-Academic Skills (effect size=0.22).

As also shown in Exhibit 4.5, at the end of the age 4 year (spring 2004), strong evidence of impacts was found again but only for CTOPPP Elision (effect size=0.15). For the early school period, there is suggestive evidence of an impact on the Woodcock-Johnson III Oral Comprehension test at the end of 1st grade (effect size=0.08). Additionally, no statistically significant impacts were identified on children's language and literacy ability as reported by teachers during the children's kindergarten and 1st grade years.

Annual Impact on the Pre-writing Measure

There is strong evidence of an impact on children's pre-writing skills at the end of the Head Start year (effect size=0.14), but there is no evidence of an effect at the end of the age 4 year. This measure was not administered in subsequent years.

Annual Impact on Parent-Reported Literacy Measure

There is strong evidence of an impact on the parent-reported literacy measure both at the end of the Head Start and age 4 years (effect sizes=0.35 and 0.16, respectively). This measure was not administered in subsequent years.

Annual Impacts on Spanish Language and Literacy Measures

There was no strong evidence of an impact of access to Head Start on either of the Spanish language tests in any year. However, there is suggestive evidence of an impact on the Batería Woodcock-Muñoz Identificación de letras y palabras at the end of kindergarten (effect size=0.26), with the Head Start group showing moderately higher scores than the control group.

Annual Impacts on Math Skills Measures

Unlike the results for children in the 4-year-old cohort, at the end of the Head Start year, there is moderate evidence of an impact on children's math skills as measured by the Woodcock-Johnson III Applied Problems (effect size=0.15) for children in the 3-year-old cohort. No subsequent impacts were found at the end of children's age 4 year, at the end of kindergarten, or at the end of 1st grade on any of the direct child assessments. Additionally, a significant impact supported by strong evidence was reported by teachers on children's math ability in an unexpected direction. Kindergarten teachers reported poorer math skills for children in the Head Start group than for those in the control group (effect size=-0.19).

Annual Impacts on School Performance Measures

Additional parent reports and teacher assessments were obtained for each child when he/she was in kindergarten and 1st grade. These included teacher assessments of the child's overall school accomplishments, teacher assessment of the child's ability in social studies and

science, and parents' report of their child's promotion to the next grade. No significant impacts were found for these ratings for children in the 3-year-old cohort in kindergarten or 1st grade.

Summary of Annual Impacts for the 3-Year-Old Cohort

Like their counterparts in the 4-year-old cohort, the annual impact estimates for children in the 3-year-old cohort suggest that access to an early year of Head Start at age three has impacts on children's cognitive development while they are in their Head Start year. However, impacts were very few in the age 4 year and were not followed by longer term benefits for children in their early school years.

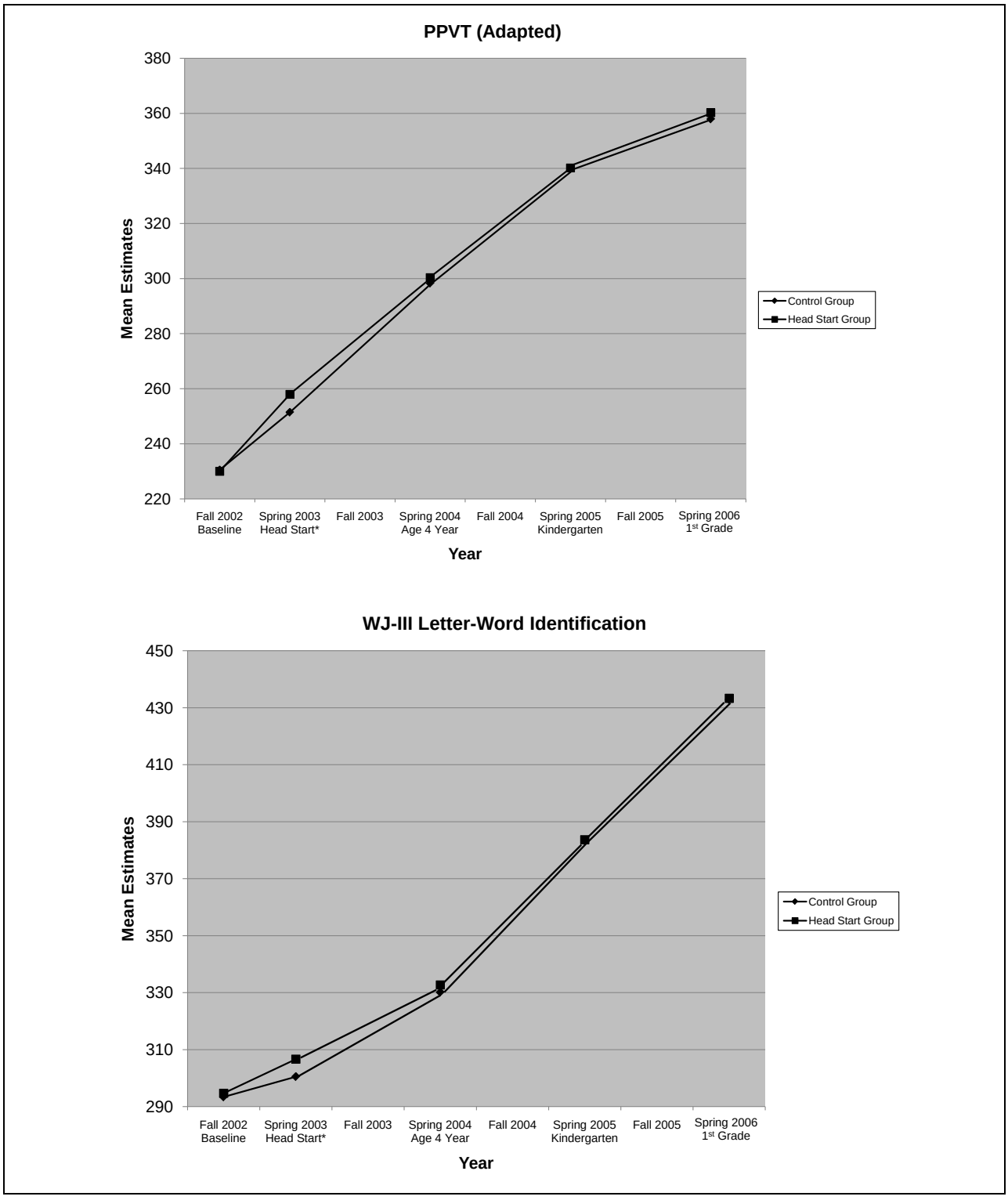
These patterns are again illustrated in graphs of the control group means and the regression-adjusted Head Start group means, by year for four child assessments that were administered longitudinally: the PPVT (Adapted), and three tests from the Woodcock-Johnson III: Letter-Word Identification, Spelling, and Applied Problems. As shown in Exhibit 4.6, there is a consistent pattern of initial program-related differences between the Head Start group and the control group at the end of the Head Start years on three of the four outcomes. This initial difference is followed by a closing of the early gains during the age 4 year and early elementary school, so that the two groups of children are essentially indistinguishable in their measured outcomes by the end of 1st grade.

Longitudinal Analyses for the 3-Year-Old Cohort

Exhibit 4.7 provides the estimated impact of access to Head Start on children's growth for the same five cognitive outcomes discussed earlier for the 4-year-old cohort. As shown, there were no significant differences ($p \leq 0.05$) in the linear growth rates over time between the Head Start group and the control group members over the five waves of observations.

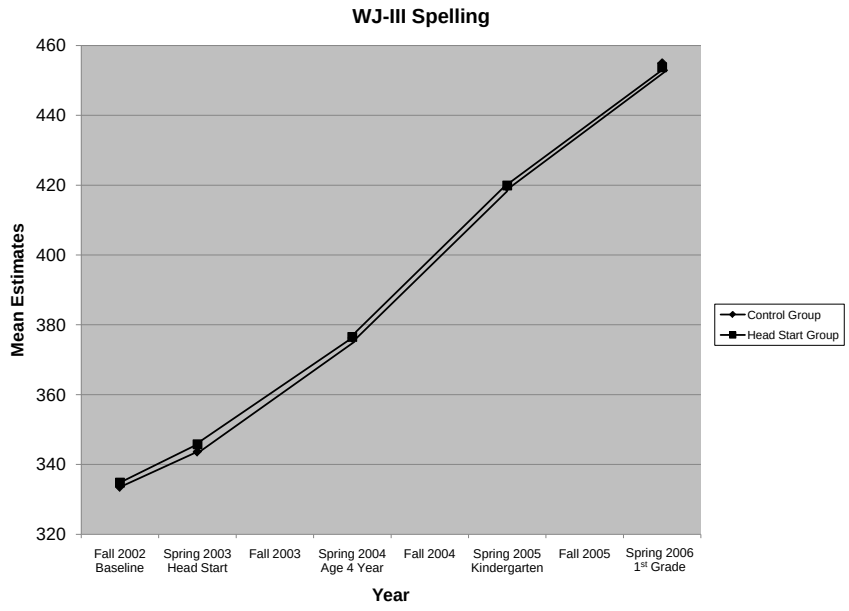
An examination of graphical plots of the annual impacts (see Exhibit 4.6), shows that the Head Start and control groups start out together at baseline, which is what we would expect since they were randomly assigned to the two groups. But, for example, at the end of the Head Start year (when the Head Start group children were allowed access to Head Start), the Head Start group performed better on the Woodcock-Johnson III Letter-Word Identification than the

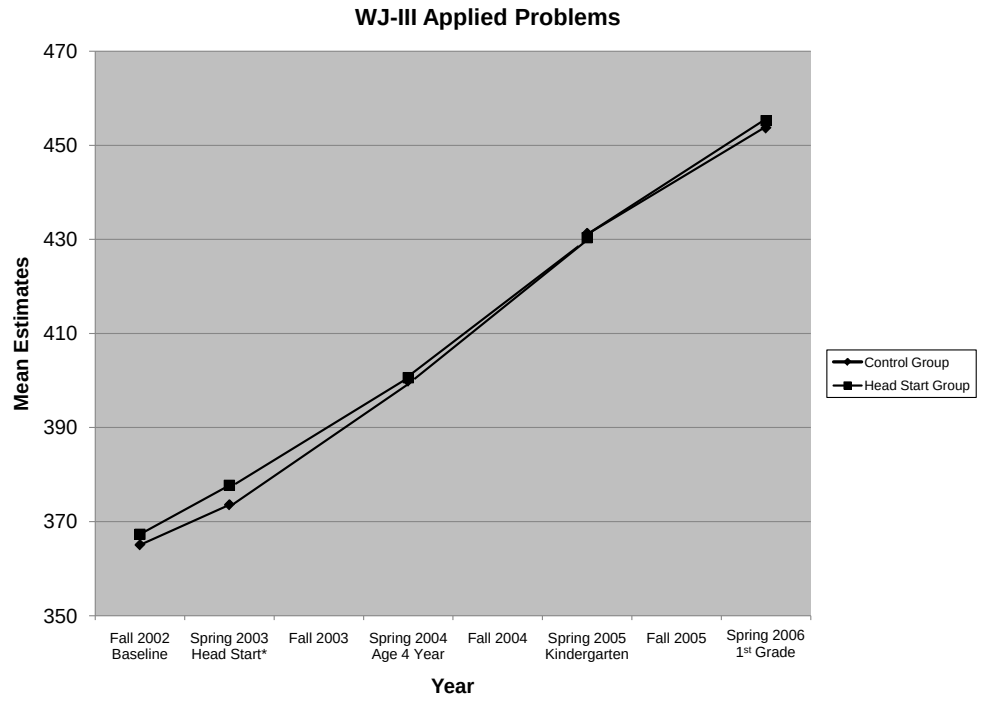
Exhibit 4.6: Estimated Annual Impacts on Cognitive Outcomes, by Year, 3-Year-Old Cohort



Key:
*Indicates a significant difference.

Exhibit 4.6: Estimated Annual Impacts on Cognitive Outcomes, by Year, 3-Year-Old Cohort (continued)





Key:
*Indicates a significant difference.

Exhibit 4.7: Estimated Impacts on Baseline English-Speaking Children's Growth for Selected Cognitive Outcomes, 3-Year-Old Cohort

Regression-Adjusted Estimated Effect on Linear Growth Between Fall 2002 and Spring 2006: 3-Year-Old Cohort (Five Waves of Data Collection)				
Outcomes	Average Annual Growth: Head Start Group	Average Annual Growth: Control Group	Estimated Effect	p-value
PPVT (Adapted)	34.48	34.38	0.10	0.80
WJ III Pre-Academic Skills	30.51	30.98	-0.47	0.14
WJ III Spelling	31.77	32.29	-0.52*	0.08
WJ III Letter-Word Identification	34.92	35.21	-0.29	0.50
WJ III Applied Problems	23.94	24.21	-0.27	0.43

Key:
*** $p \leq 0.01$
** $p \leq 0.05$
* $p \leq 0.10$
Note: Time measure is wave of data collection.

control group. By the end of their age 4 year, the two groups drew close together but the Head Start group was still slightly higher than the control group. By the end of kindergarten and 1st grade, however, the groups were indistinguishable on average performance. Thus, although access to Head Start led to improvements in 3-year-olds' performance on several assessments at the end of the Head Start year, access to Head Start did not alter children's growth trajectories, when averaged over the four year period.

For Woodcock-Johnson III Spelling, there was a significant difference at the 0.10 level, indicating a negative effect for treatment. This is somewhat puzzling given the observed, positive annual cross-sectional differences found during the Head Start year (see Exhibit 4.5). These modest annual differences may be masked, however, in a model that estimates longitudinal outcome levels and impact as part of a linear trajectory.[86] In other words, since the difference in growth in Exhibit 4.7 is over all five waves of data, the initial modest difference in growth at the end of the Head Start year combined with a lack of difference after the age 4 year, amounts to a non-significant difference in growth overall. Also, since the Head Start group mean for the Woodcock-Johnson III Spelling measure was slightly larger than that of the control group prior to Head Start, when the groups came together at wave five, it resulted in a small, yet

[86] To keep the analysis similar for both the 3-year-old and 4-year-old cohort, a linear growth model was again used.

significant negative effect for treatment. However, the graph of covariate-controlled means in Exhibit 4.6 shows that this negative effect is very slight.

Although the developmental growth trajectories of the Head Start group and control group have almost identical slopes, they drift closer together with time so that it is possible for a significant early treatment effect to accompany a non-statistically significant later effect. Along the way, however, there is no statistically significant effect of treatment on the rate of growth over time.

Impacts on Participants

This section discusses Head Start's impact on the cognitive development of children who *actually participated in the program.*[87] Moving from measures of the impact of access to Head Start (ITT estimates) to the impact of participating (IOT estimates) entails scaling the estimates by a constant factor, derived for each cohort using methods discussed in Chapter 2 and the Technical Report for the Head Start Impact Study. There is no change in the statistical significance of the estimates. For the 4-year-old cohort, this factor is about 1.5. For example, the impact of access to Head Start on the PPVT (Adapted) scores of 4-year-olds at the end of their Head Start year (spring 2003) is 3.55 (ITT estimate in Exhibit 4.2), while the impact of participating in Head Start is 5.31 (IOT estimate in Exhibit E4.2 in Appendix E). For the 3-year-old cohort, the factor is about 1.4. For example, the impact of access to Head Start on the PPVT (Adapted) is 6.53 points in spring 2003 (ITT estimate in Exhibit 4.5), while the impact of participating in Head Start is 9.36 (IOT estimate in Exhibit E4.5 in Appendix E). Effect sizes can help in interpreting the differences in magnitude between the impact of access to Head Start and the impact of participation. For example, among impacts for the two cohorts confirmed by strong evidence (found primarily in the first year of follow-up) effect sizes range from 0.13 to 0.50 for the impact of participation, in contrast to effect sizes of 0.09 to 0.35 for the impact of access.

Looking at effects on *participants* does not change the overall patterns found in the main analysis, which show that Head Start improved children's language and literacy development during the program year but not later and had only one strongly confirmed impact on math

[87] Detailed findings are presented in Appendix E. As noted earlier, some children granted access never participated in Head Start while other children randomized into the study's control group did participate.

ability in a negative direction. (For the 3-year-old cohort, kindergarten teachers reported poorer math skills for children in the Head Start group than children in the control group.)

Summary

Based on the annual impact estimates presented in this chapter, it appears that access to Head Start has an impact on 4-year-olds' language and literacy skills while they are in Head Start, but these early gains are not sustained as the children develop and move into the early school years. There is no evidence of impacts on children's math ability during the Head Start, kindergarten, or 1st grade years for these children. Like their counterparts in the 4-year-old cohort, the annual impact estimates for children in the 3-year-old cohort suggest that access to an early year of Head Start at age three has an impact on children's cognitive development while they are in their Head Start year. However, these impacts were not followed by longer term benefits for children as they develop and move into their age 4 year and early school years. Although both the Head Start and control groups of children are making progress over time, in most instances, the Head Start group scores are not statistically different from the control group scores after that first year.

Regardless of any impact of Head Start, it is helpful to understand how the study children fall in terms of their cognitive and academic skills. Comparing the skill levels of children in the Head Start Impact Study with those of the general population of 3- and 4-year-olds in the United States (including those who were not from low-income families) on the PPVT showed that the mean performance of study children was below the average performance level for all U.S. children. For example, the average 2003 PPVT score for a child in the 4-year-old Head Start group is at the 31st percentile, while the average score for a child in the 3-year-old Head Start group is at the 32nd percentile. The highest percentile reported for the average PPVT score across the years is the 36th percentile with the Head Start group percentile usually equal to or greater than the control group percentile.

The study children also lag behind other children in the nation on letter identification. For the 4-year-old cohort, 55 percent of the Head Start group can recognize all their letters by the end of their kindergarten year. For the 3-year-old cohort, 65 percent of the Head Start group can recognize all their letters. Comparing these numbers to a nationally representative sample of children from the Early Childhood Longitudinal Study-Kindergarten Cohort (ECLS-K) (U.S.

Department of Education, 2002) reported that by the spring of kindergarten, 95 percent of children know the letters of the alphabet by the end of their kindergarten year.

However, not all the measures suggest that these children are below average. On some of the Woodcock-Johnson III tests (i.e., Quantitative Concepts, Applied Problems, Math Reasoning, Passage Comprehension and Oral Comprehension) children scored below the 50th percentile at the end of 1st grade. However, they scored at or above the 50th percentile on several other tests, including Letter-Word Identification, Spelling, Word Attack and Calculation. This is in contradiction with the other norms presented here, sometimes even within the same skill set (e.g., ECLS-K letter naming measure and the W-J III Letter-Word Identification). Hence, it is unclear whether these norms suggest that this group of children from low-income families is indeed fairing as would be expected of other children their age. A complete listing of subscale norms for both cohorts and all study years is provided in the Technical Report for the Head Start Impact Study.

Chapter 5: The Impact of Head Start on Children's Social-Emotional Development

Introduction

This chapter focuses on the impact of access to and participation in Head Start on children's early social-emotional development as rated by their parents and teachers. In the spring of each year, parent/primary caregivers rated their child on: (1) total problem behavior, including aggressive, hyperactive, and withdrawn behaviors; (2) social competencies; (3) social skills and approaches to learning; and (4) the child's relationship with his/her parent. Kindergarten and 1st grade teachers rated children on: (1) aggressive, inattentive/hyperactive, low energy, oppositional, and shy/social reticence behaviors; (2) problems with peer interactions, teacher interactions, and structured learning situations; and (3) his or her relationship with the child. Because many control group children were in parent care during the Head Start years, teacher ratings are not available for the full sample of study children during this time and were, therefore, examined only when the children were in kindergarten and 1st grade.

Measures

A brief description of the social-emotional measures used in this study is provided below.[88]

- **Social Skills and Positive Approaches to Learning** (Achenbach et al., 1987). Parents were asked to rate their child's social skills and positive approaches to learning. Social skills focused on cooperative and empathic behavior, and the positive approaches to learning scale addressed curiosity, imagination, openness to new tasks and challenges, and having a positive attitude about gaining new knowledge and skills.
- **Social Competencies** (Developing Skills Checklist, 1990). Parents were asked to rate their child's social competencies, such as interactions with adults and responsibility for personal belongings.
- **Total Problem Behavior** (Achenbach et al., 1987). Parents were asked to rate their child's total problem behavior. Three subscales are included in this measure: aggressive behavior, hyperactive behavior, and withdrawn behavior.

[88] See Chapter 2 and the Technical Report for the Head Start Impact Study for detailed information on the social-emotional measures.

- **Child-Parent Relationship Scale** (Pianta, 1992). Parents were asked to rate their child's relationship with them. Two subscales can be measured: closeness and conflict, as well as a summary scale indicating the overall positive relationship.
- **Student-Teacher Relationship Scale** (Pianta, 1996). Teachers were asked to rate their relationship with each study child in their class. The subscales were the same as reported above for the Child-Parent Relationship Scale.
- **ASPI (Adjustment Scales for Pre-school Intervention)** (Lutz et al., 2000). The ASPI is designed for use by teachers to measure emotional and behavioral adjustment within usual classroom situations. Five behavioral dimensions can be measured using the ASPI: aggressive, withdrawn/low energy, shy/socially reticent, oppositional, and inattentive/hyperactive. Three situational dimensions also can be measured using the ASPI: problems with structured learning, problems with peer interaction, and problems with teacher interaction.

Many measures in the social-emotional domain are scored in such a way that a lower number indicates better functioning. For example, fewer behavior problems are preferable to more behavior problems. In contrast, measures of positive behaviors and skills—such as Social Competencies—are scored such that higher scores indicate better functioning. A higher value indicates better functioning, and a positive impact is desirable on four social-emotional measures reported in this chapter (i.e., social competencies, social skills and positive approaches to learning, closeness, and positive relationships). For the remaining measures, (i.e., total problem behavior, aggressive behavior, hyperactive behavior, withdrawn behavior, conflict, and all ASPI measures), a lower value indicates better functioning, and a negative impact is desirable.

The first set of impact estimates discussed in this chapter measures the effect of access to Head Start on the average child randomly assigned to the Head Start group (the intent to treat or ITT estimates), while the end of the chapter provides a discussion of the impact of Head Start on the children who actually participated in the program (i.e., the impact on the treated or the IOT estimates). (See Chapter 2 and the Technical Report for the Head Start Impact Study for a discussion of the methodology used for the impact estimates.)

Impacts for the 4-Year-Old Cohort

As shown in Exhibit 5.1,[89] there is no evidence of impacts from providing access to Head Start on any of the social-emotional development measures at the end of Head Start or at the end of kindergarten for the 4-year-old cohort. However, at the end of 1st grade, there is moderate evidence[90] of an impact on teacher reports of children's shy/socially reticent behavior in school. Teachers reported that Head Start group children were more shy or socially reticent than the control group children (effect size=0.19). Teachers rated six percent of the Head Start group children and four percent of the control group children at the end of 1st grade with a score indicating that the child is particularly shy or reticent.[91]

There is suggestive evidence from teachers who reported more problems for Head Start group children than control group children on teacher and child interactions (effect size=0.13) at the end of 1st grade. In contrast and supported by suggestive evidence, parents of the Head Start group children reported *less* withdrawn behavior than the parents of children in the control group (effect size=-0.13).

These findings are inconsistent, with teachers reporting that children in the Head Start group are more shy and socially reticent than control group children, while their parents are reporting that they are less withdrawn. This discrepancy could be due to any number of issues. For one, teachers and parents see children in different settings, where children may behave differently. Second, it might reflect different expectations that parents and teachers hold for children's behavior. The findings might also be due to chance. Finally, it's possible that

---

[89] Detailed tables for the impact estimates, including sample sizes, standard errors, and confidence intervals, can be found on the Administration for Children and Families, Office of Planning, Research and Evaluation Website at: http://www.acf.hhs.gov/programs/opre/hs/impact_study/index.html.

[90] Due to the large number of statistical tests, the following language was developed to report categories of statistically significant results:

- *Strong Evidence*: the estimated impact for a particular outcome is statistically significant at the typical level ($p \leq 0.05$), and this result holds up under the test for multiple comparisons (i.e., Benjamini-Hochberg procedure).
- *Moderate Evidence*: the estimated impact for a particular outcome is statistically significant at the typical level ($p \leq 0.05$), but this result *does not* hold up under the test for multiple comparisons.
- *Suggestive Evidence*: the estimated impact for a particular outcome is statistically significant under a relaxed standard ($p \leq 0.10$), and this result *may or may not* hold up under the test for multiple comparisons.

[91] Shy/socially reticent behavior is confirmed by a t-score of 60 or higher on the ASPI.

Exhibit 5.1: Estimated Impacts on Social-Emotional Outcomes by Year: 4-Year-Old Cohort

Outcomes	Mean Estimates				Regression-Adjusted Impact		Effect Size
	Head Start Group	Control Group	Head Start - Control	p-value	Impact	p-value	
Head Start Baseline (Fall 2002)							
Parent-Reported Measures							
Aggressive Behavior	2.93	3.07					
Hyperactive Behavior	1.85	2.08					
Withdrawn Behavior	0.70	0.75					
Total Problem Behavior	6.06	6.43					
Social Competencies	10.78	10.84					
Social Skills and Positive Approaches to Learning	12.33	12.22					
Head Start Year (Spring 2003)							
Parent-Reported Measures							
Aggressive Behavior+	2.73	2.86	-0.13	0.263	-0.16	0.164	-0.10
Hyperactive Behavior+	1.71	1.77	-0.06	0.502	-0.09	0.324	-0.06
Withdrawn Behavior+	0.65	0.70	-0.04	0.505	-0.04	0.575	-0.04
Total Problem Behavior	5.60	5.80	-0.20	0.406	-0.27	0.289	-0.08
Social Competencies+	11.01	11.06	-0.04	0.665	-0.04	0.566	-0.03
Social Skills and Positive Approaches to Learning	12.46	12.48	-0.02	0.891	-0.06	0.682	-0.04
Closeness	33.58	33.31	0.27	0.097	0.25	0.146	0.09
Conflict	17.46	17.71	-0.25	0.683	-0.23	0.698	-0.03
Positive Relationships+	64.05	63.47	0.58	0.411	0.56	0.419	0.07
Kindergarten Year (Spring 2004)							
Parent-Reported Measures							
Aggressive Behavior+	2.41	2.47	-0.06	0.608	-0.08	0.477	-0.05
Hyperactive Behavior	1.53	1.39	0.15	0.174	0.11	0.273	0.08
Withdrawn Behavior+	0.74	0.71	0.02	0.721	0.00	0.986	0.00

Exhibit 5.1: Estimated Impacts on Social-Emotional Outcomes by Year: 4-Year-Old Cohort (continued)

Outcomes	Mean Estimates				Regression-Adjusted Impact		Effect Size
	Head Start Group	Control Group	Head Start - Control	p-value	Impact	p-value	
Parent-Reported Measures (cont'd)							
Total Problem Behavior	5.18	4.99	0.19	0.464	0.09	0.710	0.03
Social Competencies+	11.10	11.17	-0.07	0.382	-0.03	0.770	-0.02
Social Skills and Positive Approaches to Learning+	12.66	12.63	0.03	0.780	0.07	0.483	0.05
Closeness	33.19	33.34	-0.15	0.526	-0.06	0.793	-0.02
Conflict	17.68	17.59	0.09	0.854	-0.13	0.788	-0.02
Positive Relationships+	63.38	63.65	-0.27	0.643	0.03	0.956	0.00
Teacher-Reported Measures							
ASPI-Aggressive	48.74	48.72	0.02	0.973	-0.09	0.893	-0.01
ASPI-Inattentive/ Hyperactive	50.49	50.97	-0.48	0.468	-0.69	0.286	-0.08
ASPI-Withdrawn/ Low Energy	49.22	49.08	0.15	0.824	0.10	0.888	0.01
ASPI-Oppositional	48.03	47.67	0.37	0.556	0.13	0.819	0.02
ASPI-Problems with Peer Interaction	51.16	51.70	-0.54	0.629	-0.89	0.410	-0.08
ASPI-Shy/Socially Reticent	47.81	47.13	0.68	0.390	0.64	0.418	0.08
ASPI-Problems with Structured Learning	50.86	51.26	-0.40	0.623	-0.67	0.410	-0.07
ASPI-Problems with Teacher Interaction	50.07	49.79	0.28	0.729	0.20	0.811	0.02
Closeness	30.33	30.11	0.21	0.631	0.26	0.557	0.06
Conflict	13.33	13.52	-0.19	0.759	-0.35	0.558	-0.06
Positive Relationships	64.81	64.38	0.42	0.613	0.63	0.445	0.07
1st Grade Year (Spring 2005)							
Parent-Reported Measures							
Aggressive Behavior	2.20	2.29	-0.09	0.476	-0.09	0.483	-0.05
Hyperactive Behavior	1.43	1.46	-0.03	0.784	0.00	0.972	0.00
Withdrawn Behavior+	0.71	0.83	-0.12	0.075	-0.13*	0.077	-0.13
Total Problem Behavior	4.84	5.05	-0.21	0.450	-0.19	0.453	-0.05

Exhibit 5.1: Estimated Impacts on Social-Emotional Outcomes by Year: 4-Year-Old Cohort (continued)

Outcomes	Mean Estimates				Regression-Adjusted Impact		Effect Size
	Head Start Group	Control Group	Head Start - Control	p-value	Impact	p-value	
Parent-Reported Measures (cont'd)							
Social Competencies+	11.09	11.13	-0.05	0.534	-0.02	0.753	-0.02
Social Skills and Positive Approaches to Learning	12.64	12.63	0.01	0.931	0.02	0.764	0.02
Closeness	33.21	33.26	-0.04	0.789	-0.01	0.944	0.00
Conflict	16.68	17.20	-0.52	0.368	-0.50	0.373	-0.07
Positive Relationships+	64.42	63.99	0.43	0.498	0.41	0.507	0.05
Teacher-Reported Measures							
ASPI-Aggressive	48.56	49.12	-0.56	0.381	-0.72	0.257	-0.09
ASPI-Inattentive/ Hyperactive	50.35	50.50	-0.15	0.852	-0.26	0.731	-0.03
ASPI-Withdrawn/ Low Energy	49.87	49.22	0.65	0.257	0.75	0.169	0.11
ASPI-Oppositional	47.79	47.88	-0.09	0.909	-0.36	0.637	-0.05
ASPI-Problems with Peer Interaction	51.33	51.53	-0.20	0.804	-0.38	0.630	-0.03
ASPI-Shy/Socially Reticent	48.00	46.76	1.24	0.043	1.37**	0.019	0.19
ASPI-Problems with Structured Learning	51.03	50.29	0.74	0.305	0.74	0.306	0.07
ASPI-Problems with Teacher Interaction	50.14	48.81	1.33	0.106	1.29*	0.099	0.13
Closeness	29.91	29.74	0.17	0.544	0.22	0.465	0.05
Conflict	14.22	13.92	0.30	0.543	0.09	0.838	0.01
Positive Relationships	63.54	63.61	-0.07	0.906	0.20	0.728	0.02

Key:
******* $p \leq 0.01$
** $p \leq 0.05$
* $p \leq 0.10$

+ Indicates the reliability (Cronbach alpha) for the measure is < 0.6.

Bold regression-adjusted impact value indicates the outcome passes the Benjamini-Hochberg test for multiple comparisons with a 10 percent false discovery rate.

measurement error explains the discrepancy, as the withdrawn behavior scale reported by parents has low reliability in 1st grade.

Impacts for the 3-Year-Old Cohort

As shown in Exhibit 5.2,[92] there is strong evidence of an impact of access to Head Start on parent-reported behavior at the end of the Head Start year. Parents reported less hyperactive behavior (effect size=-0.21) and fewer total problem behaviors (effect size=-0.14) for the Head Start group as compared to the control group. There is also suggestive evidence of a favorable impact on parent-reported social skills and approaches to learning (effect size= 0.11) at the end of the age 4 year.

At the end of the kindergarten year, there is moderate evidence of an impact on parents' reports of children's hyperactive behavior. Parents reported less hyperactive behavior (effect size=-0.12) for the Head Start group as compared to the control group. As in the age 4 year, there is also suggestive evidence of a favorable impact on children's social skills and approaches to learning (effect size=0.14) as reported by parents in this year.

At the end of 1st grade, there is moderate evidence of an impact on parent reports of closeness with their child (effect size=0.10), indicating more desirable or positive behavior for the Head Start group children compared to those in the control group. There is also suggestive evidence of an improvement on parent-reported positive relationships with their child (effect size=0.10) in this year.

No significant impacts were found on any of the teacher ratings in the social-emotional domain for the 3-year old cohort in any year.

[92] Detailed tables for the impact estimates, including sample sizes, standard errors, and confidence intervals, can be found on the Administration for Children and Families, Office of Planning, Research and Evaluation Website at: http://www.acf.hhs.gov/programs/opre/hs/impact_study/index.html.

Exhibit 5.2: Estimated Impacts on Social-Emotional Outcomes by Year: 3-Year-Old Cohort

Outcomes	Mean Estimates				Regression-Adjusted Impact		Effect Size
	Head Start Group	Control Group	Head Start - Control	p-value	Impact	p-value	
Head Start Baseline (Fall 2002)							
Parent-Reported Measures							
Aggressive Behavior	3.12	3.03					
Hyperactive Behavior	1.86	1.91					
Withdrawn Behavior	0.63	0.60					
Total Problem Behavior	6.16	6.09					
Social Competencies	10.74	10.70					
Social Skills and Positive Approaches to Learning	12.23	12.08					
Head Start Year (Spring 2003)							
Parent-Reported Measures							
Aggressive Behavior	2.97	3.05	-0.08	0.417	-0.10	0.274	-0.06
Hyperactive Behavior	1.71	2.00	-0.29	0.004	**-0.33*** **	0.001	-0.21
Withdrawn Behavior+	0.55	0.58	-0.02	0.708	-0.04	0.510	-0.04
Total Problem Behavior	5.80	6.24	-0.44	0.053	**-0.52*** **	0.003	-0.14
Social Competencies+	10.95	10.99	-0.04	0.540	-0.03	0.637	-0.03
Social Skills and Positive Approaches to Learning	12.41	12.38	0.03	0.740	0.04	0.745	0.02
Closeness	33.63	33.44	0.19	0.178	0.18	0.220	0.06
Conflict	18.04	18.12	-0.07	0.860	-0.05	0.893	-0.01
Positive Relationships+	63.50	63.19	0.31	0.491	0.28	0.517	0.03
Age 4 Year (2004)							
Parent-Reported Measures							
Aggressive Behavior	2.59	2.70	-0.11	0.310	-0.12	0.203	-0.07
Hyperactive Behavior+	1.64	1.73	-0.10	0.349	-0.13	0.242	-0.09
Withdrawn Behavior+	0.59	0.64	-0.05	0.467	-0.08	0.248	-0.08
Total Problem Behavior	5.29	5.63	-0.34	0.197	-0.39	0.115	-0.10

Exhibit 5.2: Estimated Impacts on Social-Emotional Outcomes by Year: 3-Year-Old Cohort (continued)

Outcomes	Mean Estimates				Regression-Adjusted Impact		Effect Size
	Head Start Group	Control Group	Head Start - Control	p-value	Impact	p-value	
Parent-Reported Measures (cont'd)							
Social Competencies	11.04	11.11	-0.06	0.447	-0.01	0.868	-0.01
Social Skills and Positive Approaches to Learning	12.59	12.46	0.13	0.177	0.19*	0.055	0.11
Closeness	33.53	33.34	0.18	0.304	0.22	0.232	0.08
Conflict	17.67	18.12	-0.44	0.412	-0.39	0.461	-0.06
Positive Relationships+	63.80	63.06	0.73	0.255	0.72	0.261	0.09
Kindergarten Year (Spring 2005)							
Parent-Reported Measures							
Aggressive Behavior	2.32	2.37	-0.05	0.695	-0.08	0.382	-0.05
Hyperactive Behavior	1.37	1.52	-0.14	0.124	-0.18**	0.048	-0.12
Withdrawn Behavior+	0.63	0.65	-0.02	0.693	-0.03	0.563	-0.03
Total Problem Behavior	4.84	5.06	-0.22	0.405	-0.26	0.246	-0.07
Social Competencies+	11.06	10.96	0.10	0.209	0.11	0.179	0.08
Social Skills and Positive Approaches to Learning	12.55	12.29	0.26	0.076	0.25*	0.075	0.14
Closeness	33.19	33.05	0.14	0.390	0.13	0.434	0.05
Conflict	17.14	17.13	0.01	0.987	-0.06	0.888	-0.01
Positive Relationships	63.88	63.82	0.06	0.893	0.12	0.803	0.02
Teacher-Reported Measures							
ASPI-Aggressive	49.02	48.66	0.36	0.444	0.40	0.318	0.05
ASPI-Inattentive/ Hyperactive	50.27	50.37	-0.10	0.852	-0.02	0.972	0.00
ASPI-Withdrawn/ Low Energy	49.09	48.66	0.44	0.320	0.52	0.241	0.08
ASPI-Oppositional	48.30	48.36	-0.05	0.896	0.03	0.953	0.00
ASPI-Problems with Peer Interaction	51.49	50.94	0.55	0.449	0.64	0.335	0.06
ASPI-Shy/Socially Reticent	47.37	47.44	-0.07	0.888	0.07	0.882	0.01
ASPI-Problems with Structured Learning	50.40	49.83	0.57	0.309	0.74	0.136	0.07

Exhibit 5.2: Estimated Impacts on Social-Emotional Outcomes by Year: 3-Year-Old Cohort (continued)

Outcomes	Mean Estimates				Regression-Adjusted Impact		Effect Size
	Head Start Group	Control Group	Head Start - Control	p-value	Impact	p-value	
Teacher-Reported Measures (cont'd)							
ASPI-Problems with Teacher Interaction	49.42	49.28	0.14	0.833	0.27	0.649	0.03
Closeness	30.02	30.25	-0.23	0.401	-0.26	0.295	-0.06
Conflict	13.99	14.02	-0.03	0.961	0.03	0.959	0.00
Positive Relationships	63.82	64.14	-0.32	0.613	-0.40	0.483	-0.04
1st Grade Year (Spring 2006)							
Parent-Report Measures							
Aggressive Behavior	2.24	2.27	-0.03	0.775	-0.05	0.624	-0.03
Hyperactive Behavior	1.38	1.49	-0.10	0.219	-0.11	0.127	-0.07
Withdrawn Behavior+	0.73	0.71	0.02	0.641	0.02	0.732	0.02
Total Problem Behavior	4.88	5.01	-0.13	0.577	-0.15	0.439	-0.04
Social Competencies	11.13	11.09	0.05	0.559	0.08	0.317	0.07
Social Skills and Positive Approaches to Learning	12.58	12.51	0.06	0.564	0.05	0.642	0.03
Closeness	33.32	33.09	0.23	0.066	0.29**	0.013	0.10
Conflict	16.75	17.28	-0.53	0.268	-0.55	0.210	-0.08
Positive Relationships	64.46	63.77	0.68	0.172	0.77*	0.098	0.10
Teacher-Reported Measures							
ASPI-Aggressive	48.81	49.15	-0.34	0.528	-0.54	0.266	-0.07
ASPI-Inattentive/ Hyperactive	50.38	50.67	-0.29	0.597	-0.45	0.402	-0.05
ASPI-Withdrawn/ Low Energy	49.51	48.95	0.57	0.287	0.44	0.383	0.06
ASPI-Oppositional	48.39	48.25	0.14	0.835	0.04	0.944	0.01
ASPI-Problems with Peer Interactions	51.96	52.20	-0.24	0.792	-0.43	0.584	-0.04
ASPI-Shy/Socially Reticent	47.36	47.05	0.31	0.533	0.21	0.623	0.03
ASPI-Problems with Structured Learning	50.76	50.55	0.21	0.789	-0.12	0.855	-0.01
ASPI-Problems with Teacher Interaction	50.00	50.13	-0.12	0.847	-0.15	0.819	-0.01

Exhibit 5.2: Estimated Impacts on Social-Emotional Outcomes by Year: 3-Year-Old Cohort (continued)

Outcomes	Mean Estimates				Regression-Adjusted Impact		Effect Size
	Head Start Group	Control Group	Head Start - Control	p-value	Impact	p-value	
Teacher-Reported Measures (cont'd)							
Closeness	29.94	29.74	0.20	0.475	0.28	0.301	0.06
Conflict	14.12	14.13	-0.01	0.983	-0.16	0.722	-0.02
Positive Relationships	63.56	63.37	0.19	0.780	0.44	0.461	0.05

Key:
*** $p \leq 0.01$
** $p \leq 0.05$
* $p \leq 0.10$

+ Indicates the reliability (Cronbach alpha) for the measure is < 0.6.

Bold regression-adjusted impact value indicates the outcome passes the Benjamini-Hochberg test for multiple comparisons with a 10 percent false discovery rate.

Impacts on Participants

This section discusses Head Start's impact on the social-emotional development of children who *actually participated in the program.*[93] Moving from measures of the impact of access to Head Start (ITT estimates) to the impact of participating (IOT estimates) changes the estimates by a factor of about 50 percent for both the 3- and 4-year-old cohorts. This factor was derived for each cohort using methods discussed in Chapter 2 and the Technical Report for the Head Start Impact Study. For example, the impact of access to Head Start on the reduction in parent-reported hyperactive behavior of 3-year-olds at the end of their Head Start year is -0.33 (ITT estimate in Exhibit 5.2), while the impact of participating in Head Start on the reduction in parent-reported hyperactive behavior is -0.47 (IOT estimate in Exhibit E5.2 in Appendix E). However, the statistical significance does not change when moving from access to participation (i.e., any statistically significant impact of access to Head Start is also statistically significant for the impact of participation).

[93] Detailed findings are presented in Appendix E. As noted in Chapter 4, some children granted access never participated in Head Start, while other children randomized into the study's control group did participate.

The larger magnitude of impact from participation, as opposed to access, can be summarized in terms of effect sizes. For example, impacts confirmed by strong evidence (found primarily in the Head Start year for the 3-year old cohort)[94] have effect sizes ranging from 0.20 to 0.30 (positive or negative) for the impact of participation, in contrast to effect sizes of 0.14 to 0.21 (positive or negative) for the impact of access.

For the 3-year-old cohort, the largest measured effect of participation backed by strong evidence[95] is an effect size of -0.30 in reducing hyperactive behavior at the end of the first year in the Head Start program (as compared to an effect size of -0.21 for the impact of access). Overall, however, the switch to effects on *participants* does not change the conclusions of the analysis of access, which showed Head Start for this age cohort reduced hyperactive and total problem behavior at the end of the Head Start year (and, for hyperactive behavior, at the end of kindergarten), and increased reported feelings of closeness and a positive relationship between parents and children at the end of 1st grade and better social skills and positive approaches to learning for the Head Start group children at the end of the age 4 year and at the end of kindergarten.

Summary

School readiness and school success require both the development of academic skills and the acquisition of positive social skills and approaches to learning (Blair, 2002). For the 4-year-olds, there is limited evidence of an impact of Head Start on children's social-emotional development with no impacts identified in the Head Start year or in kindergarten and only limited and conflicting impacts reported at the end of 1st grade.

For the 3-year olds, there is more evidence of an impact of Head Start on children's social-emotional development; however, all of the statistically significant findings are parent-reported. Parents reported strong evidence of reduced hyperactive behavior and reduced total problem behavior for children in the Head Start group during the Head Start year. The reduction of hyperactive behavior continued into kindergarten for children in the Head Start group with a

[94] As noted earlier in the chapter, for the 4-year-old cohort, no outcomes were identified that demonstrated strong evidence of impact.

[95] The discussion on effect sizes is limited to the impacts confirmed by strong evidence for illustration only. All impacts have similar increases in their effect sizes.

moderate report by parents. Parents also reported suggestive evidence of better social skills and positive approaches to learning for children in the Head Start group in both the age 4 year and in kindergarten. Finally, at the end of 1st grade, parents of Head Start group children reported evidence of a positive and closer relationship with their child than parents of children in the control group.

To provide context for the social-emotional findings, a t-score of 60 or higher for any ASPI component empirically confirms a problem with that component. The percent of empirically confirmed problems for the study children in the 4-year-old cohort at the end of 1st grade ranges from a low of four to six percent on the shy/socially reticent component to a high of 25 percent on the problems with peer interaction component.

The findings in the social-emotional domain are limited for the 4-year-old cohort, but positive change is noted for children in the 3-year old Head Start group on selected social-emotional measures from the Head Start year through 1st grade. Although the findings are not always at the highest level of significance ($p \leq 0.01$), the findings are all favorable to the Head Start group for the 3-year old cohort.

Chapter 6: The Impact of Head Start on Children's Receipt of Health Care Services and Current Health Status

Introduction

This chapter focuses on the impact of access to and participation in Head Start on children's receipt of health care services—parent-reported information on dental care and health insurance and parent-reported information on their child's current health status. No direct collection of health data was included in the study (e.g., from direct health examinations, health records, or medical provider reports).

Measures

The Head Start program seeks to provide comprehensive child development services to economically disadvantaged children and families. The program has a broad mandate to improve children's growth and development, which includes the provision of health, nutritional, and social services to enrolled children and families. Early identification and treatment of health problems is intended to reduce complications from undetected conditions and improve long-term health outcomes, thereby helping to ensure that children enter school ready to learn. Hence, Head Start programs are required to conduct early developmental and health screenings, to link children and families to needed services (including Medicaid and the State Children's Health Insurance Programs when eligible).

Measures of a child's health status and the receipt of health care services for this study are briefly described below.[96]

Parents were asked to report on two health care services and three aspects of their child's health status. Each measure is described below.

- **Whether the Child Has Health Insurance.** Parents were asked if their child was covered by Medicaid or a state health insurance program or by health insurance through their job or the job of another employed adult.

[96] See Chapter 2 and the Technical Report for the Head Start Impact Study for detailed information on the health measures.

- **Whether the Child Has Received Dental Care.**[97] Parents were asked if the child had seen a dentist since September of that year.
- **Child's Health Status Is Excellent or Very Good.** Parents were asked if, overall, their child's health was excellent, very good, good, fair, or poor. This outcome was coded "yes" for those who reported that their child's health was excellent or very good.
- **Whether the Child Needs Ongoing Medical Care.** Parents were asked if their child had an illness or condition that requires regular ongoing medical care.
- **Whether the Child Received Medical Care for an Injury in the Last Month.** Parents were asked how many times their child, in the last month, had seen a doctor or other medical professional or visited a clinic or emergency room for an injury. This outcome was coded "yes" if the parent reported any such occurrences in the last month.

The first set of impact estimates discussed in this chapter measure the effect of access to Head Start on the average child randomly assigned to the Head Start group (the intent to treat or ITT estimates), while the end of the chapter provides a discussion of the impact of Head Start on the children who actually participated in the program (i.e., the impact on the treated or the IOT estimates). (See Chapter 2 and the Technical Report for the Head Start Impact Study for a discussion of the methodology used for the impact estimates.)

Impacts for the 4-Year-Old Cohort

As shown in Exhibit 6.1,[98] there is evidence of the impact of Head Start on a child's receipt of health care services. There is strong evidence[99] of an impact of access to Head Start

---

[97] At the time of the 2002 baseline, parents were asked whether the child had ever seen a dentist.

[98] Detailed tables for the impact estimates, including sample sizes, standard errors, and confidence intervals, can be found on the Administration for Children and Families, Office of Planning, Research and Evaluation Website at: http://www.acf.hhs.gov/programs/opre/hs/impact_study/index.html.

[99] Due to the large number of statistical tests, the following language was developed to report categories of statistically significant results:

- *Strong Evidence*: the estimated impact for a particular outcome is statistically significant at the typical level ($p \leq 0.05$), and this result holds up under the test for multiple comparisons (i.e., Benjamini-Hochberg procedure).
- *Moderate Evidence*: the estimated impact for a particular outcome is statistically significant at the typical level ($p \leq 0.05$), but this result *does not* hold up under the test for multiple comparisons.
- *Suggestive Evidence*: the estimated impact for a particular outcome is statistically significant under a relaxed standard ($p \leq 0.10$), and this result *may or may not* hold up under the test for multiple comparisons.

Exhibit 6.1: Estimated Impacts on Parent-Reported Child Health Outcomes by Year: 4-Year-Old Cohort

Outcomes	Mean Estimates				Regression-Adjusted Impact		Effect Size
	Head Start Group	Control Group	Head Start - Control	p-value	Impact	p-value	
Head Start Baseline (Fall 2002)							
Parent-Reported Measures							
Child Received Dental Care	0.83	0.67					
Child Has Health Insurance Coverage	0.84	0.89					
Child's Overall Health Status Is Excellent/Good	0.79	0.79					
Child Needs Ongoing Care	0.13	0.19					
Child Had Care for Injury Last Month	0.06	0.05					
Head Start Year (Spring 2003)							
Parent-Reported Measures							
Child Received Dental Care	0.73	0.56	0.17	0.000	**0.15*** **	0.000	0.31
Child Has Health Insurance Coverage	0.89	0.88	0.01	0.473	0.01	0.733	0.02
Child's Overall Health Status Is Excellent/Good	0.79	0.82	-0.03	0.225	-0.03	0.244	-0.07
Child Needs Ongoing Care	0.12	0.11	0.01	0.656	0.01	0.422	0.05
Child Had Care for Injury Last Month	0.12	0.12	0.00	0.929	-0.02	0.409	-0.06
Kindergarten Year (Spring 2004)							
Parent-Reported Measures							
Child Received Dental Care	0.67	0.66	0.01	0.670	0.03	0.435	0.06
Child Has Health Insurance Coverage	0.90	0.86	0.04	0.060	0.04*	0.056	0.11

Exhibit 6.1: Estimated Impacts on Parent-Reported Child Health Outcomes by Year: 4-Year-Old Cohort (continued)

Outcomes	Mean Estimates				Regression-Adjusted Impact		Effect Size
	Head Start Group	Control Group	Head Start - Control	p-value	Impact	p-value	
Parent-Reported Measures (cont'd)							
Child's Overall Health Status Is Excellent/Good	0.82	0.76	0.06	0.094	0.05*	0.098	0.13
Child Needs Ongoing Care	0.12	0.13	-0.01	0.619	-0.02	0.432	-0.06
Child Had Care for Injury Last Month	0.12	0.10	0.02	0.380	0.02	0.547	0.05
1st Grade Year (Spring 2005)							
Parent-Reported Measures							
Child Received Dental Care	0.68	0.66	0.02	0.524	0.02	0.550	0.03
Child Has Health Insurance Coverage	0.89	0.85	0.04	0.073	0.04**	0.044	0.11
Child's Overall Health Status Is Excellent/Good	0.81	0.81	0.00	0.951	-0.01	0.858	-0.01
Child Needs Ongoing Care	0.14	0.13	0.02	0.476	0.02	0.291	0.07
Child Had Care for Injury Last Month	0.15	0.13	0.02	0.260	0.02	0.303	0.06

Key:
*** $p \leq 0.01$
** $p \leq 0.05$
* $p \leq 0.10$

Bold regression-adjusted impact value indicates the outcome passes the Benjamini-Hochberg test for multiple comparisons with a 10 percent false discovery rate.

on children's receipt of dental care since September of their Head Start year—a difference of 15 percentage points in favor of the Head Start group (effect size=0.31). However, this impact is not sustained in kindergarten and 1st grade.

There also is suggestive evidence at the end of kindergarten of an impact on health insurance coverage (effect size=0.11). This impact continues into 1st grade, when there is

moderate evidence of an impact on health insurance coverage—a difference of four percentage points in favor of the Head Start group (effect size=0.11).

With regard to health status, there is only limited evidence of an impact of Head Start for the 4-year-old cohort. At the end of kindergarten, there is suggestive evidence that Head Start increases the likelihood that children are in very good or excellent health. In this year, 82 percent of parents of children in the Head Start group reported their child's overall health as excellent or very good as compared to 76 percent of parents of control group children (effect size=0.13).

Impacts for the 3-Year-Old Cohort

Similar to the 4-year-old cohort, there is evidence of the impact of Head Start on the receipt of dental care since September of that year for the 3-year-old cohort. As shown in Exhibit 6.2,[100] there is strong evidence of an impact of access to Head Start on children's receipt of dental care at the end of the Head Start year and at the end of children's age 4 year—differences of 17 (effect size=0.33) and ten (effect size=0.20) percentage points respectively in favor of the Head Start group. However, these impacts were no longer significant by the time the children were in kindergarten or 1st grade.

There is limited evidence that Head Start increases health insurance coverage for the 3-year-old cohort, at least temporarily. There was moderate evidence of an impact on health insurance coverage at the end of kindergarten, with a difference of four percentage points in favor of the Head Start group (effect size=0.14). There were no impacts on health insurance in any other year for this age cohort.

Evidence that Head Start affects children's health status is also limited for this age cohort. At the end of the Head Start year, there was moderate evidence that Head Start children were more likely to be in very good or excellent health (effect size=0.11). Eighty-one percent of the Head Start group parents reported that their child's overall health was excellent or very good compared with 76 percent of the control group parents. There were no impacts on health status in any other years for this age group.

---

[100] Detailed tables for the impact estimates, including sample sizes, standard errors, and confidence intervals, can be found on the Administration for Children and Families, Office of Planning, Research and Evaluation Website at: http://www.acf.hhs.gov/programs/opre/hs/impact_study/index.html.

Exhibit 6.2: Estimated Impacts on Child Health Outcomes by Year: 3-Year-Old Cohort

Outcomes	Mean Estimates				Regression-Adjusted Impact		Effect Size
	Head Start Group	Control Group	Head Start - Control	p-value	Impact	p-value	
Head Start Baseline (Fall 2002)							
Parent-Reported Measures							
Child Received Dental Care	0.76	0.58					
Child Has Health Insurance Coverage	0.91	0.87					
Child's Overall Health Status Is Excellent/Good	0.80	0.78					
Child Needs Ongoing Care	0.18	0.22					
Child Had Care for Injury Last Month	0.06	0.07					
Head Start Year (Spring 2003)							
Parent-Reported Measures							
Child Received Dental Care	0.68	0.52	0.17	0.000	**0.17*** **	0.000	0.33
Child Has Health Insurance Coverage	0.92	0.92	0.01	0.738	0.00	0.803	0.01
Child's Overall Health Status Is Excellent/Good	0.81	0.76	0.05	0.084	0.05**	0.045	0.11
Child Needs Ongoing Care	0.13	0.13	0.00	0.910	0.00	0.988	0.00
Child Had Care for Injury Last Month	0.09	0.09	0.00	0.918	-0.01	0.699	-0.02
Age 4 Year (Spring 2004)							
Parent-Reported Measures							
Child Received Dental Care	0.74	0.65	0.09	0.001	**0.10*** **	0.001	0.20
Child Has Health Insurance Coverage	0.93	0.92	0.01	0.460	0.00	0.935	0.00
Child's Overall Health Status Is Excellent/Good	0.83	0.83	0.00	0.897	0.00	0.851	0.01
Child Needs Ongoing Care	0.13	0.13	0.00	0.862	0.01	0.739	0.02

Exhibit 6.2: Estimated Impacts on Child Health Outcomes by Year: 3-Year-Old Cohort (continued)

	Mean Estimates				Regression-Adjusted Impact		
	Head Start Group	Control Group	Head Start - Control	p-value	Impact	p-value	Effect Size
Parent-Reported Measures (cont'd)							
Child Had Care for Injury Last Month	0.12	0.09	0.03	0.039	0.03*	0.089	0.10
Kindergarten Year (Spring 2005)							
Parent-Reported Measures							
Child Received Dental Care	0.75	0.72	0.04	0.114	0.03	0.270	0.06
Child Has Health Insurance Coverage	0.93	0.90	0.03	0.087	0.04**	0.044	0.14
Child's Overall Health Status Is Excellent/Good	0.81	0.83	-0.01	0.660	0.00	0.889	-0.01
Child Needs Ongoing Care	0.15	0.19	-0.04	0.051	-0.03	0.114	-0.07
Child Had Care for Injury Last Month	0.13	0.13	0.00	0.836	0.00	0.985	0.00
1st Grade Year (Spring 2006)							
Parent-Reported Measures							
Child Received Dental Care	0.74	0.73	0.02	0.514	0.01	0.786	0.02
Child Has Health Insurance Coverage	0.93	0.92	0.02	0.187	0.02	0.252	0.06
Child's Overall Health Status Is Excellent/Good	0.84	0.84	0.00	0.900	0.02	0.434	0.04
Child Needs Ongoing Care	0.16	0.17	-0.01	0.705	-0.01	0.578	-0.03
Child Had Care for Injury Last Month	0.09	0.07	0.02	0.370	0.02	0.294	0.07

Key:
*** $p \leq 0.01$
** $p \leq 0.05$
* $p \leq 0.10$

Bold regression-adjusted impact value indicates the outcome passes the Benjamini-Hochberg test for multiple comparisons with a 10 percent false discovery rate.

There is also one counter-intuitive finding. At the end of the age 4 year, there is suggestive evidence that Head Start increases the likelihood that children were treated for an injury in the last month (effect size=0.10). The interpretation of this finding is unclear, as it could reflect either an increase in injuries or an increase in care-seeking, or both.

Impacts on Participants

This section discusses Head Start's impact on health status and receipt of health services for children who *actually participated in the program.*[101] Moving from measures of the average impact of access to Head Start (ITT estimates) to the average impact of participating (IOT estimates) increases all findings by about 50 percent. For example, the impact of access to Head Start on dental care for 4-year-olds at the end of Head Start is 0.15 (ITT estimate in Exhibit 6.1) while the impact of participating in Head Start on dental care is 0.22 (IOT estimate in Exhibit E6.1 in Appendix E). All increases in magnitude for participants follow this proportion for the 4-year-old cohort. Similarly, the impact of access to Head Start on dental care in spring 2003 for the 3-year-old cohort is 0.17 (ITT estimate in Exhibit 6.2), while the impact for participating in Head Start on dental care is 0.24 (IOT estimate in Exhibit E6.2 in Appendix E), with all changes in magnitude for that cohort following the same proportion. As noted in previous chapters, the statistical significance does not change when moving from access to participation (i.e., any statistically significant impact of access to Head Start is also statistically significant for the impact of participation). Yet, these estimates suggest that the identified benefits of Head Start are likely to be larger than reported for the ITT impacts.

The larger magnitude of impact from participation, as opposed to access, can be summarized in terms of the percentage of children receiving dental care. For example, for children with access to Head Start in the 4-year-old cohort, there is a 15 percent difference in receipt of dental care favoring the Head Start group. Participation in Head Start changes the impact to a difference of 22 percent between the Head Start group and the control group for this cohort. A similar pattern exists for the 3-year-old cohort. With access to Head Start, there is a 17 percent difference on receipt of dental care in favor of the Head Start group, while

[101] Detailed findings are presented in Appendix E. As noted in Chapter 4, some children granted access never participated in Head Start, while other children randomized into the study's control group did participate.

participation in Head Start results in a difference of 24 percent. The switch to effects on participants does not change the overall conclusion for either cohort from the main analysis.

Summary

The relatively large impact on children's receipt of dental care for both age cohorts when children are in Head Start is particularly important in light of numerous studies that have documented substantial inadequacy in the level of dental services received by low-income and minority children, who are most at risk of having untreated cavities compared with other children. For example, a GAO study published in 2000 reported that among children ages two through five who had family incomes below $10,000, nearly one in three had at least one decayed tooth that had not been treated (U.S. General Accounting Office, 2000). In contrast, only one in ten children of the same age from families with incomes of $35,000 or higher had untreated cavities. This disparity is recognized in the Healthy People 2010 objectives, one of which is to "*Increase the proportion of low-income children and adolescents who received any preventive dental service during the past year*" (U.S. Department of Health and Human Services, 2000). In fact, the proportion of Head Start children who received dental care in this study exceeded the 66 percent target in Healthy People 2010. For the 4-year-olds, 73 percent of children in the Head Start group had seen a dentist since September in their Head Start year. For the 3-year-olds, 68 percent of children in the Head Start group had seen a dentist since September of their Head Start year, and 74 percent had seen a dentist since September of their age 4 year.

Head Start also appears to have led to increases in children's health insurance coverage during the early school years. In 2004, an estimated 7.8 million children were uninsured, a decline of 1.8 million from 1999 (The Urban Institute, 2007) due in large part to eligibility for Medicaid or the State Children's Health Insurance Program (SCHIP.) During the kindergarten year, 90 percent of the Head Start group children in the 4-year-old cohort and 93 percent of Head Start group children in the 3-year-old cohort were covered by health insurance as reported by their parents. Suggestive evidence of a significant impact is indicated for the 4-year-old cohort, while moderate evidence of a significant impact is indicated for the 3-year-old cohort. Notably, for the 4-year-old cohort, this impact continues through the end of 1st grade.

Other findings in the health domain could be of interest to the Head Start community. There is suggestive evidence that providing access to Head Start at age 4 improves children's health status in kindergarten and that access at age 3 improves health status at the end of the Head Start year and may increase the likelihood that children receive care for an injury.

Comparing the health status of children in the Head Start Impact Study with children in the general population demonstrates that Head Start children are about equal to other children on general health status. In the ECLS-K study (U.S. Department of Education, 2000), 83 percent of parents reported that their child's health status was good or excellent at the beginning of kindergarten. At the end of kindergarten in the Head Start Impact Study, 82 percent of the Head Start group parents for the 4-year-old cohort and 81 percent of the Head Start group parents for the 3-year-old cohort reported that their child's health was good or excellent. These numbers compare favorably, suggesting that the low-income parents in this study do not see their children as more or less likely to be in good health than parents nationally. However, it is important to note that we have only very limited measures of health in this study, and cannot make definitive statements about how the health of the study children compares to other children nationally on many health issues (e.g., obesity, asthma).

Chapter 7: The Impact of Head Start on Parenting Practices

Introduction

This chapter focuses on the impact of access to and participation in Head Start on the practices of the parents of the study children, including child disciplinary practices, educational supports, parental safety practices, school communication and parent participation, and parenting styles. Data are derived from both parent and teacher reports.

Measures

As noted above, the measures used in this report to assess the impact of Head Start on parenting practices focus on five key parenting constructs that are briefly described below.[102]

- **Disciplinary Practices.** Parents were asked to report on the use of spanking and time out.
- **Educational Activities and Supports.** Parents were asked to report on the use of educational activities such as frequency of reading to their child and participation in cultural enrichment activities with their child.
- **Safety Practices.** Parents were asked to report on the use of safety precautions ranging from the use of car seats to having a working smoke detector in the home. These items were combined to create a safety practices scale.
- **School Contact and Communication/Parent Participation.** Teachers were asked two questions related to school communication with the parents (i.e., frequency of parent-initiated contact about the child's progress and frequency of teacher-initiated contact with the child's home about problems with the child) and two questions related to parent participation (i.e., parents' attendance at school or class activities and parents serving as volunteers for class or school activities).
- **Parenting Styles.** Parents were asked to respond to selected items from the Child Rearing Practices Report (CRPR) (Block, 1965). Indicators of authoritative, authoritarian, permissive, and neglectful parenting are created from this measure.

The skill-based dimensions of parenting practices emphasizing cognitive stimulation, child discipline, and child safety are common elements of parent education offered through Head Start and thus have the potential to be affected by parents' access to the program. Prior research with similar populations has shown significant associations between these domains and

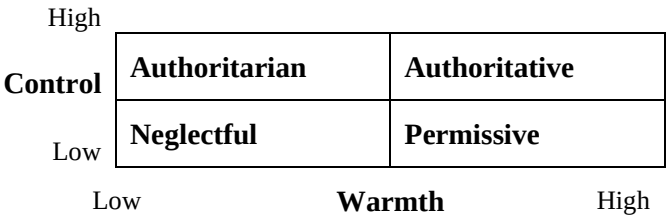
[102] See Chapter 2 and the Technical Report for the Head Start Impact Study for detailed information on the parenting practices measures.

children's cognitive and social-emotional development, including in samples of Head Start children (Eamon, 2005; U.S. Department of Health and Human Services, 2003).

The inclusion of parenting styles provides an opportunity to investigate parent's overall approach to childrearing rather than the use of any particular parenting practice. Specifically, parenting styles differ from parenting practices in that parenting practices refer to single behaviors or actions (e.g., spanking, reading to the child) whereas parenting styles capture an overall climate in which parenting practices occur.

For this study, parenting styles were derived using Baumrind's typology (1971) to assess the degree to which the two dimensions of warmth (nurturing and supportive behaviors) and control (disciplinary strategies, restrictive behaviors) exist in the parent-child relationship, based on parents' reports of their approaches to parenting. These two dimensions form four matrix cells that define different parenting styles (Figure 7.1): (1) "authoritative" parents are identified as being high in both warmth and control, (2) "permissive" parents are high in warmth but low in control, (3) "authoritarian" parents exhibit high levels of control but low levels of warmth, and (4) "neglectful" parents are low in both warmth and control. Evidence suggests that authoritative parenting styles produce the most desirable outcomes in children. For example, one study reported a positive association between authoritative parenting and academic performance, whereas children of neglectful parents had the worst academic performance of all parenting styles (Pong et al., 2005). Similarly, in a study of six-to-seven-year-old children, authoritative parenting was related to characteristics critical to school success (e.g., task persistence, lack of failure expectation) (Onatsu-Arvilommi et al., 1998). Positive outcomes are not restricted to the academic domain. Authoritative parenting style has also been linked to a reduction in child behavior problems for young children in poverty (Linver et al., 2002). Although prior studies have linked authoritarian parenting styles to Black parents (McLoyd, 1990), recent studies (Amato & Fowler, 2002; Deater-Deckard et al., 1996; Lareau, 2002; Pinderhughes et al., 2000) have reported that parenting styles "differ more as a function of socio-economic status (SES) rather than race or culture, with both White and minority low-income parents exhibiting parenting repertoires that are more directive, controlling, and punitive than those of their middle-income counterparts" (Coolahan et al., 2002, p.358).

Figure 7.1: Baumrind's Typology for Parenting Styles



The first set of impact estimates discussed in this chapter measures the effect of Head Start on the average child randomly assigned to the Head Start Head Start group (i.e., the intent to treat or ITT estimates), while the end of the chapter provides a discussion of the impact of Head Start on the children who actually participated in the program (i.e., the impact on the treated or the IOT estimates). (See Chapter 2 and the Technical Report for the Head Start Impact Study for a discussion of the methodology used for the impact estimates.)

Impacts for the 4-Year-Old Cohort

With regard to disciplinary practices (Exhibit 7.1),[103] the use of both spanking and time out declined each year for children in both the Head Start and control groups. That is, parents of children in the 4-year-old cohort used disciplinary strategies like spanking and time out less frequently as their children got older, regardless of whether they received access to Head Start.

There is only limited evidence that Head Start affected parenting practices for the 4-year-old cohort. There was moderate evidence[104] that access to Head Start led to a decrease in the use

[103] Detailed tables for the impact estimates, including sample sizes, standard errors, and confidence intervals, can be found on the Administration for Children and Families, Office of Planning, Research and Evaluation Website at: http://www.acf.hhs.gov/programs/opre/hs/impact_study/index.html.

[104] Due to the large number of statistical tests, the following language was developed to report categories of statistically significant results:

- *Strong Evidence*: the estimated impact for a particular outcome is statistically significant at the typical level ($p \leq 0.05$), and this result holds up under the test for multiple comparisons (i.e., Benjamini-Hochberg procedure).
- *Moderate Evidence*: the estimated impact for a particular outcome is statistically significant at the typical level ($p \leq 0.05$), but this result *does not* hold up under the test for multiple comparisons.
- *Suggestive Evidence*: the estimated impact for a particular outcome is statistically significant under a relaxed standard ($p \leq 0.10$), and this result *may or may not* hold up under the test for multiple comparisons.

Exhibit 7.1: Estimated Impacts on Parenting Outcomes by Year: 4-Year-Old Cohort

Outcomes	Mean Estimates				Regression-Adjusted Impact		Effect Size
	Head Start Group	Control Group	Head Start - Control	p-value	Impact	p-value	
Head Start Baseline (Fall 2002)							
Parent-Reported Measures							
Parent Spanked Child in Last Week	0.41	0.43					
Parent Used Time Out in Last Week	0.62	0.61					
Parent Read to Child in Last Week	0.37	0.34					
Parental Safety Practices Scale	3.66	3.66					
Family Cultural Enrichment Scale	3.44	3.41					
Head Start Year (Spring 2003)							
Parent-Reported Measures							
Parent Spanked Child in Last Week	0.38	0.38	0.00	0.946	-0.01	0.750	-0.02
Parent Used Time Out in Last Week	0.61	0.68	-0.07	0.044	-0.08**	0.025	-0.17
Parent Read to Child in Last Week	0.36	0.33	0.03	0.391	0.03	0.396	0.06
Parental Safety Practices Scale+	3.73	3.71	0.02	0.364	0.03	0.382	0.08
Family Cultural Enrichment Scale+	4.01	3.89	0.11	0.237	0.08	0.368	0.06
Kindergarten Year (Spring 2004)							
Parent-Reported Measures							
Parent Spanked Child in Last Week	0.29	0.31	-0.02	0.481	0.00	0.869	-0.01
Parent Used Time Out in Last Week	0.58	0.57	0.01	0.858	0.01	0.689	0.02
Parent Read to Child in Last Week	0.35	0.38	-0.03	0.440	-0.03	0.385	-0.07
Parental Safety Practices Scale+	3.72	3.68	0.03	0.251	0.04	0.156	0.11
Family Cultural Enrichment Scale+	4.11	3.97	0.14	0.153	0.14	0.142	0.10
Parenting Style: Authoritarian	0.08	0.08	0.00	0.873	0.00	0.981	0.00
Parenting Style: Authoritative	0.68	0.63	0.05	0.205	0.05	0.164	0.10

Exhibit 7.1: Estimated Impacts on Parenting Outcomes by Year: 4-Year-Old Cohort (continued)

Outcomes	Mean Estimates				Regression-Adjusted Impact		Effect Size
	Head Start Group	Control Group	Head Start - Control	p-value	Impact	p-value	
Parent-Reported Measures (cont'd)							
Parenting Style: Neglectful	0.09	0.12	-0.03	0.200	-0.03	0.211	-0.09
Parenting Style: Permissive	0.18	0.19	-0.01	0.548	-0.02	0.447	-0.05
Teacher-Reported Measures							
School Contact and Communication	0.83	0.84	-0.01	0.848	-0.01	0.845	-0.02
Parent Participation	0.89	0.90	-0.01	0.557	-0.01	0.841	-0.02
1st Grade Year (Spring 2005)							
Parent-Reported Measures							
Parent Spanked Child in Last Week	0.21	0.22	0.00	0.881	0.00	0.976	0.00
Parent Used Time Out in Last Week	0.49	0.53	-0.04	0.305	-0.04	0.322	-0.08
Parent Read to Child in Last Week	0.40	0.43	-0.03	0.407	-0.01	0.733	-0.02
Family Cultural Enrichment Scale+	4.00	3.94	0.06	0.529	0.04	0.612	0.03
Parenting Style: Authoritarian	0.07	0.10	-0.03	0.156	-0.03	0.199	-0.10
Parenting Style: Authoritative	0.67	0.64	0.03	0.302	0.04	0.158	0.09
Parenting Style: Neglectful	0.07	0.09	-0.01	0.416	-0.02	0.327	-0.06
Parenting Style: Permissive	0.19	0.18	0.01	0.579	0.00	0.936	0.00

Exhibit 7.1: Estimated Impacts on Parenting Outcomes by Year: 4-Year-Old Cohort (continued)

Outcomes	Mean Estimates				Regression-Adjusted Impact		Effect Size
	Head Start Group	Control Group	Head Start - Control	p-value	Impact	p-value	
Teacher-Reported Measures							
School Contact and Communication	0.79	0.81	-0.02	0.635	-0.02	0.570	-0.06
Parent Participation+	0.86	0.87	-0.01	0.753	-0.01	0.817	-0.02

Key:
*** $p \leq 0.01$
** $p \leq 0.05$
* $p \leq 0.10$

+ Indicates the reliability (Cronbach alpha) for the measure is <0.6.

Bold regression-adjusted impact value indicates the outcome passes the Benjamini-Hochberg test for multiple comparisons with a 10 percent false discovery rate.

Note: The derivation of the variable, Parent Read to Child in Last Week, was revised for this report. The findings in this report may differ from the findings in the *Head Start Impact Study: First Year Findings* (U.S. Department of Health and Human Services, 2005). See Chapter 2 and the Technical Report for the Head Start Impact Study for details of this change.

of time out measured at the end of the Head Start year. There was an eight percentage point difference in the use of time out between the Head Start and control groups (effect size=-0.17). The interpretation of a decrease in the use of time out as a disciplinary technique is not clear. Parents may have decreased the use of time out because their children's behavior improved, or they may have switched to other (likely less positive) parenting techniques (although no increases in spanking were found).

There are no other statistically significant impacts on any of the other parenting practice outcomes for children in the 4-year-old cohort in any year.

Impacts for the 3-Year-Old Cohort

As in the 4-year-old cohort, there were reductions in the use of both spanking and time out each successive year for both the Head Start and control groups in the 3-year-old cohort

(Exhibit 7.2).[105] That is, parents of children in the 3-year-old cohort, regardless of whether they received access to Head Start, used disciplinary strategies of spanking and time out less frequently as their children progressed to early elementary school.

There is also strong evidence of an impact of access to Head Start on parent's use of spanking at the end of the Head Start year. Parents of children in the Head Start group were seven percentage points less likely to have spanked their children in the last week than parents in the control group (effect size=-0.14). There were no impacts on disciplinary practices at the end of the age 4 year, but evidence of impacts re-emerged in kindergarten and 1st grade. There is moderate evidence that Head Start led to a decrease in the use of time out at the end of kindergarten (effect size=-0.13); suggestive evidence of a reduction in spanking at the end of kindergarten (effect size=-0.09), and suggestive evidence that the reduction in time out continues through the end of 1st grade (effect size=-0.11). As noted for the 4-year-olds, changes in the use of discipline may reflect changes in children's behavior, changes in parents' reactions, or both.

In the area of educational supports, there is strong evidence of an impact of Head Start on two parent-reported measures at the end of the Head Start year. Parents of children in the Head Start group were seven percentage points more likely to have read to their child in the last week than parents in the control group (effect size=0.15). Further, parents in the Head Start group were more likely to have involved their child in cultural enrichment activities during the past month than parents of children in the control group (effect size=0.18). These findings are not sustained through the age 4 year, or into kindergarten or 1st grade.

Finally, there is evidence that Head Start had a favorable impact on the parenting styles used by parents of children in the 3-year-old cohort. At the end of the age 4 year and the end of 1st grade, parents in the Head Start group were less likely to use an authoritarian parenting style (characterized by low warmth and high parental control) than parents in the control group. The strength of the evidence of these impacts varied by year, while the effect sizes were similar (effect size=-0.14 supported by strong evidence in the age 4 year and effect size=-0.11 supported by moderate evidence in 1st grade).

---

[105] Detailed tables for the impact estimates, including sample sizes, stand errors, and confidence intervals, can be found on the Administration for Children and Families, Office of Planning, Research and Evaluation Website at: http://www.acf.hhs.gov/programs/opre/hs/impact_study/index.html.

Exhibit 7.2: Estimated Impacts on Parenting Outcomes by Year: 3-Year-Old Cohort

Outcomes	Mean Estimates				Regression-Adjusted Impact		Effect Size
	Head Start Group	Control Group	Head Start - Control	p-value	Impact	p-value	
Head Star Baseline (Fall 2002)							
Parent-Reported Measures							
Parent Spanked Child in Last Week	0.45	0.50					
Parent Used Time Out in Last Week	0.61	0.66					
Parent Read to Child in Last Week	0.36	0.32					
Parental Safety Practices Scale	3.67	3.65					
Family Cultural Enrichment Scale	3.41	3.24					
Head Start Year (Spring 2003)							
Parent-Reported Measures							
Parent Spanked Child in Last Week	0.42	0.48	-0.07	0.037	**-0.07****	0.025	-0.14
Parent Used Time Out in Last Week	0.62	0.66	-0.04	0.211	-0.04	0.205	-0.08
Parent Read to Child in Last Week	0.35	0.28	0.07	0.045	**0.07****	0.030	0.15
Parental Safety Practices Scale+	3.73	3.70	0.03	0.170	0.03	0.146	0.10
Family Cultural Enrichment Scale+	3.78	3.55	0.23	0.003	**0.25*****	0.000	0.18
Age 4 Year (Spring 2004)							
Parent-Reported Measures							
Parent Spanked Child in Last Week	0.35	0.33	0.02	0.602	0.01	0.635	0.03
Parent Used Time Out in Last Week	0.61	0.64	-0.03	0.338	-0.02	0.355	-0.05
Parent Read to Child in Last Week	0.34	0.35	0.00	0.917	-0.01	0.827	-0.01
Parental Safety Practices Scale+	3.73	3.71	0.03	0.282	0.02	0.313	0.06
Family Cultural Enrichment Scale+	3.94	3.87	0.07	0.415	0.04	0.593	0.03
Parenting Style: Authoritarian	0.04	0.08	-0.04	0.006	**-0.04*****	0.005	-0.14
Parenting Style: Authoritative	0.72	0.67	0.05	0.081	0.04	0.186	0.08

Exhibit 7.2: Estimated Impacts on Parenting Outcomes by Year: 3-Year-Old Cohort (continued)

Outcomes	Mean Estimates				Regression-Adjusted Impact		Effect Size
	Head Start Group	Control Group	Head Start - Control	p-value	Impact	p-value	
Parent-Reported Measures (cont'd)							
Parenting Style: Neglectful	0.06	0.06	0.00	0.823	0.00	0.826	-0.01
Parenting Style: Permissive	0.19	0.20	0.00	0.882	0.00	0.893	0.01
Kindergarten Year (Spring 2005)							
Parent-Reported Measures							
Parent Spanked Child in Last Week	0.26	0.31	-0.04	0.073	-0.04*	0.070	-0.09
Parent Used Time Out in Last Week	0.52	0.58	-0.06	0.022	-0.07**	0.013	-0.13
Parent Read to Child in Last Week	0.36	0.32	0.03	0.302	0.03	0.305	0.07
Parental Safety Practices Scale+	3.72	3.71	0.01	0.673	0.01	0.714	0.02
Family Cultural Enrichment Scale+	3.93	3.93	0.00	0.988	0.00	0.968	0.00
Parenting Style: Authoritarian	0.06	0.07	-0.01	0.545	0.00	0.950	0.00
Parenting Style: Authoritative	0.67	0.66	0.01	0.656	0.00	0.905	-0.01
Parenting Style: Neglectful	0.06	0.08	-0.02	0.229	-0.02	0.202	-0.09
Parenting Style: Permissive	0.21	0.19	0.02	0.475	0.03	0.310	0.07
Teacher-Reported Measures							
School Contact and Communication	0.82	0.82	0.00	0.957	0.00	0.879	0.01
Parent Participation+	0.87	0.87	-0.01	0.806	0.00	0.886	-0.01
1st Grade Year (Spring 2006)							
Parent-Reported Measures							
Parent Spanked Child in Last Week	0.21	0.23	-0.03	0.252	-0.03	0.183	-0.07
Parent Used Time Out in Last Week	0.48	0.53	-0.06	0.080	-0.05*	0.075	-0.11
Parent Read to Child in Last Week	0.38	0.36	0.02	0.525	0.01	0.746	0.02
Family Cultural Enrichment Scale+	3.92	3.87	0.05	0.590	0.01	0.879	0.01

Exhibit 7.2: Estimated Impacts on Parenting Outcomes by Year: 3-Year-Old Cohort (continued)

Outcomes	Mean Estimates				Regression-Adjusted Impact		Effect Size
	Head Start Group	Control Group	Head Start - Control	p-value	Impact	p-value	
Parent-Reported Measures (cont'd)							
Parenting Style: Authoritarian	0.05	0.07	-0.03	0.031	-0.03**	0.046	-0.11
Parenting Style: Authoritative	0.71	0.69	0.03	0.389	0.00	0.877	0.01
Parenting Style: Neglectful	0.06	0.06	0.00	0.758	0.00	0.974	0.00
Parenting Style: Permissive	0.18	0.17	0.01	0.743	0.02	0.320	0.06
Teacher-Reported Measures							
School Contact and Communication	0.82	0.79	0.03	0.353	0.02	0.453	0.05
Parent Participation+	0.85	0.84	0.01	0.739	0.01	0.626	0.04

Key:
******* $p \leq 0.01$
** $p \leq 0.05$
* $p \leq 0.10$

+ Indicates the reliability (Cronbach alpha) for the measure is <0.06.

Bold regression-adjusted impact value indicates the outcome passes the Benjamini-Hochberg test for multiple comparisons with a 10 percent false discovery rate.

Note: The derivation of the variable, Parent Read to Child in Last Week, was revised for this report. The findings in this report may differ from the findings in the *Head Start Impact Study: First Year Findings* (U.S. Department of Health and Human Services, 2005). See Chapter 2 and the Technical Report for the Head Start Impact Study for details of this change.

Impacts on Participants

This section discusses Head Start's impact on parenting practices for parents of children who *actually participated in the program.*[106] Moving from measures of the average impact of access to Head Start (ITT estimates) to the average impact of participating (IOT estimates) changes all findings by about 50 percent. For example within the 3-year-old cohort, the impact of Head Start on the probability of a parent spanking his/her child at the end of the Head Start

[106] Detailed findings are presented in Appendix E. As noted in Chapter 4, some children granted access (e.g., Head Start group) never participated in Head Start, while other children randomized into the study's control group did participate.

year is -0.07 (ITT estimate in Exhibit 7.2) while the impact of participating in Head Start on the probability of a parent spanking his/her child is -0.10 (IOT estimate in Exhibit E7.2 in Appendix E). All changes in magnitude for participants follow this proportion for the 3-year-old cohort. A similar change occurs for 4-year-old impact magnitudes. The statistical significance does not change when moving from access to participation (i.e., any statistically significant impact of access to Head Start is also statistically significant for the impact of participation). However, wherever statistically significant impacts were found, the effect of Head Start is larger than found in the ITT estimates.

The larger magnitude of impact from participation, as opposed to access, can be summarized in terms of effect sizes. For example, impacts in the parenting domain confirmed by strong evidence[107] have effect sizes ranging from 0.20 to 0.26 (positive or negative) for the impact of participation, in contrast to effect sizes of 0.14 to 0.18 (positive or negative) for the impact of access.

For instance, for the 3-year-olds at the end of the Head Start year, the effect size for the reduction in spanking was -0.14 for the ITT (impact of access) estimates and -0.20 for the IOT (participation) estimates. Likewise, the increases in parents' reading to their children had an effect size of 0.15 for the ITT estimates and 0.22 for the IOT estimates. Although the effect sizes are larger, the effect sizes for participation do not change the conclusions reported for the access or ITT impact estimates.

Summary

The second objective in the Head Start Program Performance Measures is to "strengthen families as the primary nurturers of their children" (U.S. Department of Health and Human Services, 2001). To achieve this objective, Head Start works with parents to improve their parenting skills through targeted enhancement of parents' abilities to interact with their children in positive and supportive ways and to provide age-appropriate discipline.

Impacts on parenting are limited for the 4-year-old cohort; however, there is evidence that Head Start improved parenting for the 3-year-old cohort, with impacts on parenting styles,

[107] The discussion on effect sizes is limited to the impacts confirmed by strong evidence for illustration only. All impacts have similar increases in their effect sizes.

parental discipline strategies, and educational support activities in the home that encourage the social and educational development of children.

For the 4-year-old cohort, there was a single impact on parenting: parents in the Head Start group were less likely to use time out at the end of the Head Start year than parents in the control group.

For the 3-year-old cohort, impacts on parenting were more prevalent. Head Start decreased authoritarian parenting both in the age 4 year and at the end of 1st grade. Head Start also led to less frequent use of spanking at the end of the Head Start year and kindergarten and less time out at the end of kindergarten and 1st grade. Further, Head Start increased the likelihood that parents were reading to their children and engaging in cultural enrichment activities at the end of the Head Start year.

Research has demonstrated that reading to children has a positive effect on their literacy outcomes (Denton, Reaney & West, 2001; Snow, Burns, & Griffin, 1998). Although, Head Start group parents report reading to their child, the percentage is less often than that reported by parents of kindergarten-age children in the general population. In the ECLS-K study (U.S. Department of Education, 2000), about 45 percent of parents reported reading to their children every day and 80 percent of parents reported reading to their children at least three times per week. For the Head Start Impact Study at the end of kindergarten, 35 percent of Head Start group parents of children in the 4-year-old cohort reported reading to their children every day while 69 percent of these parents reported reading to their children at least three times per week. For the 3-year-old cohort, 34 percent of the Head Start group parents reported reading to their children every day while 65 percent of these parents reported reading to their children at least three times per week.

Decreases in the use of an authoritarian parenting style and spanking and increases in reading and cultural enrichment are favorable findings for the children provided access to Head Start at age three. However, the interpretation of impacts on time out is less clear. It is possible that these decreases suggest that parents are using less positive disciplinary approaches. Yet, it is possible that the reductions in time out reflect the improvements in children's behavior and parent-child relationships reported in the social emotional chapter.

Chapter 8: Variation in Impacts: Child and Parent Subgroups

Summary of Approach and Findings

Previous chapters presented impact estimates for the effect of Head Start on the average child randomly assigned to the Head Start group. This chapter looks instead for differences in impacts among different types of children and parents to respond to the congressional mandate that the Head Start Impact Study look for *". . .possible sources of variation in impact of the Head Start program."*

For this analysis, seven dimensions were used to define subgroups: (1) whether a child had low pre-academic skills at the start of Head Start (referred to as children in the lowest quartile), (2) whether the child was a Dual Language Learner at the start of Head Start, (3) whether the child had special needs (as reported by the parent at the start of Head Start), (4) biological mother's/caregiver's race/ethnicity, (5) reported level of depressive symptoms for the child's parent/caregiver, (6) a composite index of household risks, and (7) urbanicity. These dimensions were based on data collected at the time of random assignment.

This chapter highlights patterns of impacts for these seven dimensions. There is no scientific consensus for what constitutes a pattern of impacts. Yet, given the large number of comparisons tested (almost 10,000, taking into consideration the study's two cohorts, four time points for measuring outcomes, and multiple outcomes), 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, this chapter concentrates on *differential impacts*, that is, impacts where there was a statistically significant difference in Head Start's effects for one subgroup compared to another. We pay particular attention to multiple impacts that occur across domains, or outcomes, or

persist into kindergarten or 1st grade. Additionally, any differences between subgroups that meet the standard of strong evidence[108] are noted in Exhibits 8.4 through 8.17.

A brief summary of the key subgroup findings sustained beyond the Head Start years follows:

4-Year-Old Cohort

In the 4-year-old cohort, several subgroups showed both: (1) differential impacts from other subgroups and (2) a pattern of positive impacts across domains, outcomes, and/or years:

- **Children of parents with reported mild depressive symptoms** experienced favorable cognitive impacts through the end of 1st grade.
- **Black children** experienced favorable impacts in the social-emotional domain at the end of kindergarten as reported by teachers.
- **Dual Language Learners** experienced health benefits at the end of kindergarten and 1st grade, although early impacts were mixed.
- **Children in the lowest academic quartile at baseline** showed benefits of Head Start in the social-emotional domain through the end of 1st grade.

Some subgroups in the 4-year-old cohort experienced mixed results or a pattern of unfavorable impacts. For example, White children experienced unfavorable impacts on several teacher-reported social-emotional measures at the end of 1st grade, one unfavorable impact each in the cognitive and parenting domains at the end of kindergarten, and an unfavorable impact in the health domain at the end of the Head Start year.

[108] Due to the large number of statistical tests, the following language was developed to report categories of statistically significant results:

- *Strong Evidence*: the estimated impact for a particular outcome is statistically significant at the typical level ($p \leq 0.05$), and this result holds up under the test for multiple comparisons (i.e., Benjamini-Hochberg procedure).
- *Moderate Evidence*: the estimated impact for a particular outcome is statistically significant at the typical level ($p \leq 0.05$), but this result *does not* hold up under the test for multiple comparisons.
- *Suggestive Evidence*: the estimated impact for a particular outcome is statistically significant under a relaxed standard ($p \leq 0.10$), and this result *may or may not* hold up under the test for multiple comparisons.

3-Year-Old Cohort

More subgroup patterns were found for the 3-year-old cohort, where several subgroups showed both: (1) differential impacts from other subgroups and (2) a pattern of positive impacts across domains, outcomes and/or years:

- **Children with special needs** benefited from Head Start in the math and social-emotional areas at the end of 1st grade.
- **Children of parents with no reported depressive symptoms** experienced sustained benefits of Head Start in the cognitive, social-emotional, and parenting domains through the end of 1st grade.
- **Children from high risk households** showed sustained favorable cognitive impacts through the end of 1st grade.
- **Children in non-urban settings** showed long-term cognitive benefits from Head Start through the end of 1st grade and some benefits in the social-emotional domain during the Head Start years.
- **Dual Language Learners** benefited in the cognitive domain through the end of kindergarten.

Three subgroups showed sustained negative impacts:

- **Children of parents with reported moderate depressive symptoms** experienced sustained negative impacts of Head Start in the cognitive, social-emotional, and health domains and mixed impacts in the parenting domain. These children were more often not promoted to the next grade as reported by their parents. Where more than one method was used to assess, these negative impacts were found across reporters or types of assessment.
- **Children of parents with reported mild depressive symptoms** experienced unfavorable impacts at the end of kindergarten and 1st grade in the social-emotional domain as reported by teachers.
- **Children from moderate risk households** experienced sustained positive benefits in the social-emotional domain during most years, yet at the same time experienced several sustained negative impacts in the cognitive domain during their school years.

Exhibits 8.1 and 8.2 provide a summary of the subgroup findings by age cohort.

Exhibit 8.1: Summary of the Subgroup Findings for the 4-Year-Old Cohort

Dimensions	Subgroup	Year	Cognitive			Social-Emotional		Health	Parenting	
			Language and Literacy	Math Skills	School Performance	Parent Report	Teacher Report	Parent Report	Parent Report	Teacher Report
Child's Pre-Academic Skills	In Lowest Quartile	HS			n/a	2/9	n/a			n/a
		K	1/9	1/3			4/11		1/9 1/9	
		1			1/5		2/11		1/8	
	Not in Lowest Quartile	HS			n/a		n/a			n/a
		K								
		1		1/4						
Child's Home Language	English Speaking	HS			n/a		n/a			n/a
		K								
		1								
	Dual Language Learners	HS	1/9		n/a		n/a	1/5 1/5	1/5	n/a
		K			1/5			1/5		
		1						1/5	1/8	
Special Needs	Special Needs	HS			n/a		n/a			n/a
		K				1/9	1/11	1/5		
		1						1/5		
	Not Special Needs	HS			n/a		n/a			n/a
		K					1/11			
		1								
Biological Mother Race/ Ethnicity	White	HS			n/a		n/a	1/5		n/a
		K	1/9				1/11		1/9 1/9	
		1					3/11			
	Black	HS	1/9		n/a		n/a			n/a
		K	2/9				5/11			
		1						1/5	2/8	
	Hispanic	HS	1/9	1/2	n/a		n/a			n/a
		K							1/9	
		1				1/9			1/8	

Key:

Numeral indicates the number of significant outcomes out of the total number of outcomes for that cell.

Yellow cell indicates significant favorable impact(s).

Red cell indicates significant unfavorable impact(s).

Light blue cell indicates a significant impact but the meaning is unclear.

n/a indicates not applicable (i.e., data were not collected for this cell).

Exhibit 8.1: Summary of the Subgroup Findings for the 4-Year-Old Cohort (continued)

Dimensions	Subgroup	Year	Cognitive			Social-Emotional		Health	Parenting	
			Language and Literacy	Math Skills	School Performance	Parent Report	Teacher Report	Parent Report	Parent Report	Teacher Report
Parent Depressive Symptoms	No Symptoms	HS	1/9		n/a		n/a			n/a
		K				1/9			1/9	
		1								
	Mild Symptoms	HS	3/9	1/2	n/a	2/9	n/a			n/a
		K				1/9	1/11			
		1	3/11	1/4		1/9				
	Moderate Symptoms	HS	1/9		n/a		n/a			n/a
		K						1/5		1/2
		1					1/11			1/2
	Severe Symptoms	HS	3/9		n/a		n/a	1/5		n/a
		K				1/9		1/5	1/9	
		1		1/4		1/9	1/11	2/5	1/8	
Household Risk	Low/No Risk	HS	2/9		n/a	1/9	n/a	1/5 1/5		n/a
		K							1/9	
		1					1/11		1/8	
	Moderate Risk	HS	1/9	1/2	n/a		n/a			n/a
		K					1/11 1/11			
		1	1/11					1/5	1/8	
	High Risk	HS			n/a	2/9	n/a	1/5		n/a
		K				1/9	2/11			
		1					1/11	1/5		
Urbanicity	Urban	HS	1/9	1/2	n/a		n/a	1/5		n/a
		K								
		1								
	Not Urban	HS	1/9		n/a		n/a	1/5	1/5	n/a
		K	1/9			1/9				1/2
		1			1/5					1/2

Key:

Numeral indicates the number of significant outcomes out of the total number of outcomes for that cell.

Yellow cell indicates significant favorable impact(s).

Red cell indicates significant unfavorable impact(s).

Light blue cell indicates a significant impact but the meaning is unclear.

n/a indicates not applicable (i.e., data were not collected for this cell).

Exhibit 8.2: Summary of the Subgroup Findings for the 3-Year-Old Cohort

Dimensions	Subgroup	Year	Cognitive			Social-Emotional		Health	Parenting	
			Language and Literacy	Math Skills	School Performance	Parent Report	Teacher Report	Parent Report	Parent Report	Teacher Report
Child's Pre-Academic Skills	In Lowest Quartile	HS			n/a	1/9	n/a			n/a
		Age 4	4/9	1/2	n/a	1/9	n/a		2/9	n/a
		K								
		1								
	Not in Lowest Quartile	HS			n/a		n/a		1/5	n/a
		Age 4			n/a		n/a			n/a
		K					2/11			
		1								
Child's Home Language	English Speaking	HS	1/9		n/a	1/9	n/a	1/5	1/5	n/a
		Age 4			n/a	1/9	n/a			n/a
		K		2/3						
		1							1/8	
	Dual Language Learners	HS	2/9		n/a		n/a	2/5		n/a
		Age 4	5/9	1/2	n/a	1/9	n/a			n/a
		K	2/9					1/5		
		1								1/2
Special Needs	Special Needs	HS			n/a		n/a		1/5	n/a
		Age 4			n/a		n/a	1/5	1/9	n/a
		K							1/9 2/9	
		1		4/4			4/11		1/8	
	Not Special Needs	HS			n/a		n/a			n/a
		Age 4			n/a		n/a			n/a
		K								
		1								

Key:

Numeral indicates the number of significant outcomes out of the total number of outcomes for that cell.

Yellow cell indicates significant favorable impact(s).

Red cell indicates significant unfavorable impact(s).

Light blue cell indicates a significant impact but the meaning is unclear.

n/a indicates not applicable (i.e., data were not collected for this cell).

Exhibit 8.2: Summary of the Subgroup Findings for the 3-Year-Old Cohort (continued)

Dimensions	Subgroup	Year	Cognitive			Social-Emotional		Health	Parenting	
			Language and Literacy	Math Skills	School Performance	Parent Report	Teacher Report	Parent Report	Parent Report	Teacher Report
Biological Mother Race Ethnicity	White	HS	1/9		n/a	3/9	n/a	1/5 1/5		n/a
		Age 4	1/9		n/a	1/9	n/a			n/a
		K	1/9			1/9		1/5	1/9	
		1	1/11			1/9 2/9		1/5		
	Black	HS	1/9		n/a	1/9	n/a	1/5		n/a
		Age 4			n/a		n/a			n/a
		K	1/9	1/3				1/5	2/9 1/9	
		1				4/9			1/8 1/8	
	Hispanic	HS	1/9		n/a		n/a	1/5 1/5		n/a
		Age 4	2/9	1/2	n/a		n/a			n/a
		K	1/9				1/11		3/9	
		1		1/4		1/9 1/9	1/11			1/2
Parent Depressive Symptoms	No Symptoms	HS	1/9		n/a	1/9	n/a		1/5	n/a
		Age 4	5/9	1/2	n/a	2/9	n/a		1/9	n/a
		K	1/9			5/9			1/9	
		1	7/11	2/4			1/11		1/8	
	Mild Symptoms	HS			n/a	1/9	n/a			n/a
		Age 4	1/9		n/a		n/a			n/a
		K				1/9	3/11		1/9	
		1			3/5		3/11			
	Moderate Symptoms	HS			n/a		n/a		1/5	n/a
		Age 4	4/9		n/a	1/9	n/a	2/5		n/a
		K			2/5	1/9	5/11		1/9 1/9 1/9	1/2
		1	9/11		1/5		2/11	1/5		
	Severe Symptoms	HS	1/9		n/a	1/9	n/a	2/5		n/a
		Age 4	1/9		n/a	1/9	n/a			n/a
		K			1/5					
		1					2/11	1/5		

Key:

Numeral indicates the number of significant outcomes out of the total number of outcomes for that cell.

Yellow cell indicates significant favorable impact(s).

Red cell indicates significant unfavorable impact(s).

Light blue cell indicates a significant impact but the meaning is unclear.

n/a indicates not applicable (i.e., data were not collected for this cell).

Exhibit 8.2: Summary of the Subgroup Findings for the 3-Year-Old Cohort (continued)

Dimensions	Subgroup	Year	Cognitive			Social-Emotional		Health	Parenting	
			Language and Literacy	Math Skills	School Performance	Parent Report	Teacher Report	Parent Report	Parent Report	Teacher Report
Household Risk	Low/No Risk	HS	2/9	1/2	n/a		n/a	1/5		n/a
		Age 4	3/9		n/a		n/a			n/a
		K								
		1								
	Moderate Risk	HS			n/a	1/9	n/a	1/5		n/a
		Age 4			n/a	1/9	n/a	1/5		n/a
		K	1/9	2/3						
		1	1/11	1/4		3/9	1/11		1/8	
	High Risk	HS	2/9		n/a		n/a	1/5		n/a
		Age 4	1/9		n/a		n/a	1/5		n/a
		K	1/9				2/11			
		1	5/11							
Urbanicity	Urban	HS	3/9		n/a	1/9	n/a	1/5		n/a
		Age 4			n/a		n/a			n/a
		K								
		1								
	Not Urban	HS	4/9		n/a	3/9	n/a	1/5	1/5	n/a
		Age 4		1/2	n/a	1/9	n/a	1/5		n/a
		K	1/9					1/5		
		1	6/11	1/4						

Key:
Numeral indicates the number of significant outcomes out of the total number of outcomes for that cell.
Yellow cell indicates significant favorable impact(s).
Red cell indicates significant unfavorable impact(s).
Light blue cell indicates a significant impact but the meaning is unclear.
n/a indicates not applicable (i.e., data were not collected for this cell).

Introduction

In addition to an interest in the overall national impact of Head Start on children's school readiness, Congress mandated an examination of how impacts vary for different types of children and families. Identifying subgroups that benefit more or less from Head Start may have important policy and program implications. It provides a sense of Head Start's strengths and it can suggest areas where the program needs to be strengthened or enhanced to ensure that all participants advance in their development. This analysis examines patterns among subgroups to identify who benefits from Head Start in the cognitive, social-emotional, health and parenting domains.

Subgroups Examined

With regard to participant characteristics, the following seven dimensions were identified as being of primary policy interest and used to create subgroups (all were measured at baseline):

- Child's Pre-Academic Skills—based on whether the child scored in the lowest quartile of the study population on the baseline assessment of the Woodcock-Johnson III Pre-Academic Skills (comprising of three tests: Letter-Word Identification, Spelling, and Applied Problems). Two subgroups were created using this test score: the child was in the lowest quartile subgroup, or the child was not in the lowest quartile subgroup.
- Child's home language—based on the language in which the child was assessed for the baseline assessment in fall 2002. Two subgroups were created: the child was English speaking, or the child was a Dual Language Learner (See Chapter 2 and the Technical Report for the Head Start Impact Study for how the language for the baseline assessment was determined.)
- Special needs—based on the parent's response to the following question on the baseline interview, "*Did a doctor or other health or education professional ever tell you that [CHILD] has any special needs or disabilities—for example, physical, emotional, language, hearing, learning difficulty, or other special needs?"* Two subgroups were created: the child was reported to have special needs, or the child was not reported to have special needs.
- Biological mother/caregiver race/ethnicity—based on the race of the person identified as being most responsible for the care of the child at the time of the baseline parent interview.[109] Three categories were created: White or other,[110] Black, and Hispanic.

[109] The primary caregiver is the child's biological mother for 96 percent of the study children.

[110] Other race (N=94 for the 3-year-old cohort and N=85 for the 4-year-old cohort) was combined with White because the number of other race respondents was too small to study independently.

- Parent/caregiver-reported depressive symptoms—determined from responses to the baseline parent/caregiver interview using the shortened version (12 items) of the Center for Epidemiologic Studies-Depression scale (CES-D) (Seligman, 1993[111]). Four subgroups were created from the scale: (1) no depressive symptoms (score of 0-4), (2) mild depressive symptoms (score of 5-9), (3) moderate depressive symptoms (score of 10-14), and (4) severe depressive symptoms (score of 15-36).
- Household risk index—determined by the number of the following characteristics reported in the baseline parent interview: (1) receipt of TANF or Food Stamps, (2) neither parent in household has high school diploma or a GED, (3) neither parent in household is employed or in school, (4) the child's biological mother/caregiver is a single parent, and (5) the child's biological mother was age 19 or younger when child was born. A child's family score could range from 0 to 5 points. Three categories were created: low/no risk (0-2 risk factors), moderate risk (3 risk factors), and high risk (4-5 risk factors).
- Urbanicity—based on the location of the Head Start center at which the family applied for admission. If the center was located in a Census-defined urbanized area, the family was considered to live in an urban area; if not, the family was considered not to live in an urban area. Thus, two subgroups were defined.

Before presenting the findings, Exhibit 8-3 provides the distribution of children across the subgroups by age cohort and status as a part of the Head Start group or the control group.

[111] The four depressive symptoms categories are reported on page 101 in the above reference for the 20 item CES-D. The cut points were proportionately adjusted for the shortened version of the CES-D for use in ECLS-B, FACES, and HSIS.

Exhibit 8.3: Distribution of Children Across the Subgroups by Age Cohort and Random Assignment Status

Subgroup	4-Year-Old Cohort		3-Year-Old Cohort	
	Head Start Group	Control Group	Head Start Group	Control Group
Child's Pre-Academic Skills				
Lowest Quartile	249 (24%)	159 (26%)	312 (24%)	189 (25%)
Non-lowest Quartile	775 (76%)	455 (74%)	987 (76%)	582 (75%)
Child's Home Language				
English Speaking	695 (68%)	418 (68%)	996 (77%)	593 (77%)
Dual Language Learners	329 (32%)	196 (32%)	303 (23%)	178 (23%)
Special Needs				
Special Needs	146 (14%)	74 (12%)	171 (13%)	77 (10%)
Non-special Needs	878 (86%)	540 (88%)	1,128 (87%)	694 (90%)
Biological Mother/Caregiver Race/Ethnicity				
White Children	374 (37%)	220 (36%)	396 (31%)	276 (36%)
Black Children	229 (22%)	134 (22%)	486 (37%)	256 (33%)
Hispanic Children	421 (41%)	260 (42%)	417 (32%)	239 (31%)
Parent/Caregiver Reported Depressive Symptoms				
No Depressive Symptoms	478 (50%)	279 (51%)	635 (53%)	377 (56%)
Mild Depressive Symptoms	250 (26%)	144 (27%)	310 (26%)	155 (23%)
Moderate Depressive Symptoms	127 (13%)	65 12%)	149 (12%)	83 (12%)
Severe Depressive Symptoms	98 (10%)	55 (10%)	107 (9%)	62 (9%)
Household Risk Index				
Low/No Household Risk	744 (73%)	456 (74%)	956 (74%)	568 (74%)
Moderate Household Risk	204 (20%)	110 (18%)	234 (18%)	145 (19%)
High Household Risk	76 (7%)	48 (8%)	109 (8%)	58 (7%)
Urbanicity				
Urban	872 (85%)	530 (86%)	1,077 (83%)	629 (82%)
Not Urban	152 (15%)	84 (14%)	222 (17%)	142 (18%)

Note: Numbers are based on the spring 2003 analysis sample. Due to rounding, the sum of the percents may not equal to 100 percent.

▪ Outcome Domains

As discussed in Chapter 2 and the Technical Report for the Head Start Impact Study, the outcomes are organized into four domains:

Cognitive Outcomes

- Direct child assessments of language and literacy and math skills;
- School performance, reported by teachers and parents;

Social-Emotional Outcomes

- Social-emotional measures provided by parents and teachers;

Health Outcomes[112]

- Health measures, using information obtained from parent reports;

Parenting Outcomes[113]

- Parenting measures provided by parents and teachers.

Analytic Approach

The determination of what constitutes a pattern of differential impacts between subgroups is not simple, and there is no scientific consensus on how best to make these determinations. Each of the seven dimensions comprises at least two separate subgroups of children (e.g., children with and without special needs), and there are at least three statistical tests conducted on each outcome for each subgroup: (e.g., *within group* impacts tested for children with and without special needs separately, as well as a test of the difference in impacts between children with and without special needs, the *difference of difference* test[114]). For dimensions that have more than two subgroups of children, such as race/ethnicity, the number of tests is even greater. All subgroups were analyzed for the 3- and 4-year-old cohorts separately.

Given the many outcome measures, the multiple years of data collection, the two age cohorts, and the seven subgroup dimensions, the number of statistical tests for these analyses was extremely large (nearly 10,000). When so many statistical tests and analyses are conducted, it is

[112] The interpretation of two health outcomes (child needs ongoing medical care and medical care for a child injury in the last month) is unclear. The outcome indicating that the child needs ongoing medical care could be interpreted as a positive outcome (i.e., the parent is aware of an illness that requires ongoing care). However, it may also indicate an increase in conditions needing medical care. The outcome indicating that the child needed medical care for an injury could be interpreted as an unfavorable outcome, indicating that the child is more likely to be injured. Alternatively, it may indicate that parents are more likely to get medical attention when the child has injuries, which would be a favorable outcome.

[113] The interpretation of one parenting outcome (use of time out in the last week) is unclear. A reduction in the use of time out may indicate that the behavior of children is improving, so there is less need for time out, or that the parent is using other, less positive disciplinary methods.

[114] This is referred to as a difference of difference test because of the nature of the comparison. For each subgroup within the given subgroup dimension—children with and without special needs—there is first a test of the *within-group impact.* This test analyzes the difference between the Head Start and control groups for special needs children only or for non-special needs children only. Then these within-group impacts are compared to one another, resulting in a test of the difference of those impacts. The difference in difference test tells us whether an impact is significantly larger for one subgroup than for another.

important to guard against Type I errors, statistically significant findings that reflect chance variations rather than true differences. As discussed in Chapter 2, the Benjamini-Hochberg procedure was used to address this problem. Findings from the Benjamin-Hochberg procedure, using a 10 percent false discovery rate, are highlighted in Exhibits 8.4 through 8.17. However, since subgroups are smaller in size and therefore have lower power than the full sample to detect true effects, the risk of Type II error (i.e., the finding that there are no differential impacts between subgroups when in fact there are) is greater, and the use of multiple comparison procedures increases the risk of a Type II error. Due to these limitations, 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.

We have aimed to make this chapter useful and readable by being more selective in both the exhibits and narrative presented here. Accordingly, the exhibits include only those outcomes with both a statistically significant difference in impacts between subgroups and a statistically significant impact for at least one subgroup in the comparison. The text highlights patterns in the results, such as differences in impact across domains, outcomes, or years. Detecting patterns is more of an art than a science, and different individuals may see different findings in these results. All the subgroup findings, including effect sizes, are available on the Administration for and Families, Office of Planning, Research and Evaluation Website at: http://www.acf.hhs.gov/programs/opre/hs/impact_study/index.html. Readers may choose to review these findings themselves and draw their own conclusions.

Exhibit Entries

The first seven exhibits (Exhibits 8.4 through 8.10) show the results for each subgroup for the 4-year-old cohort. The second seven exhibits (Exhibits 8.11 through 8.17) show the results for each subgroup for the 3-year-old cohort. Each exhibit is organized around the four outcome domains, corresponding with the groupings for which the Benjamini-Hochberg test was conducted. As noted above, exhibit entries are limited to impacts that (1) significantly differ in magnitude between subgroups being compared and (2) significantly differ from zero for at least one of the subgroups being compared. Within this framework, each of the exhibits provides the following information: (1) the number of outcome variables included in each sub-domain for each subgroup and year (in parentheses in the first column), (2) the outcome variables for which

evidence exists of significant *differences in impact between subgroups* ($p \leq 0.10$), in instances in which at least one of the subgroups involved had a significant non-zero impact itself, (3) the regression-adjusted estimate of the magnitude of impact on each of the individual subgroups involved and the statistical significance of these estimates (in the third and fourth columns and fifth or sixth columns when needed), and (4) the statistical significance of these differences in impact between subgroups (in the last column). Items (3) and (4) appear in bold for findings that hold up under the Benjamini-Hochberg test for multiple comparisons and hence meet the standard of strong evidence of program impact.

Differential Impacts for the 4-Year-Old Cohort

Child's Pre-Academic Skills. The impacts on children in the lowest quartile of pre-academic skills at baseline differed from those on other children in ways that varied by domain. Exhibit 8.4 shows that access to Head Start benefited children in the lowest quartile more than other children in the social-emotional domain through the end of 1st grade. At the end of the Head Start year, there were favorable impacts for the lowest quartile children compared to their non-lowest quartile counterparts on parent's report of their relationship with the child. In the school years, teacher reports showed more favorable impacts for lowest quartile children on oppositional behavior, problems with peer interaction, conflict, and positive relationships with the teacher than non-lowest quartile children.

There were fewer differential impacts by children's initial pre-academic skills in the cognitive and parenting domains (there were none in health). However, those that did exist were unfavorable for children in the lowest quartile in the cognitive domain (e.g., poorer math skills), and the impacts were mixed in the parenting domain.

The effect sizes[115] for the significant within subgroup differences for the social-emotional domain range from 0.24 to 0.43 for children in the lowest quartile, with the majority of effect sizes greater than or equal to 0.30

Child's Home Language. Exhibit 8.5 shows the differences in the effect of Head Start for subgroups defined by the child's home language. There is some evidence that Dual

[115] The effect sizes for all impact differentials and within-subgroup differences are reported on the Administration for Children and Families, Office of Planning, Research and Evaluation Website at: http://www.acf.hhs.gov/programs/opre/hs/impact_study/index.html..

Language Learners benefited more in the health domain than their English-speaking counterparts. Head Start had a favorable impact on the health insurance coverage of Dual Language Learners at the end of the Head Start and kindergarten years and on the receipt of dental care at the end of 1st grade. However, Head Start had an unfavorable impact on the health status of Dual Language Learners at the end of the Head Start year.

There was no evidence of differential patterns for Dual Language Learners in the cognitive, parenting or social-emotional domains, although there were a few statistically significant impacts.

The effect sizes for the significant within subgroup differences for the health domain range from 0.17 to 0.32 for Dual Language Learners, with the majority of effect sizes less than 0.24.

Special Needs. Exhibit 8.6 indicates little evidence of differential impacts for children with and without special needs.

Biological Mother/Caregiver Race/Ethnicity. As shown in Exhibit 8.7, there was a pattern of more favorable impacts from Head Start for Black children in the social-emotional domain at the end of kindergarten as reported by teachers. Black children in the Head Start group were reported to have reduced inattentiveness, fewer problems with structured learning, peer interactions, or teacher interactions; and better relationship with teachers. The majority of these impacts were significantly different from those on White and Hispanic children.

In the other domains, there was little evidence of a consistent pattern, despite a few favorable impacts for Blacks in these domains.

In contrast to a more positive picture for Black children, White children were more likely to experience negative impacts of Head Start at the end of 1st grade, particularly in the social-emotional domain. At the end of 1st grade, teachers reported unfavorable impacts on withdrawn behavior for White children compared to Black children and teachers also reported unfavorable impacts on shy/socially reticent behavior and problems with teacher interactions for White children compared to Hispanic children.

The effect sizes for the significant within subgroup differences for the social-emotional domain range from 0.36 to 0.57 for Black children, with the majority of effect sizes equal to or

greater than 0.48. For White children, the effect sizes for significant impacts for the social-emotional domain range from 0.29 to 0.37.

Parent/Caregiver-Reported Depressive Symptoms. Children of parents with reported mild depressive symptoms showed more favorable cognitive benefits at the end of Head Start and 1st grade than children of parents with other levels of reported depressive symptoms, as presented in Exhibit 8.8. In the cognitive domain, each subgroup showed some favorable impacts at the end of the Head Start year, but the clearest pattern found was for children whose parents reported mild depressive symptoms. At the end of the Head Start year, this subgroup experienced several benefits of Head Start compared to children of parents with other levels of reported depressive symptoms. These impacts included improvements in the Counting Bears and Letter Naming tasks and more improved scores on the Woodcock-Johnson III Pre-Academic Skills and Letter-Word Identification. No cognitive impacts were found in kindergarten for any of the subgroups, but benefits re-appeared for children of parents with reported mild depressive symptoms at the end of 1st grade. Impacts in the other domains were mixed or neutral for the other subgroups.

The effect sizes for the significant within subgroup differences for all domains as well as the cognitive domain range from 0.20 to 0.41 for children of parents who reported mild depressive symptoms, with the majority of effect sizes less than or equal to 0.32.

Household Risk Index. Exhibit 8.9 presents differential impacts for different household risk subgroups across all domains. There was no pattern of differential impacts for this subgroup dimension.

Urbanicity. As shown in Exhibit 8.10, there were few differential impacts related to urbanicity, and those that were found were often conflicting within and across domains.

Differential Impacts for the 3-Year-Old Cohort

Child's Pre-Academic Skills. Children who entered Head Start in the lowest quartile of pre-academic skills experienced more favorable cognitive impacts than other children, although these differences were found only in the age 4 year. There were few differential impacts in the social-emotional, health, or parenting domains noted in Exhibit 8.11.

The effect sizes for the significant within subgroup differences for the cognitive domain range from 0.14 to 0.23 for children in the lowest quartile.

Child's Home Language. Exhibit 8.12 illustrates that Head Start had more favorable impacts on Dual Language Learners than on English-speaking children in the cognitive domain, though the impacts lasted only through the end of kindergarten. The following favorable impacts were significant in the cognitive domain at the end of the age 4 year for Dual Language Learners (CTOPPP Elision, Letter Naming, Woodcock-Johnson III Pre-Academic Skills, Woodcock-Johnson III Letter-Word Identification, Counting Bears, and Woodcock-Johnson III Spelling). Some of these favorable cognitive impacts continued into kindergarten but fade by 1st grade.

There were fewer favorable differential impacts in the social-emotional, health, and parenting domains for Dual Language Learners.

The effect sizes for the significant within subgroup differences for the cognitive domain range from 0.18 to 0.42 for Dual Language Learners with most effect sizes less than 0.34.

Special Needs. Special needs children experienced benefits of Head Start at the end of 1st grade, particularly in the cognitive and social-emotional domains (Exhibit 8.13) compared to non-special needs children. Of particular interest is the fact that these differences between special needs and non-special needs children did not appear until the 1st grade.

In the cognitive domain, special needs children experienced favorable impacts on all direct assessments of math skills conducted in the 1st grade compared to other children. These favorable impacts were experienced on the test scores for the Woodcock-Johnson III tests of Applied Problems, Calculation, Quantitative Concepts, and the Math Reasoning.

According to teacher reports, special needs children also experienced favorable impacts from Head Start in the social-emotional domain at the end of 1st grade. As a result of access to Head Start, special needs children showed a reduction in inattention/hyperactivity, in problems with structured learning, and in conflict with teachers as well as an increase in positive teacher relationships. Children without special needs did not experience any of these benefits.

There were also several parenting impacts for special needs children through 1st grade, although their interpretation is unclear. Statistically significant impacts of Head Start on special needs children's parenting experiences were found in all years of the study. However, some of

these impacts appeared to be favorable (e.g., less spanking), whereas others were unfavorable or unclear (e.g., less reading to child, more authoritarian and less authoritative parenting styles, less time out).

For children with special needs, the effect sizes for the significant within subgroup differences range from 0.34 to 0.45 for the cognitive domain, 0.37 to 0.48 for the social-emotional domain, and 0.25 to 0.61 for the parenting domain, with most under 0.38 in the parenting domain.

Biological Mother/Caregiver Race/Ethnicity. There was no clear pattern of benefit for one racial or ethnic subgroup over the others as presented in Exhibit 8.14. In fact, each of the subgroups appears to benefit in some way, and often these benefits differed from the other subgroups.

Parent/Caregiver-Reported Depressive Symptoms. Exhibit 8.15 shows a pattern that Head Start had lasting favorable impacts in the cognitive, social-emotional, and parenting domains for children of parents with no reported depressive symptoms compared to other children. There is also reason for concern for children of parents who reported moderate depressive symptoms, as there were many lasting unfavorable impacts in all domains for this subgroup, including a decrease in promotion to the next grade. Additionally, children of parents with reported mild depressive symptoms experienced unfavorable impacts in the cognitive and social-emotional domain at the end of kindergarten and 1st grade as reported by teachers.

In the cognitive domain, children of parents with no reported depressive symptoms benefited from Head Start on many direct assessments of language, literacy, and math skills in all years, and especially at the end of 1st grade. The majority of cognitive differential impacts for children of parents with no reported depressive symptoms was significantly different from the impacts on children of parents with moderate depressive symptoms and sometimes children of parents with reported mild depressive symptoms.

Head Start also had lasting favorable social-emotional impacts for the children of parents with no reported depressive symptoms compared to the children of parents in other reported depressive symptoms subgroups. For example, parent reports suggested that Head Start reduced hyperactive behavior for children of parents with no reported depressive symptoms at the end of the Head Start years and kindergarten.

In the parenting domain, one favorable impact was noted at the end of each year for children of parents with no reported depressive symptoms relative to other children. Impacts on these variables were neutral for the other subgroups. At the end of the Head Start year, children of parents with no reported depressive symptoms benefited more from Head Start on parental use of safety practices than children of parents with reported mild or moderate depressive symptoms.

In contrast to these more positive patterns, we found consistent, sustained, negative, cognitive impacts for the subgroup of children whose parents reported moderate depressive symptoms. These children experienced several negative impacts on direct assessments of language and literacy skills, as well as teachers' reports of academic ability and progress especially in 1st grade. Additionally, Head Start group parents reporting moderate depressive symptoms were more likely to report that their child was not promoted into second grade than control group parents in this subgroup.

Furthermore, in the social-emotional domain, parents and teachers both reported unfavorable impacts for this subgroup compared to other children at the end of the Head Start and school years. In the health domain, the impacts were mixed. There were unfavorable health impacts at the end of the age 4 year and one impact at the end of 1st grade.* Finally, within this subgroup of children, kindergarten teachers reported that parents with moderate depressive symptoms were in less communication with the school than parents with other reported levels of reported depressive symptoms. However, children of parents with moderate depressive symptoms benefited from Head Start on increased reading to the child at the end of the Head Start year (compared to children of parents with reported severe depressive symptoms) and at the end of kindergarten (compared to children of parents with no reported depressive symptoms).

For all domains, the effect sizes for the significant within subgroup differences range from 0.12 to 0.29 for children of parents with no reported depressive symptoms, 0.20 to 0.33 for children of parents with reported mild depressive symptoms, and 0.21 to 0.54 for children of parents with reported moderate depressive symptoms, with most less than 0.45 for this subgroup.

Household Risk Index. Exhibit 8.16 presents repeated evidence of differential impacts that vary by domain for households at different levels of economic risk.

* Interpretation of this impact is unclear.

Children from high risk households experienced sustained, favorable benefits from Head Start in the cognitive domain through the end of 1st grade. For example, in general, children from high risk households demonstrated favorable impacts on the Woodcock-Johnson III Spelling test in every year compared with their counterparts from other levels of household risk, except the age 4 year. By 1st grade, children from high risk households experienced benefits in five direct assessments of academic skills. In contrast, favorable cognitive impacts for children from low/no risk households faded by the end of the age 4 year, whereas impacts for the moderate household risk subgroup moved from neutral in the first two years to negative in kindergarten and 1st grade.

The social-emotional impacts, in contrast, were the reverse. Children from moderate risk households benefited more from Head Start than children in high risk households in the reduction of total problem behavior at the end of the Head Start year and at the end of 1st grade (compared to children from low/no risk households). Additional significant impacts included aggressive behavior (at the end of the age 4 year and 1st grade), hyperactive behavior and problems with peer interactions (at the end of 1st grade). There were no social-emotional impacts at the end of kindergarten for this subgroup.

In the health domain, all subgroups showed impacts at the end of the Head Start year. However, all of the health impacts faded by the end of Head Start.

The effect sizes for the significant within subgroup differences for the cognitive domain range from 0.31 to 0.51 for children from high risk households. The effect sizes for the significant within subgroup differences for the social-emotional domain range from 0.23 to 0.35 for children from moderate risk households.

Urbanicity. Exhibit 8.17 shows favorable impacts for children in non-urban settings in all domains, particularly at the end of the Head Start year. Furthermore, in the cognitive domain, the children in non-urban settings experienced favorable impacts through the end of 1st grade while their counterparts in the urban settings experienced no significant impacts in the cognitive domain after the Head Start year. Limited evidence for additional favorable impacts was found at the end of the Head Start years in the social-emotional and health domains. Children in non-urban settings demonstrated more favorable social-emotional impacts than their urban counterparts on aggressive behavior in both Head Start years, and social competencies and total

problem behavior at the end of the Head Start year. In the health domain, there were two impacts for children in non-urban settings: an increase in children's health status at the end of the age 4 year and an increase in care for an injury at the end of the Head Start year (the interpretation of the latter impact is unclear). Children in non-urban settings also had less need for ongoing care at the end of kindergarten.

The effect sizes for the significant within subgroup differences range from 0.22 to 0.49 for children from non-urban areas Most of these effect sizes range from 0.22 to 0.36 with the exception of larger impacts (0.42 to 0.49) on three cognitive impacts in the Head Start year.

Summary

This chapter reviewed the evidence of differential impacts for key subgroups of children. We looked for evidence of patterns of impacts for each subgroup to get a better sense of who benefits most from Head Start.

Many subgroups experienced sustained benefits of Head Start in one or all domains and multiple outcome measures. These subgroups were found in both age cohorts, although they were more common in the 3-year-old cohort. Among the 4-year-olds, the subgroups that stand out are children of parents with reported mild depressive symptoms, Black children, Dual Language Learners, and children in the lowest academic quartile at baseline. Among the 3-year-old cohort, the most notable subgroups with positive effects are children with special needs, children of parents with no reported depressive symptoms, children from high risk households, and children in non-urban settings. In the 3-year-old cohort, children in the lowest academic quartile at baseline and Dual Language Learners also experienced important favorable impacts though they were not sustained through 1st grade. We also identified several subgroups that experienced a mixture of favorable and unfavorable impacts, including 4-year-old children in the lowest academic quartile at baseline, and 3-year-old children from moderate risk households.

Finally, we identified several subgroups that experienced solely – or primarily – negative impacts of Head Start. These include White children in the 4-year-old cohort and 3-year-old children of parents with mild or moderate depressive symptoms. In fact, the 3-year-old children of parents with reported moderate depressive symptoms experienced consistent, sustained

negative impacts throughout the life of the study on those variables for which there were differential impacts across subgroups.

Exhibit 8.4: Significant Subgroup Impacts and Differences in Impact for the 4-Year-Old Cohort: Child's Pre-Academic Skills

Year (Number of Outcomes Examined)	Outcome Measure	Impact in Subgroup A (Head Start – Control) Children not in Lowest Quartile	Impact in Subgroup B (Head Start –Control) Children in Lowest Quartile	Statistically Significant Differences in Impacts Between Subgroups
COGNITIVE				
Direct Child Assessment				
Head Start Year (11)	--	--	--	--
Kindergarten (12)	WJ III Oral Comprehension	0.19	-4.24**	A – B**
	WJ III Math Reasoning	0.96	-3.28*	A – B*
1st grade (15)	WJ III Applied Problems	2.36*	-3.85	A – B*
School Performance				
Head Start Year (0)				
Kindergarten (5)	--	--	--	--
1st grade (5)	Math Ability	-0.01	-0.15**	A – B*
SOCIAL-EMOTIONAL				
Parent Report				
Head Start Year (9)	Closeness	-0.04	1.08**	A – B**
	Positive Relationships	-0.16	2.60*	A – B*
Kindergarten (9)	--	--	--	--
1st grade (9)	--	--	--	--
Teacher Report				
Head Start Year (0)				
Kindergarten (11)	ASPI-Oppositional	0.74	-1.65*	A – B*
	ASPI-Problems with Peer Interaction	0.22	-4.17***	A – B**
	Conflict	0.38	-2.52*	A – B**
	Positive Relationships	-0.29	3.38**	A – B**
1st grade (11)	Closeness	-0.17	1.36*	A – B*
	Positive Relationships	-0.57	2.44*	A – B*
HEALTH				
Head Start Year (5)	--	--	--	--
Kindergarten (5)	--	--	--	--
1st grade (5)	--	--	--	--

Exhibit 8.4: Significant Subgroup Impacts and Differences in Impact for the 4-Year-Old Cohort: Child's Pre-Academic Skills (continued)

Year (Number of Outcomes Examined)	Outcome Measure	Impact in Subgroup A (Head Start – Control) Children not in Lowest Quartile	Impact in Subgroup B (Head Start –Control) Children in Lowest Quartile	Statistically Significant Differences in Impacts Between Subgroups
PARENTING				
Parent Report				
Head Start Year (5)	--	--	--	--
Kindergarten (9)	Parent Spanked Child in Last Week	0.04	**-0.13*** **	A – B**
	Parent Used Time Out in Last Week	-0.02	0.10*	A – B*
1st grade (8)	Parent Read to Child in Last Week	0.03	-0.12*	A – B*
Teacher Report				
Head Start Year (0)				
Kindergarten (2)	--	--	--	--
1st grade (2)	--	--	--	--

Key:
*** $p \leq 0.01$
** $p \leq 0.05$
* $p \leq 0.10$

Bold exhibit entry indicates that the impact passes the Benjamini-Hochberg test for multiple comparisons with a 10 percent false discovery rate.

Gray cell indicates no data available.

Exhibit 8.5: Significant Subgroup Impacts and Differences in Impact for the 4-Year-Old Cohort: Child's Home Language

Year (Number of Outcomes Examined)	Outcome Measure	Impact in Subgroup A (Head Start – Control) Dual Language Learners	Impact in Subgroup B (Head Start –Control) Children English Speaking	Statistically Significant Differences in Impacts Between Subgroups
COGNITIVE				
Direct Child Assessment				
Head Start Year (11)	PPVT (Adapted)	**8.97*** **	1.23	**A – B** **
Kindergarten (12)	--	--		--
1st grade (15)	--	--	--	--
School Performance				
Head Start Year (0)				
Kindergarten (5)	Math Ability	0.15**	0.00	A – B*
1st grade (5)	--	--	--	--
SOCIAL-EMOTIONAL				
Parent Report				
Head Start Year (9)	--	--	--	--
Kindergarten (9)	--	--	--	--
1st grade (9)	--	--	--	--
Teacher Report				
Head Start Year (0)				
Kindergarten (11)	--	--	--	--
1st grade (11)	--	--	--	--
HEALTH				
Head Start Year (5)	Child's Overall Health Status is Excellent/Good Child Has Health Insurance Coverage	**-0.08*** ** 0.07*	-0.01 -0.02	A – B* A – B*
Kindergarten (5)	Child Has Health Insurance Coverage	0.11**	0.01	A – B*
1st grade (5)	Child Received Dental Care	0.08*	-0.01	A – B*

Exhibit 8.5: Significant Subgroup Impacts and Differences in Impact for the 4-Year-Old Cohort: Child's Home Language (continued)

Year (Number of Outcomes Examined)	Outcome Measure	Impact in Subgroup A (Head Start – Control) Dual Language Learners	Impact in Subgroup B (Head Start –Control) Children English Speaking	Statistically Significant Differences in Impacts Between Subgroups
PARENTING				
Parent Report				
Head Start Year (5)	Parental Safety Practices Scale	0.09*	-0.00	A – B*
Kindergarten (9)	--	--	--	--
1st grade (8)	Parenting Style: Neglectful	-0.06**	0.00	A – B*
Teacher Report				
Head Start Year (0)				
Kindergarten (2)	--	--	--	--
1st grade (2)	--	--	--	--

Key:
******* $p \leq 0.01$
** $p \leq 0.05$
* $p \leq 0.10$

Bold exhibit entry indicates that the impact passes the Benjamini-Hochberg test for multiple comparisons with a 10 percent false discovery rate.

Gray cell indicates no data available.

Exhibit 8.6: Significant Subgroup Impacts and Differences in Impact for the 4-Year-Old Cohort: Special Needs

Year (Number of Outcomes Examined)	Outcome Measure	Impact in Subgroup A (Head Start – Control) Children with No Special Needs	Impact in Subgroup B (Head Start –Control) Children with Special Needs	Statistically Significant Differences in Impacts Between Subgroups
COGNITIVE				
Direct Child Assessment				
Head Start Year (11)	--	--	--	--
Kindergarten (12)	--	--	--	--
1st grade (15)	--	--	--	--
School Performance				
Head Start Year (0)				
Kindergarten (5)	--	--	--	--
1st grade (5)	--	--	--	--
SOCIAL-EMOTIONAL				
Parent Report				
Head Start Year (9)	--	--	--	--
Kindergarten (9)	Social Competencies	0.04	-0.46**	A – B***
1st grade (9)	--	--	--	--
Teacher Report				
Head Start Year (0)				
Kindergarten (11)	ASPI – Problems with Structured Learning	-1.45*	4.60*	A – B**
1st grade (11)	--	--	--	--
HEALTH				
Head Start Year (5)	--	--	--	--
Kindergarten (5)	Child's Overall Health Status Is Excellent/Good	0.04	**0.17****	A – B*
	Child Had Care for Injury Last Month	0.00	**0.09****	A – B*
1st grade (5)	Child Needs Ongoing Care	0.00	0.16*	A – B*
	Child Had Care for Injury Last Month	-0.01	**0.22*****	**A – B*****

Exhibit 8.6: Significant Subgroup Impacts and Differences in Impact for the 4-Year-Old Cohort: Special Needs (continued)

Year (Number of Outcomes Examined)	Outcome Measure	Impact in Subgroup A (Head Start – Control) Children with No Special Needs	Impact in Subgroup B (Head Start –Control) Children with Special Needs	Statistically Significant Differences in Impacts Between Subgroups
PARENTING				
Parent Report				
Head Start Year (5)	--	--	--	--
Kindergarten (9)	Parenting Style: Neglectful	-0.01	-0.14*	A – B*
1st grade (8)	--	--	--	--
Teacher Report				
Head Start Year (0)				
Kindergarten (2)	--	--	--	--
1st grade (2)	--	--	--	--

Key:
******* $p \leq 0.01$
** $p \leq 0.05$
* $p \leq 0.10$

Bold exhibit entry indicates that the impact passes the Benjamini-Hochberg test for multiple comparisons with a 10 percent false discovery rate.

Gray cell indicates no data available.

Exhibit 8.7: Significant Subgroup Impacts and Differences in Impact for the 4-Year-Old Cohort: Biological Mother/ Caregiver Race/Ethnicity

Year (Number of Outcomes Examined)	Outcome Measure	Impact in Subgroup A (Head Start – Control) White/Other Children	Impact in Subgroup B (Head Start – Control) Black Children	Impact in Subgroup C (Head Start – Control) Hispanic Children	Statistically Significant Differences in Impacts Between Subgroups
COGNITIVE					
Direct Child Assessment					
Head Start Year (11)	PPVT (Adapted)	0.90	-2.04	**9.35*** **	A – C** B – C**
	Counting Bears	-0.06	--	0.09*	A – C*
	WJ III Spelling	1.85	**8.62*** **	--	A – B*
Kindergarten (12)	WJ III Spelling	-4.90*	6.62**	-1.55	**A – B*** ** B – C*
	WJ III Basic Reading Skills	-2.84	5.95*	-3.44	A – B* B – C**
1st grade (15)					
School Performance					
Head Start Year (0)					
Kindergarten (5)	--	--	--	--	--
1st grade (5)	--	--	--	--	--
SOCIAL-EMOTIONAL					
Parent Report					
Head Start Year (9)	--	--	--	--	--
Kindergarten (9)	--	--	--	--	--
1st grade (9)	Social Skills and Positive Approaches to Learning	0.22	0.33	-0.32***	A – C** B – C***

Exhibit 8.7: Significant Subgroup Impacts and Differences in Impact for the 4-Year-Old Cohort: Biological Mother/ Caregiver Race/Ethnicity (continued)

Year (Number of Outcomes Examined)	Outcome Measure	Impact in Subgroup A (Head Start – Control) White/Other Children	Impact in Subgroup B (Head Start – Control) Black Children	Impact in Subgroup C (Head Start – Control) Hispanic Children	Statistically Significant Differences in Impacts Between Subgroups
Teacher Report					
Head Start Year (0)					
Kindergarten (11)	ASPI-Inattentive/Hyperactive	-0.36	-4.60**	0.86	A – B** **B – C*****
	ASPI-Problems with Structured Learning	1.02	-5.34**	0.07	A – B*** B – C**
	ASPI-Problems with Peer Interaction	1.47	-6.12*	--	A – B**
	Positive Relationships	-0.56	4.07*	0.08	A – B* B – C*
	ASPI-Problems with Teacher Interaction	1.85	-3.37*	0.38	A – B** B – C*
	ASPI-Shy/Socially Reticent	2.25*	--	-0.84	A – C*
1st grade (11)	ASPI-Withdrawn/Low Energy	2.25**	-1.51	--	A – B**
	ASPI-Shy/Socially Reticent	2.73***	--	0.67	A – C*
	ASPI-Problems with Teacher Interaction	3.12**	--	0.08	A – C**
HEALTH					
Head Start Year (5)	Child's Overall Health Status Is Excellent/Good	**-0.09****	0.03	--	A – B**
Kindergarten (5)	--	--	--	--	--
1st grade (5)	Child's Overall Health Status Is Excellent/Good	-0.03	0.10*	-0.05	A – B* B – C*

Exhibit 8.7: Significant Subgroup Impacts and Differences in Impact for the 4-Year-Old Cohort: Biological Mother/ Caregiver Race/Ethnicity (continued)

Year (Number of Outcomes Examined)	Outcome Measure	Impact in Subgroup A (Head Start – Control) White/Other Children	Impact in Subgroup B (Head Start – Control) Black Children	Impact in Subgroup C (Head Start – Control) Hispanic Children	Statistically Significant Differences in Impacts Between Subgroups
PARENTING					
Parent Report					
Head Start Year (5)	--	--	--	--	--
Kindergarten (9)	Parenting Style: Neglectful	-0.09**	--	-0.00	A – C*
	Parent Read to Child in Last Week	-0.10*	0.07	--	A – B*
	Family Cultural Enrichment Scale	--	-0.15	0.34**	B – C*
1st grade (8)	Parenting Style: Neglectful	--	0.05	-0.05**	**B – C****
	Parent Spanked Child in Last Week	0.03	**-0.11*****	0.05	**A – B*****
					B – C***
	Parenting Style: Permissive	0.03	-0.10*	0.04	A – B*
					B – C*
Teacher Report					
Head Start Year (0)					
Kindergarten (2)	--	--	--	--	--
1st grade (2)	--	--	--	--	--

Key:
*** $p \leq 0.01$
** $p \leq 0.05$
* $p \leq 0.10$

Bold exhibit entry indicates that the impact passes the Benjamini-Hochberg test for multiple comparisons with a 10 percent false discovery rate.

Gray cell indicates no data available.

Exhibit 8.8: Significant Subgroup Impacts and Differences in Impact for the 4-Year-Old Cohort: Parent/Caregiver Reported Depressive Symptoms

Year (Number of Outcomes Examined)	Outcome Measure	Impact in Subgroup A (Head Start – Control) No Symptoms	Impact in Subgroup B (Head Start – Control) Mild Symptoms	Impact in Subgroup C (Head Start – Control) Moderate Symptoms	Impact in Subgroup D (Head Start – Control) Severe Symptoms	Statistically Significant Differences in Impacts Between Subgroups
COGNITIVE						
Direct Child Assessment						
Head Start Year (11)	PPVT (Adapted)	5.30**	0.20	--	--	A – B*
	Counting Bears	--	**0.16****	-0.06	--	B – C*
	Letter Naming	0.69	**3.90*****	**4.75*****	**4.64*****	A – B* **A – C*****
						A – D**
	WJ III Pre-Academic Skills	1.06	**5.68****	--	--	A – B*
	WJ III Letter Word Identification	2.90	**7.26****	--	**13.61*****	**A – D*****
						B – D*
	Color Identification	0.02	--	--	0.20**	A – D*
Kindergarten (12)	--	--	--	--	--	--
1st grade (15)	WJ III Academic Application	-0.98	3.88*	--	-3.23	A – B*
						B – D**
	WJ III Quantitative Concepts	1.13	0.56	--	-4.81*	A – D*
						B – D*
	WJ III Passage Comprehension	-2.60	6.37*	-5.82	-3.44	A – B**
						B – C*
						B – D**
	WJ III Applied Problems	--	3.87**	--	-3.96	B – D**
	WJ III Word Attack	--	7.56*	--	-5.55	B – D*
School Performance						
Head Start Year (0)						
Kindergarten (5)	--	--	--	--	--	--
1st grade (5)	--	--	--	--	--	--

Exhibit 8.8: Significant Subgroup Impacts and Differences in Impact for the 4-Year-Old Cohort: Parent/Caregiver Reported Depressive Symptoms (continued)

Year (Number of Outcomes Examined)	Outcome Measure	Impact in Subgroup A (Head Start – Control) No Symptoms	Impact in Subgroup B (Head Start – Control) Mild Symptoms	Impact in Subgroup C (Head Start – Control) Moderate Symptoms	Impact in Subgroup D (Head Start – Control) Severe Symptoms	Statistically Significant Differences in Impacts Between Subgroups
SOCIAL-EMOTIONAL						
Parent Report						
Head Start Year (9)	Hyperactive Behavior	0.02	**-0.46****	--	--	A – B**
	Total Problem Behavior	-0.11	**-1.05*****	0.70	--	A – B* B – C**
Kindergarten (9)	Withdrawn Behavior	-0.18**	0.19*	0.36	-0.38	**A – B***** A – C** B – D**
	Social Competencies	0.11	--	--	-0.48**	A – D**
1st grade (9)	Social Competencies	0.03	-0.32***	--	--	A – B**
	Aggressive Behavior	--	0.02	0.20	-0.75*	B – D* C – D**
Teacher Report						
Head Start Year (0)						
Kindergarten (11)	ASPI-Inattentive/ Hyperactive	0.74	-2.63*	--	--	A – B*
1st grade (11)	ASPI–Withdrawn/Low Energy	-0.18	0.21	4.23**	--	A – C** B – C*
	ASPI–Aggressive	0.05	-0.16	--	-3.98**	A – D* B – D*

Exhibit 8.8: Significant Subgroup Impacts and Differences in Impact for the 4-Year-Old Cohort: Parent/Caregiver Reported Depressive Symptoms (continued)

Year (Number of Outcomes Examined)	Outcome Measure	Impact in Subgroup A (Head Start – Control) No Symptoms	Impact in Subgroup B (Head Start – Control) Mild Symptoms	Impact in Subgroup C (Head Start – Control) Moderate Symptoms	Impact in Subgroup D (Head Start – Control) Severe Symptoms	Statistically Significant Differences in Impacts Between Subgroups
HEALTH						
Head Start Year (5)	Child Needs Ongoing Care	0.02	-0.00	--	0.10**	A – D* B – D**
Kindergarten (5)	Child Had Care for Injury Last Month Child Received Dental Care	-- 0.02	-- -0.06	0.14** --	-0.07 0.21**	C – D* A – D* B – D*
1st grade (5)	Child Needs Ongoing Care Child Had Care for Injury Last Month	0.01 --	-- -0.05	-- -0.00	**0.13**** **0.16****	A – D* B – D** C – D*
PARENTING						
Parent Report						
Head Start Year (5)	--	--	--	--	--	--
Kindergarten (9)	Parenting Style: Neglectful	-0.04*	--	0.08	-0.17*	A – C** C – D**
1st grade (8)	Parenting Style: Authoritarian	--	-0.01	0.02	-0.15*	B – D* C – D**

Exhibit 8.8: Significant Subgroup Impacts and Differences in Impact for the 4-Year-Old Cohort: Parent/Caregiver Reported Depressive Symptoms (continued)

Year (Number of Outcomes Examined)	Outcome Measure	Impact in Subgroup A (Head Start – Control) No Symptoms	Impact in Subgroup B (Head Start – Control) Mild Symptoms	Impact in Subgroup C (Head Start – Control) Moderate Symptoms	Impact in Subgroup D (Head Start – Control) Severe Symptoms	Statistically Significant Differences in Impacts Between Subgroups
Teacher Report						
Head Start Year (0)						
Kindergarten (2)	Parent Participation	0.04	0.05	-0.12**	--	A – C** B – C**
1st grade (2)	Parent Participation	--	0.01	-0.14*	0.15	B – C* C – D**

Key:
*** $p \leq 0.01$
** $p \leq 0.05$
* $p \leq 0.10$

Bold exhibit entry indicates that the impact passes the Benjamini-Hochberg test for multiple comparisons with a 10 percent false discovery rate.

Gray cell indicates no data available.

Exhibit 8.9: Significant Subgroup Impacts and Differences in Impact for the 4-Year-Old Cohort: Household Risk Index

Year (Number of Outcomes Examined)	Outcome Measure	Impact in Subgroup A (Head Start – Control) Low/No Household Risk	Impact in Subgroup B (Head Start – Control) Moderate Household Risk	Impact in Subgroup C (Head Start – Control) High Household Risk	Statistically Significant Differences in Impacts Between Subgroups
COGNITIVE					
Direct Child Assessment					
Head Start Year (11)	Letter Naming	**2.10****	**4.49*****	0.24	A – B* B – C*
	Color Identification	**0.10*****	--	-0.08	A – C*
	WJ III Applied Problems	--	5.28**	-2.76	B – C*
Kindergarten (12)	--	--	--	--	--
1st grade (15)	PPVT (Adapted)	1.04	11.73**	--	A – B**
School Performance					
Head Start Year (0)					
Kindergarten (5)	--	--	--	--	--
1st grade (5)	--	--	--	--	--
SOCIAL-EMOTIONAL					
Parent Report					
Head Start Year (9)	Closeness	0.49**	-0.37	-0.55	A – B* A – C**
	Social Skills and Positive Approaches to Learning	-0.08	0.26	-0.66**	A – C* B – C**
	Social Competencies	0.00	--	-0.40*	A – C*
Kindergarten (9)	Positive Relationships	0.46	--	-2.66*	A – C*
1st grade (9)	--	--	--	--	--

Exhibit 8.9: Significant Subgroup Impacts and Differences in Impact for the 4-Year-Old Cohort: Household Risk Index (continued)

Year (Number of Outcomes Examined)	Outcome Measure	Impact in Subgroup A (Head Start – Control) Low/No Household Risk	Impact in Subgroup B (Head Start – Control) Moderate Household Risk	Impact in Subgroup C (Head Start – Control) High Household Risk	Statistically Significant Differences in Impacts Between Subgroups
Teacher Report					
Head Start Year (0)					
Kindergarten (11)	ASPI-Inattentive/Hyperactive	0.14	-3.93*	--	A – B*
	ASPI-Withdrawn/Low Energy	-0.41	2.45*	--	A – B**
	Closeness	0.07	--	3.41**	A – C**
	ASPI-Aggressive	0.63	--	-5.69**	A – C**
1st grade (11)	ASPI-Inattentive/Hyperactive	0.10	--	-3.50*	A – C*
	ASPI-Withdrawn/Low Energy	1.35**	-1.46	--	A – B*
HEALTH					
Head Start Year (5)	Child Received Dental Care	**0.16*** **	0.01	**0.36*** **	A – B**
					A – C*
					B – C* **
	Child's Overall Health Status Is Excellent/Good	**-0.06****	--	0.18	A – C*
Kindergarten (5)	--	--	--	--	--
1st grade (5)	Child Received Dental Care	0.04	-0.13*	0.14	A – B**
					B – C*
	Child Had Care for Injury Last Month	0.01	0.01	**0.20*** **	**A – C****
					B – C*
PARENTING					
Parent Report					
Head Start Year (5)	--	--	--	--	--
Kindergarten (9)	Parental Cultural Enrichment Scale	0.24**	--	-0.31	A – C*
	Parenting Style: Authoritarian	-0.01	--	0.10*	A – C*
1st grade (8)	Parenting Style: Authoritative	0.08**	-0.05	--	A – B*
	Parenting Style: Permissive	-0.03	0.13**	--	A – B**
	Parenting Style: Neglectful	--	-0.09*	0.09	B – C**

Exhibit 8.9: Significant Subgroup Impacts and Differences in Impact for the 4-Year-Old Cohort: Household Risk Index (continued)

Year (Number of Outcomes Examined)	Outcome Measure	Impact in Subgroup A (Head Start – Control) Low/No Household Risk	Impact in Subgroup B (Head Start – Control) Moderate Household Risk	Impact in Subgroup C (Head Start – Control) High Household Risk	Statistically Significant Differences in Impacts Between Subgroups
Teacher Report					
Head Start Year (0)					
Kindergarten (2)	--	--	--	--	--
1st grade (2)	--	--	--	--	--

Key:
*** $p \leq 0.01$
** $p \leq 0.05$
* $p \leq 0.10$

Bold exhibit entry indicates that the impact passes the Benjamini-Hochberg test for multiple comparisons with a 10 percent false discovery rate.

Gray cell indicates no data available.

Exhibit 8.10: Significant Subgroup Impacts and Differences in Impact for the 4-Year-Old Cohort: Urbanicity

Year (Number of Outcomes Examined)	Outcome Measure	Impact in Subgroup A (Head Start – Control) Not Urban	Impact in Subgroup B (Head Start –Control) Urban	Statistically Significant Differences in Impacts Between Subgroups
COGNITIVE				
Direct Child Assessment				
Head Start Year (11)	Letter Naming	**5.20*** **	**1.82** **	A – B**
	Counting Bears	-0.12	0.07**	A – B*
Kindergarten (12)	WJ III Spelling	-7.86**	1.07	A – B**
1st grade (15)	--	--	--	--
School Performance				
Head Start Year (0)				
Kindergarten (5)	--	--	--	--
1st grade (5)	Math Ability	-0.14**	-0.03	A – B*
SOCIAL-EMOTIONAL				
Parent Report				
Head Start Year (9)	--	--	--	--
Kindergarten (9)	Total Problem Behavior	0.90**	-0.09	A – B**
1st grade (9)	--	--	--	--
Teacher Report				
Head Start Year (0)				
Kindergarten (11)	--	--	--	--
1st grade (11)	--	--	--	--
HEALTH				
Head Start Year (5)	Child Received Dental Care	**0.32*** **	**0.12*** **	A – B*
Kindergarten (5)	--	--	--	--
1st grade (5)	--	--	--	--
PARENTING				
Parent Report				
Head Start Year (5)	Parent Spanked Child in Last Week	-0.12*	0.01	A – B**
Kindergarten (9)	--	--	--	--
1st grade (8)	--	--	--	--

Exhibit 8.10: Significant Subgroup Impacts and Differences in Impact for the 4-Year-Old Cohort: Urbanicity (continued)

Year (Number of Outcomes Examined)	Outcome Measure	Impact in Subgroup A (Head Start – Control) Not Urban	Impact in Subgroup B (Head Start –Control) Urban	Statistically Significant Differences in Impacts Between Subgroups
Teacher Report				
Head Start Year (0)				
Kindergarten (2)	Parent Participation	-0.11**	0.01	**A – B*** **
1st grade (2)	School Contact and Communication	-0.15*	0.00	A – B*

Key:
*** $p \leq 0.01$
** $p \leq 0.05$
* $p \leq 0.10$

Bold exhibit entry indicates that the impact passes the Benjamini-Hochberg test for multiple comparisons with a 10 percent false discovery rate.

Gray cell indicates no data available.

Exhibit 8.11: Significant Subgroup Impacts and Differences in Impact for the 3-Year-Old Cohort: Child's Pre-Academic Skills

Year (Number of Outcomes Examined)	Outcome Measure	Impact in Subgroup A (Head Start – Control) Children not in Lowest Quartile	Impact in Subgroup B (Head Start –Control) Children in Lowest Quartile	Statistically Significant Differences in Impacts Between Subgroups
COGNITIVE				
Direct Child Assessment				
Head Start Year (11)	--	--	--	--
Age 4 Year (11)	PPVT (Adapted)	-0.39	**8.87*** **	A – B**
	WJ III Applied Problems	-0.73	**5.03** **	A – B*
	WJ III Oral Comprehension	-0.50	**2.31** **	A – B*
	WJ III Pre-Academic Skills	0.02	**4.09** **	A – B*
	WJ III Letter Word Identification	1.07	**6.61** **	A – B*
Kindergarten (12)	--	--	--	--
1st grade (15)	--	--	--	--
School Performance				
Head Start Year (0)				
Age 4 Year (0)				
Kindergarten (5)	--	--	--	--
1st grade (5)	--	--	--	--
SOCIAL-EMOTIONAL				
Parent Report				
Head Start Year (9)	Withdrawn Behavior	0.02	-0.20*	A – B*
Age 4 Year (9)	Social Skills and Positive Approaches to Learning	0.06	**0.59*** **	A – B**
Kindergarten (9)	--	--	--	--
1st grade (9)	--	--	--	--
Teacher Report				
Head Start Year (0)				
Age 4 Year (0)				
Kindergarten (11)	ASPI-Aggressive	0.99*	-1.49	A – B*
	ASPI-Problems with Peer Interaction	1.56**	-2.36	A – B**
1st grade (11)	--	--	--	--

Exhibit 8.11: Significant Subgroup Impacts and Differences in Impact for the 3-Year-Old Cohort: Child's Pre-Academic Skills (continued)

Year (Number of Outcomes Examined)	Outcome Measure	Impact in Subgroup A (Head Start – Control) Children not in Lowest Quartile	Impact in Subgroup B (Head Start –Control) Children in Lowest Quartile	Statistically Significant Differences in Impacts Between Subgroups
HEALTH				
Head Start Year (5)	--	--	--	--
Age 4 Year (5)	--	--	--	--
Kindergarten (5)	--	--	--	--
1st grade (5)	--	--	--	--
PARENTING				
Parent Report				
Head Start Year (5)	Parent Used Time Out in Last Week	-0.07**	0.04	A – B*
Age 4 Year (9)	Parent Read to Child in Last Week	-0.04	0.10*	A – B**
	Parenting Style: Authoritarian	-0.02	**-0.09*** **	A – B**
Kindergarten (9)	--	--	--	--
1st grade (8)	--	--	--	--
Teacher Report				
Head Start Year (0)				
Age 4 Year (0)				
Kindergarten (2)	--	--	--	--
1st grade (2)	--	--	--	--

Key:
*** $p \leq 0.01$
** $p \leq 0.05$
* $p \leq 0.10$

Bold exhibit entry indicates that the impact passes the Benjamini-Hochberg test for multiple comparisons with a 10 percent false discovery rate.

Gray cell indicates no data available.

Exhibit 8.12: Significant Subgroup Impacts and Differences in Impact for the 3-Year-Old Cohort: Child's Home Language

Year (Number of Outcomes Examined)	Outcome Measure	Impact in Subgroup A (Head Start – Control) Dual Language Learners	Impact in Subgroup B (Head Start –Control) Children English Speaking	Statistically Significant Differences in Impacts Between Subgroups
COGNITIVE				
Direct Child Assessment				
Head Start Year (11)	PPVT (Adapted)	**11.87*** **	**4.97** **	A – B*
	WJ III Spelling	7.66**	0.62	A – B*
Age 4 Year (11)	CTOPPP Elision	**23.20*** **	4.00	**A – B** **
	Letter Naming	**3.20*** **	0.17	**A – B** **
	WJ III Pre-Academic Skills	**5.72** **	-0.07	**A – B** **
	WJ III Letter Word Identification	**8.59*** **	0.79	**A – B** **
	Counting Bears	**0.16*** **	-0.04	**A – B*** **
	WJ III Spelling	**4.75***	-1.03	**A – B***
Kindergarten (12)	WJ III Applied Problems	4.20	-2.56*	A – B**
	WJ III Pre-Academic Skills	5.20*	-1.68	A – B**
	WJ III Spelling	5.20*	-1.07	A – B*
	WJ III Math Reasoning	2.78	-2.07*	A – B**
1st grade (15)	--	--	--	--
School Performance				
Head Start Year (0)				
Age 4 Year (0)				
Kindergarten (5)	--	--	--	--
1st grade (5)	--	--	--	--
SOCIAL-EMOTIONAL				
Parent Report				
Head Start Year (9)	Withdrawn Behavior	0.19	-0.10*	A – B*
Age 4 Year (9)	Withdrawn Behavior	-0.24*	-0.03	A – B*
	Closeness	-0.77	0.54**	A – B**
Kindergarten (9)	--	--	--	--
1st grade (9)	--	--	--	--

Exhibit 8.12: Significant Subgroup Impacts and Differences in Impact for the 3-Year-Old Cohort: Child's Home Language (continued)

Year (Number of Outcomes Examined)	Outcome Measure	Impact in Subgroup A (Head Start – Control) Dual Language Learners	Impact in Subgroup B (Head Start –Control) Children English Speaking	Statistically Significant Differences in Impacts Between Subgroups
Teacher Report				
Head Start Year (0)				
Age 4 Year (0)				
Kindergarten (11)	--	--	--	--
1st grade (11)	--	--	--	--
HEALTH				
Head Start Year (5)	Child Received Dental Care	**0.24*** **	**0.14*** **	A – B*
	Child's Overall Health Status Is Excellent/Good	**0.13*** **	0.02	A – B*
Age 4 Year (5)	--	--	--	--
Kindergarten (5)	Child Had Care for Injury Last Month	0.05*	-0.02	A – B*
1st grade (5)	--	--	--	--
PARENTING				
Parent Report				
Head Start Year (5)	Parental Safety Practices Scale	-0.02	**0.05***	A – B*
Age 4 Year (9)	--	--	--	--
Kindergarten (9)	--	--	--	--
1st grade (8)	Parenting Style: Permissive	-0.07	0.05**	A – B*
Teacher Report				
Head Start Year (0)				
Age 4 Year (0)				
Kindergarten (2)	--	--	--	--
1st grade (2)	School Contact and Communication	**0.15*** **	-0.02	**A – B*** **

Key:
******* $p \leq 0.01$
** $p \leq 0.05$
* $p \leq 0.10$

Bold exhibit entry indicates that the impact passes the Benjamini-Hochberg test for multiple comparisons with a 10 percent false discovery rate.
Gray cell indicates no data available.

Exhibit 8.13: Significant Subgroup Impacts and Differences in Impact for the 3-Year-Old Cohort: Special Needs

Year (Number of Outcomes Examined)	Outcome Measure	Impact in Subgroup A (Head Start – Control) Children With No Special Needs	Impact in Subgroup B (Head Start –Control) Children With Special Needs	Statistically Significant Differences in Impacts Between Subgroups
COGNITIVE				
Direct Child Assessment				
Head Start Year (11)	--	--	--	--
Age 4 Year (11)	--	--	--	--
Kindergarten (12)	--	--	--	--
1st grade (15)	WJ III Applied Problems	0.66	9.11*	A – B*
	WJ III Calculation	-0.70	5.86*	A – B*
	WJ III Quantitative Concept	0.08	6.64*	A – B*
	WJ III Math Reasoning	0.39	7.88**	A – B*
School Performance				
Head Start Year (0)				
Age 4 Year (0)				
Kindergarten (5)	--	--	--	--
1st grade (5)	--	--	--	--
Social-Emotional				
Parent Report				
Head Start Year (9)	--	--	--	--
Age 4 Year (9)	--	--	--	--
Kindergarten (9)	--	--	--	--
1st grade (9)	--	--	--	--
Teacher Report				
Head Start Year (0)				
Age 4 Year (0)				
Kindergarten (11)	--	--	--	--
1st grade (11)	ASPI-Inattentive/Hyperactive	-0.01	-3.84*	A – B*
	Conflict	0.23	-3.10*	A – B**
	Positive Relationships	0.05	3.56*	A – B*
	ASPI- Problems with Structured Learning	0.54	-5.01**	A – B**

Exhibit 8.13: Significant Subgroup Impacts and Differences in Impact for the 3-Year-Old Cohort: Special Needs (continued)

Year (Number of Outcomes Examined)	Outcome Measure	Impact in Subgroup A (Head Start – Control) Children With No Special Needs	Impact in Subgroup B (Head Start –Control) Children With Special Needs	Statistically Significant Differences in Impacts Between Subgroups
HEALTH				
Head Start Year (5)	--	--	--	--
Age 4 Year (5)	Child Needs Ongoing Care	-0.01	0.15*	A – B**
Kindergarten (5)	Child Had Care for Injury Last Month	-0.01	**0.10*** **	**A – B*** **
1st grade (5)	--	--	--	--
PARENTING				
Parent Report				
Head Start Year (5)	Parent Used Time Out in Last Week	-0.03	-0.15**	A – B*
Age 4 Year (9)	Parent Read to Child in Last Week	0.02	-0.17**	A – B**
Kindergarten (9)	Parent Spanked Child in Last Week	-0.03	-0.17**	A – B*
	Parenting Style: Authoritarian	-0.01	0.09**	A – B**
	Parenting Style: Authoritative	0.01	-0.12*	A – B*
1st grade (8)	Parent Used Time Out in Last Week	-0.02	**-0.30*** **	**A – B*** **
Teacher Report				
Head Start Year (0)				
Age 4 Year (0)				
Kindergarten (2)	--	--	--	--
1st grade (2)	--	--	--	--

Key:
*** $p \leq 0.01$
** $p \leq 0.05$
* $p \leq 0.10$

Bold exhibit entry indicates that the impact passes the Benjamini-Hochberg test for multiple comparisons with a 10 percent false discovery rate.

Gray cell indicates no data available.

Exhibit 8.14: Significant Subgroup Impacts and Differences in Impact for the 3-Year-Old Cohort: Biological Mother/ Caregiver Race/Ethnicity

Year (Number of Outcomes Examined)	Outcome Measure	Impact in Subgroup A (Head Start – Control) White/Other Children	Impact in Subgroup B (Head Start – Control) Black Children	Impact in Subgroup C (Head Start – Control) Hispanic Children	Statistically Significant Differences in Impacts Between Subgroups
COGNITIVE					
Direct Child Assessment					
Head Start Year (11)	WJ III Oral Comprehension	2.41**	-2.07	--	A – B**
	WJ III Spelling	0.26	0.19	**6.70****	A – C *
					B – C**
	CTOPPP Elision	--	7.20*	-2.40	B – C*
Age 4 Year (11)	PPVT (Adapted)	--	-1.94	5.26*	B – C*
	WJ III Spelling	-2.40	--	3.64*	A – C*
	McCarthy Draw-a-Design	-0.38*	--	0.32	A – C**
	Counting Bears	--	-0.07	0.09**	B – C**
Kindergarten (12)	Letter Naming	0.38	--	-1.15**	A – C**
	WJ III Oral Comprehension	3.19**	-1.31	-0.34	A – B**
					A – C*
	WJ III Word Attack	3.39	-6.76*	--	A – B**
	WJ III Applied Problems	--	-3.86*	2.94	B – C**
1st grade (15)	PPVT (Adapted)	6.37**	-0.47	0.97	A – B*
					A – C*
	WJ III Quantitative Concept	--	-0.84	3.18*	B – C*
School Performance					
Head Start Year (0)					
Age 4 Year (0)					
Kindergarten (5)	--	--	--	--	--
1st grade (5)	--	--	--	--	--

Exhibit 8.14: Significant Subgroup Impacts and Differences in Impact for the 3-Year-Old Cohort: Biological Mother/ Caregiver Race/Ethnicity (continued)

Year (Number of Outcomes Examined)	Outcome Measure	Impact in Subgroup A (Head Start – Control) White/Other Children	Impact in Subgroup B (Head Start – Control) Black Children	Impact in Subgroup C (Head Start – Control) Hispanic Children	Statistically Significant Differences in Impacts Between Subgroups
SOCIAL-EMOTIONAL					
Parent Report					
Head Start Year (9)	Social Competencies	0.20*	-0.24*	--	A – B**
	Social Skills and Positive Approaches to Learning	0.30**	--	-0.19	A – C**
	Closeness	0.48*	--	-0.28	A – C*
Age 4 Year (9)	Closeness	0.68**	--	-0.35	A – C*
Kindergarten (9)	Closeness	0.60**	--	-0.12	A – C*
1st grade (9)	Conflict	0.47	-1.64*	--	A – B*
	Aggressive Behavior	0.40**	-0.33*	-0.22	**A – B*** **
					A – C**
	Total Problem Behavior	0.77*	-0.76**	-0.46	A – B**
					A – C**
	Positive Relationships	-0.50	2.18**	--	A – B**
	Hyperactive Behavior	0.18	--	-0.31**	A – C**
	Social Skills and Positive Approaches to Learning	0.34**	--	-0.23**	**A – C*** **
Teacher Report					
Head Start Year (0)					
Age 4 Year (0)					
Kindergarten (11)	ASPI-Inattentive/Hyperactive	1.07	0.61	-1.99**	A – C**
					B – C*
1st grade (11)	ASPI-Withdrawn/Low Energy	-0.62	--	1.56**	A – C**

Exhibit 8.14: Significant Subgroup Impacts and Differences in Impact for the 3-Year-Old Cohort: Biological Mother/ Caregiver Race/Ethnicity (continued)

Year (Number of Outcomes Examined)	Outcome Measure	Impact in Subgroup A (Head Start – Control) White/Other Children	Impact in Subgroup B (Head Start – Control) Black Children	Impact in Subgroup C (Head Start – Control) Hispanic Children	Statistically Significant Differences in Impacts Between Subgroups
HEALTH					
Head Start Year (5)	Child Had Care for Injury Last Month	**0.07*** **	-0.03	**-0.07***	**A – B**** **A – C*****
	Child Received Dental Care	**0.22*** **	0.09*	--	A – B*
	Child's Overall Health Status Is Excellent/Good	0.02	--	**0.12*** **	A – C*
Age 4 Year (5)	--	--	--	--	--
Kindergarten (5)	Child Has Health Insurance Coverage	-0.01	**0.08*** **	--	**A – B****
	Child Needs Ongoing Care	**-0.10*** **	--	0.05	**A – C****
1st grade (5)	Child Needs Ongoing Care	**-0.12*** **	0.05	0.04	**A – B***** **A – C*****
PARENTING					
Parent Report					
Head Start Year (9)	--	--	--	--	--
Age 4 Year (5)	--	--	--	--	--
Kindergarten (9)	Parenting Style: Authoritative	0.09*	-0.06	-0.03	A – B** A – C*
	Parental Cultural Enrichment Scale	-0.04	-0.18	0.22*	A – C* B – C**
	Parent Spanked Child in Last Week	-0.00	--	**-0.08****	A – C*
	Parenting Style: Permissive	-0.03	0.07*	--	A – B*
	Parent Used Time Out in Last Week	-0.00	**-0.12*** **	--	A – B**
	Parental Safety Practices Scale	0.01	**-0.07*** **	**0.09****	A – B* **B – C*****

Exhibit 8.14: Significant Subgroup Impacts and Differences in Impact for the 3-Year-Old Cohort: Biological Mother/ Caregiver Race/Ethnicity (continued)

Year (Number of Outcomes Examined)	Outcome Measure	Impact in Subgroup A (Head Start – Control) White/Other Children	Impact in Subgroup B (Head Start – Control) Black Children	Impact in Subgroup C (Head Start – Control) Hispanic Children	Statistically Significant Differences in Impacts Between Subgroups
1st grade (8)	Parent Spanked Child in Last Week	0.05	-0.07**	--	A – B**
	Parenting Style: Authoritative	--	-0.08*	0.05	B – C**
Teacher Report					
Head Start Year (0)					
Age 4 Year (0)					
Kindergarten (2)	--	--	--	--	--
1st grade (2)	School Contact and Communication	-0.02	-0.02	0.11**	A – C* B – C*

Key:
*** $p \leq 0.01$
** $p \leq 0.05$
* $p \leq 0.10$

Bold exhibit entry indicates that the impact passes the Benjamini-Hochberg test for multiple comparisons with a 10 percent false discovery rate.

Gray cell indicates no data available.

Exhibit 8.15: Significant Subgroup Impacts and Differences in Impact for the 3-Year-Old Cohort: Parent/Caregiver Reported Depressive Symptoms

Year (Number of Outcomes Examined)	Outcome Measure	Impact in Subgroup A (Head Start – Control) No Symptoms	Impact in Subgroup B (Head Start – Control) Mild Symptoms	Impact in Subgroup C (Head Start – Control) Moderate Symptoms	Impact in Subgroup D (Head Start – Control) Severe Symptoms	Statistically Significant Differences in Impacts Between Subgroups
COGNITIVE						
Direct Child Assessment						
Head Start Year (11)	CTOPPP Elision	**9.01****	--	-5.37	--	A – C*
	Letter Naming	1.00	--	--	**2.76*****	A – D*
Age 4 Year (11)	CTOPPP Elision	**9.93*****	11.34*	-2.55	12.54*	**A – C****
						B – C*
						C – D*
	WJ III Applied Problems	3.15*	--	-2.84	--	A – C*
	WJ III Letter Word Identification	5.23**	--	-5.70	--	**A – C*****
	WJ III Pre-Academic Skills	3.52**	1.85	-5.51*	--	**A – C****
						B – C*
	WJ III Spelling	3.12	0.06	**-10.81*****	--	**A – C*****
						B –C**
	Letter Naming	1.14*	--	-1.96	--	**A – C***
	McCarthy Draw-a-Design	0.00	0.24	-0.52**	--	A – C*
						B – C**
	PPVT (Adapted)	**7.38*****	1.16	-9.81**	--	**A – C*****
						B – C*
Kindergarten (12)	WJ III Letter Word Identification	4.73**	-4.28	--	--	**A – B*****

Exhibit 8.15: Significant Subgroup Impacts and Differences in Impact for the 3-Year-Old Cohort: Parent/Caregiver Reported Depressive Symptoms (continued)

Year (Number of Outcomes Examined)	Outcome Measure	Impact in Subgroup A (Head Start – Control) No Symptoms	Impact in Subgroup B (Head Start – Control) Mild Symptoms	Impact in Subgroup C (Head Start – Control) Moderate Symptoms	Impact in Subgroup D (Head Start – Control) Severe Symptoms	Statistically Significant Differences in Impacts Between Subgroups
1st grade (15)	PPVT (Adapted)	**6.47****	-0.46	-6.48	--	A – B* **A – C****
	WJ III Pre-Academic Skills	3.96*	-3.80	-7.53*	5.81	A – B* **A – C***** C – D*
	WJ III Basic Reading Skills	4.98*	-5.93	-10.03*	8.82	A – B* **A – C***** C – D*
	WJ III Applied Problems	**4.16****	-3.37	-2.07	--	A – B ** **A – C***
	WJ III Letter Word Identification	6.30*	-6.09	-13.61**	6.87	A – B* **A – C***** C – D*
	WJ III Oral Comprehension	**3.45****	1.65	-5.15***	1.85	**A – C**** B – C** C – D*
	WJ III Academic Applications	**3.45****	--	-5.06**	3.98	**A – C***** C – D*
	WJ III Math Reasoning	**3.22****	--	-3.15	--	**A – C****
	WJ III Academic Skills	2.27	--	-8.45**	5.46	**A – C***** C – D*
	WJ III Writing Sample	1.53	--	-4.43**	--	**A – C*****
	WJ III Passage Comprehension	**4.89****	--	-9.26**	2.15	**A – C***** C – D*
	WJ III Spelling	1.76	--	-8.16*	--	**A – C*****

Exhibit 8.15: Significant Subgroup Impacts and Differences in Impact for the 3-Year-Old Cohort: Parent/Caregiver Reported Depressive Symptoms (continued)

Year (Number of Outcomes Examined)	Outcome Measure	Impact in Subgroup A (Head Start – Control) No Symptoms	Impact in Subgroup B (Head Start – Control) Mild Symptoms	Impact in Subgroup C (Head Start – Control) Moderate Symptoms	Impact in Subgroup D (Head Start – Control) Severe Symptoms	Statistically Significant Differences in Impacts Between Subgroups
School Performance						
Head Start Year (0)						
Age 4 Year (0)						
Kindergarten (5)	Language and Literacy Ability	--	0.06	**-0.16****	--	B – C*
	Math Ability	--	0.02	**-0.19****	-0.19*	B – C** B – D*
1st grade (5)	Language and Literacy Ability	0.03	-0.11*	--	--	A – B**
	Math Ability	-0.01	**-0.11****	--	--	A – B*
	Promotion	0.02	--	**-0.11****	0.06	**A – C***** **C – D****
	Science and Social Studies Ability		**-0.11****	--	0.11	B – D*
SOCIAL-EMOTIONAL						
Parent Report						
Head Start Year (9)	Hyperactive Behavior	-0.33**	-0.46**	0.27	-0.77*	A – C* B – C** C – D**
Age 4 Year (9)	Hyperactive Behavior	**-0.37*****	-0.03	0.39**	0.27	A – B* **A – C***** A – D* B – C*
	Total Problem Behavior	-0.61**	--	0.38	--	A – C*
	Closeness	0.26	0.13	--	1.57**	A – D* B – D*

Exhibit 8.15: Significant Subgroup Impacts and Differences in Impact for the 3-Year-Old Cohort: Parent/Caregiver Reported Depressive Symptoms (continued)

Year (Number of Outcomes Examined)	Outcome Measure	Impact in Subgroup A (Head Start – Control) No Symptoms	Impact in Subgroup B (Head Start – Control) Mild Symptoms	Impact in Subgroup C (Head Start – Control) Moderate Symptoms	Impact in Subgroup D (Head Start – Control) Severe Symptoms	Statistically Significant Differences in Impacts Between Subgroups
Kindergarten (9)	Closeness	0.83**	-0.65**	-0.90**	--	A – B*** A – C***
	Social Competencies	0.22*	-0.10	--	-0.39	A – B * A – D*
	Social Skills and Positive Approaches to Learning	**0.46*** **	-0.06	--	--	A – B **
	Hyperactive Behavior	-0.31**	--	0.34	--	A – C**
	Aggressive Behavior	-0.27**	--	--	0.68	A – D **
1st grade (9)	--	--	--	--	--	--
Teacher Report						
Head Start Year (0)						
Age 4 Year (0)						
Kindergarten (11)	ASPI-Problem with Peer Interaction	0.08	2.45*	5.01**	-2.89	A – B* A – C** C – D**
	Conflict	--	1.82*	1.97*	-2.84	B – D* C – D**
	Positive Relationships	--	-2.55**	-3.38**	3.65	B – D* C – D**
	ASPI-Aggressive	-0.28	--	3.23**	-1.69	A – C** C – D*
	ASPI-Oppositional	--	--	2.44*	-3.34	C – D**

Exhibit 8.15: Significant Subgroup Impacts and Differences in Impact for the 3-Year-Old Cohort: Parent/Caregiver Reported Depressive Symptoms (continued)

Year (Number of Outcomes Examined)	Outcome Measure	Impact in Subgroup A (Head Start – Control) No Symptoms	Impact in Subgroup B (Head Start – Control) Mild Symptoms	Impact in Subgroup C (Head Start – Control) Moderate Symptoms	Impact in Subgroup D (Head Start – Control) Severe Symptoms	Statistically Significant Differences in Impacts Between Subgroups
1st grade (11)	ASPI-Withdrawn/Low Energy	0.07	1.98**	--	--	A – B*
	ASPI-Shy/Social Reticent	0.39	**2.36*** **	--	-2.18	A – B* B – D**
	Positive Relationships	1.11*	-1.55	-3.32*	5.27*	A – B* A – C** B – D* C – D**
	Conflict	-0.53	0.90	2.79**	-4.19*	A – C** A – D* B – D** C – D**
	ASPI- Problems with Structured Learner	0.01	**3.08*** **	-1.39	-3.30	A – B** B – C* B – D**
HEALTH						
Head Start Year (5)	Child Has Health Insurance Coverage	--	--	-0.03	0.12*	C – D**
	Child Received Dental Care	--	--	0.08	**0.28*** **	C – D*
Age 4 Year (5)	Child's Overall Health Status Is Excellent/Good	0.04	-0.02	**-0.16*** **	0.11	**A – C*** ** B – C* C – D**
	Child Has Health Insurance Coverage	0.03	0.03	**-0.06** **	--	**A – C** ** B – C*
Kindergarten (5)	--	--	--	--	--	--

Exhibit 8.15: Significant Subgroup Impacts and Differences in Impact for the 3-Year-Old Cohort: Parent/Caregiver Reported Depressive Symptoms (continued)

Year (Number of Outcomes Examined)	Outcome Measure	Impact in Subgroup A (Head Start – Control) No Symptoms	Impact in Subgroup B (Head Start – Control) Mild Symptoms	Impact in Subgroup C (Head Start – Control) Moderate Symptoms	Impact in Subgroup D (Head Start – Control) Severe Symptoms	Statistically Significant Differences in Impacts Between Subgroups
1st grade (5)	Child Had Care for Injury in Last Month	0.01	-0.01	**0.11****	--	A – C* **B – C****
	Child Needs Ongoing Care	--	0.06	--	-0.15*	B – D**
PARENTING						
Parent Report						
Head Start Year (5)	Parent Read to Child in Last Week	--	--	**0.16*****	-0.03	C – D*
	Parental Safety Practices Scale	**0.08*****	-0.02	-0.06	--	**A – B**** **A – C*****
Age 4 Year (9)	Parental Safety Practices Scale	0.06**	-0.04	--	--	A – B*
Kindergarten (9)	Parent Read to Child in Last Week	-0.03	--	0.13*	--	A – C*
	Parental Safety Practices Scale	-0.01	0.09**	--	--	A – B**
	Parenting Style: Neglect	-0.05**	--	--	0.09	A – D**
	Parenting Style: Permissive	--	-0.08	0.10*	--	B – C**
	Parent Used Time Out in Last Week	--	--	-0.17**	0.14	C – D**
1st grade (8)	Parent Spanked Child in Last Week	-0.08**	--	--	0.10	A – D**
Teacher Report						
Head Start Year (0)						
Age 4 Year (0)						

Exhibit 8.15: Significant Subgroup Impacts and Differences in Impact for the 3-Year-Old Cohort: Parent/Caregiver Reported Depressive Symptoms (continued)

Year (Number of Outcomes Examined)	Outcome Measure	Impact in Subgroup A (Head Start – Control) No Symptoms	Impact in Subgroup B (Head Start – Control) Mild Symptoms	Impact in Subgroup C (Head Start – Control) Moderate Symptoms	Impact in Subgroup D (Head Start – Control) Severe Symptoms	Statistically Significant Differences in Impacts Between Subgroups
Kindergarten (2)	School Contact and Communication	-0.00	0.06	**-0.13****	0.10	A – C** B – C* C – D**
1st grade (2)	--	--	--	--	--	--

Key:
******* $p \leq 0.01$
** $p \leq 0.05$
***** $p \leq 0.10$

Bold exhibit entry indicates that the impact passes the Benjamini-Hochberg test for multiple comparisons with a 10 percent false discovery rate.

Gray cell indicates no data available.

Exhibit 8.16: Significant Subgroup Impacts and Differences in Impact for the 3-Year-Old Cohort: Household Risk Index

Year (Number of Outcomes Examined)	Outcome Measure	Impact in Subgroup A (Head Start – Control) Low/No Household Risk	Impact in Subgroup B (Head Start – Control) Moderate Household Risk	Impact in Subgroup C (Head Start – Control) High Household Risk	Statistically Significant Differences in Impacts Between Subgroups
COGNITIVE					
Direct Child Assessment					
Head Start Year (11)	Counting Bears	**0.06****	-0.07	--	A – B*
	CTOPPP Elision	**7.81****	-6.14	--	A – B**
	McCarthy Draw-a-Design	**0.16****	0.02	0.59**	A – C*
					B – C**
	WJ III Spelling	--	-1.67	8.84**	B – C**
Age 4 Year (11)	WJ III Letter Word Identification	4.28**	-3.00	--	A – B*
	CTOPPP Elision	7.19**	6.95	**25.78*****	A – C**
					B – C*
	WJ III Pre-Academic Skills	2.77*	-4.65	--	A – B**
Kindergarten (12)	WJ III Applied Problems	0.30	-5.08**	--	A – B*
	WJ III Math Reasoning	-0.16	-3.94**	--	A – B*
	Letter Naming	-0.04	-2.05**	1.84	A – B**
					B – C***
	WJ III Spelling	0.30	-1.68	8.57**	A – C*
					B – C*
1st grade (15)	WJ III Calculation	0.79	-4.71*	--	A – B**
	WJ III Spelling	-0.58	-8.56**	8.85**	A – B*
					A – C**
					B – C***
	WJ III Academic Skills	--	-5.46	7.26*	B – C**
	WJ III Passage Comprehension	--	-4.85	7.87**	B – C**
	WJ III Pre-Academic Skills	--	-4.01	7.62*	B – C**
	WJ III Academic Applications	--	-2.77	5.44*	B – C**

Exhibit 8.16: Significant Subgroup Impacts and Differences in Impact for the 3-Year-Old Cohort: Household Risk Index (continued)

Year (Number of Outcomes Examined)	Outcome Measure	Impact in Subgroup A (Head Start – Control) Low/No Household Risk	Impact in Subgroup B (Head Start – Control) Moderate Household Risk	Impact in Subgroup C (Head Start – Control) High Household Risk	Statistically Significant Differences in Impacts Between Subgroups
School Performance					
Head Start Year (0)					
Age 4 Year (0)					
Kindergarten (5)	--	--	--	--	--
1st grade (5)	--	--	--	--	--
SOCIAL-EMOTIONAL					
Parent Report					
Head Start Year (9)	Total Problem Behavior	--	-1.08**	0.38	B – C*
Age 4 Year (9)	Aggressive Behavior	-0.04	-0.50*	--	A – B*
Kindergarten (9)	--	--	--	--	--
1st grade (9)	Aggressive Behavior	0.04	-0.57**	0.15	A – B** B – C*
	Hyperactive Behavior	-0.01	**-0.50*** **	--	A – B**
	Total Problems Behavior	0.16	**-1.59*** **	--	**A – B*** **
Teacher Report					
Head Start Year (0)					
Age 4 Year (0)					
Kindergarten (11)	Conflict	-0.40	--	3.05*	A – C*
	Positive Relationships	-0.04	--	-4.06*	A – C*
1st grade (11)	ASPI—Problems with Peer Interaction	0.46	-4.09**	--	A – B**

Exhibit 8.16: Significant Subgroup Impacts and Differences in Impact for the 3-Year-Old Cohort: Household Risk Index (continued)

Year (Number of Outcomes Examined)	Outcome Measure	Impact in Subgroup A (Head Start – Control) Low/No Household Risk	Impact in Subgroup B (Head Start – Control) Moderate Household Risk	Impact in Subgroup C (Head Start – Control) High Household Risk	Statistically Significant Differences in Impacts Between Subgroups
HEALTH					
Head Start Year (5)	Child Received Dental Care	**0.19*** **	0.01	**0.25** **	**A – B ** **
					B – C **
	Child Had Care for Injury in Last Month	0.01	**-0.11** **	0.09	**A – B** **
					B – C **
Age 4 Year (5)	Child Had Care for Injury in Last Month	0.02	--	0.12**	A – C*
	Child Needs Ongoing Care	-0.01	0.09*	-0.07	A – B*
					B – C*
Kindergarten (5)	--	--	--	--	--
1st grade (5)	--	--	--	--	--
PARENTING					
Parent Report					
Head Start Year (5)	--	--	--	--	--
Age 4 Year (9)	--	--	--	--	--
Kindergarten (9)	--	--	--	--	--
1st grade (8)	Parent Spanked Child in Last Week	-0.01	**-0.15*** **	0.12	A – B **
					B – C * **
Teacher Report					
Head Start Year (0)					
Age 4 Year (0)					
Kindergarten (2)	--	--	--	--	--
1st grade (2)	--	--	--	--	--

Key:
*** $p \leq 0.01$
** $p \leq 0.05$
* $p \leq 0.10$

Bold exhibit entry indicates that the impact passes the Benjamini-Hochberg test for multiple comparisons with a 10 percent false discovery rate.
Gray cell indicates no data available.

Exhibit 8.17: Significant Subgroup Impacts and Differences in Impact for the 3-Year-Old Cohort: Urbanicity

Year (Number of Outcomes Examined)	Outcome Measure	Impact in Subgroup A (Head Start – Control) Not Urban	Impact in Subgroup B (Head Start –Control) Urban	Statistically Significant Differences in Impacts Between Subgroups
COGNITIVE				
Direct Child Assessment				
Head Start Year (11)	WJ III Pre-Academic Skills	**8.37*** **	3.21**	A – B*
	McCarthy Draw-a-Design	**0.42*** **	0.09	A – B**
	Letter Naming	**3.11*** **	1.18**	A – B**
	WJ III Letter Word Identification	**11.75*** **	**4.72*** **	A – B**
Age 4 Year (11)	WJ III Applied Problems	5.21**	-0.34	A – B*
Kindergarten (12)	WJ III Spelling	6.05*	-0.94	A – B*
1st grade (15)	WJ III Pre-Academic Skills	7.06**	-1.45	A – B**
	WJ III Spelling	6.04*	-3.01	A – B**
	WJ III Academic Applications	5.00**	-0.34	A – B*
	WJ III Word Attack	10.08*	-3.29	A – B**
	WJ III Writing	4.67**	-1.28	A – B**
	WJ III Basic Reading	9.64*	-2.53	A – B**
	WJ III Math Reasoning	5.04**	0.26	A – B*
School Performance				
Head Start Year (0)				
Age 4 Year (0)				
Kindergarten (5)	--	--	--	--
1st grade (5)	--	--	--	--
SOCIAL-EMOTIONAL				
Parent Report				
Head Start Year (9)	Aggressive Behavior	**-0.40****	-0.02	A – B**
	Social Competencies	**0.28****	-0.11	**A – B*** **
	Total Problem Behavior	**-1.04*** **	-0.38**	A – B*
Age 4 Year (9)	Aggressive Behavior	-0.45**	-0.05	A – B*
Kindergarten (9)	--	--	--	--
1st grade (9)	--	--	--	--

Exhibit 8.17: Significant Subgroup Impacts and Differences in Impact for the 3-Year-Old Cohort: Urbanicity (continued)

Year (Number of Outcomes Examined)	Outcome Measure	Impact in Subgroup A (Head Start – Control) Not Urban	Impact in Subgroup B (Head Start –Control) Urban	Statistically Significant Differences in Impacts Between Subgroups
Teacher Report				
Head Start Year (0)				
Age 4 Year (0)				
Kindergarten (11)	--	--	--	--
1st grade (11)	--	--	--	--
HEALTH				
Head Start Year (5)	Child Had Care for Injury in Last Month	**0.08****	-0.03*	**A – B*****
Age 4 Year (5)	Child's Overall Health Status Is Excellent/Good	**0.11****	-0.02	**A – B****
Kindergarten (5)	Child Needs Ongoing Care	**-0.11****	-0.01	A – B**
1st grade (5)	--	--	--	--
PARENTING				
Parent Report				
Head Start Year (5)	Parental Safety Practices Scale	**0.09****	0.02	A – B *
Age 4 Year (9)	--	--	--	--
Kindergarten (9)	--	--	--	--
1st grade (8)	--	--	--	--
Teacher Report				
Head Start Year (0)				
Age 4 Year (0)				
Kindergarten (2)	--	--	--	--
1st grade (2)	--	--	--	--

Key:
******* $p \leq 0.01$
** $p \leq 0.05$
* $p \leq 0.10$

Bold exhibit entry indicates that the impact passes the Benjamini-Hochberg test for multiple comparisons with a 10 percent false discovery rate.

Gray cell indicates no data available.

Chapter 9: Conclusions

Introduction

Head Start seeks to improve educational and developmental outcomes for children from severely economically disadvantaged families. To be eligible, children must come from families that have incomes that "are below the poverty line, or are eligible or, in the absence of child care, would potentially be eligible for public assistance" {Sec. 645. [42 U.S.C. 9840] (a)(1)}. As stated in the *Head Start Act*,[116] the intent of the program is to "promote the school readiness of low-income children by enhancing their cognitive, social, and emotional development (1) in a learning environment that supports children's growth in language, literacy, mathematics, science, social and emotional functioning, creative arts, physical skills, and approaches to learning and (2) through the provision to low-income children and their families of health, educational, nutritional, social, and other services that are determined, based on family needs assessments, to be necessary." {Sec. 645. [42 U.S.C. 9840] (a)(1)}

This chapter summarizes the results discussed in this report to provide an overall cohesive story about what we have learned regarding the extent to which access to Head Start meets these stated goals. The narrative begins with the context within which impacts on child and family outcomes are expected to occur through an assessment of how the availability of Head Start shapes the preschool and early elementary school experiences of newly entering 3- and 4-year-old eligible children. We then examine the extent to which Head Start supports children's school readiness by looking at its impact on children's cognitive, social-emotional, and health outcomes, and on the parenting practices of their primary caregivers, at the end of the Head Start years when children are about to enter kindergarten. Next, we examine how these early effects carry through into the early years of elementary school by examining the subsequent impact of access to Head Start on children's development and school performance at the end of kindergarten and 1st grade. Finally, we examine how impacts vary by child and family characteristics at the end of preschool and again at the end of kindergarten and 1st grade.

[116] Public Law 110-134 *Improving Head Start for School Readiness Act of 2007*, signed into law on December 12, 2007.

How Does Gaining Access to Head Start Affect Children's Head Start and Early Elementary School Experiences?

Head Start Experiences

Providing access to Head Start was found to have a positive impact on children's preschool experiences. There are 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 when they were initially 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 are quite large, but are also affected 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).[117] 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 age 4 year. Most of the children (both Head Start and control group) were in some type of center-based care by the second 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 experience measures all but vanished.

While on average having access to Head Start resulted in more positive experiences for children, not all children in the Head Start group had the same quality of experience. The majority (70 percent) of Head Start children in both cohorts were in centers with overall average ECERS-R scores of at least a five on a seven-point scale, indicating a good or better environment. Most children were also in classrooms that emphasized language and literacy and math activities—approximately 60 percent were in classrooms that provided an emphasis on language and literacy and math activities (in which teachers reported providing more than half of

[117] For these analyses, children in parent care were included and given a score of zero.

a list of 12 language and literacy activities and eight math activities three times per week). Conversely, the remaining Head Start group children did not experience centers with such high ECERS-R scores or as strong an emphasis on language and literacy or math activities. There is also diversity in the training and qualification of the teachers serving the Head Start group children. Approximately 30 percent of the Head Start children had teachers with a BA degree, and 30 percent had teachers with at least an associate's degree, leaving approximately 40 percent of the children with teachers without a postsecondary degree. Slightly more than one-third of the 3-year-old cohort, and 40 percent of the 4-year-old cohort, assigned to the Head Start group had teachers who had received 25 hours or more of training in the last year. Consequently, the nature and quality of the treatment varies—for some children it is very good, while for other children it could be less so. Both the overall high quality, on average, and the variation may be important in understanding impacts on child and family outcomes.

Experiences in Kindergarten and 1st Grade

For this study, the data collected on children's school age experiences was limited—a subsequent examination of children through the end of 3rd grade will shed greater light on this issue. Based on the information available, the results show that providing access to Head Start did not appear to have much impact on the types of schools children attended. On average, both Head Start and control group children attended public schools where the percentage of students at or above proficient on state assessments in math and reading were in the middle of the respective state averages, indicating that most of 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 between the Head Start and control group children's school's average math proficiency scores, favoring the control 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 meals) and were in schools with higher proportions of minority 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 few differences that were found varied, sometimes favoring the control group and sometimes favoring the Head Start group.

What Is the Impact of Head Start on Children and Families?

Impacts at the End of the Head Start Years

There is clear evidence that Head Start has a statistically significant impact on children's language and literacy development while children are in Head Start, and these effects, albeit modest in magnitude, are found for both age cohorts during their first year of admission to the program. However, at the end of the age 4 year for the younger children (the 3-year-old cohort) the overall positive effects are limited to a single language outcome (CTOPPP).

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

With regard to children's social-emotional development, the story is primarily about the younger children. There are no detectable impacts in the social-emotional domain for the 4-year-old cohort. Yet, there is strong evidence that Head Start leads to reductions in 3-year-olds' problem behaviors and hyperactivity and suggestive evidence of an increase in social skills and positive approaches to learning at the end of the age 4 year.

In the health domain, there is very strong evidence of relatively large impacts on children's receipt of dental care for both age cohorts. The effects for the 3-year-olds are found at the end of both the tHead Start year and the age 4 year. For the 3-year-old cohort, there also is evidence of a positive impact on children's health status (as reported by parents) at the end of the Head Start year.

Finally, the evidence related to the program's impact on parenting practices is, like the impacts on children's social-emotional development, primarily a story about the younger children. For children in the 4-year-old cohort, the only observed effect is moderate evidence of less parental use of time out as a discipline strategy. Alternatively, for the 3-year-old cohort there is (1) very strong evidence of an increase in the frequency of both parents' reading to their child and involving them in cultural enrichment activities at the end of the Head Start year, (2) very strong evidence of less use of spanking also at end of the Head Start year, and (3) very strong evidence of less use of the negative authoritarian parenting style at the end of the age 4 year.

Impacts at the End of Kindergarten and 1st Grade

Despite the early, positive, cognitive effects, subsequent direct assessments and teacher ratings show only weak evidence of an impact and only at the end of 1st grade. There was only a single remaining increase in language skills for the 3-year-old cohort and for the 4-year-old cohort in 1st grade. These impacts were on measures of oral comprehension and receptive vocabulary for the 3-year-old and 4-year-old cohorts, respectively.

This pattern of limited cognitive impacts in the school years may suggest that the magnitude of the initial cognitive impacts may not have been sufficiently potent for the early gains Head Start children made to be sustained as they developed and moved into the elementary school years.

With regard to social-emotional development, the early impacts found for the 3-year-old cohort appear to have continued into early elementary school. At the end of kindergarten, there is again evidence of less parent-reported hyperactive behavior and better social skills and positive approaches to learning (although the latter is merely suggestive evidence). At the end of 1st grade, there is moderate evidence of an increase in parent-child closeness and suggestive evidence of improved child-parent relationships.

For the 4-year-old cohort, the picture is less clear. Teachers report that Head Start children are more reticent and have poorer teacher-child relationships at the end of 1st grade than control group children. Yet, parents' reports suggest that the Head Start group is less withdrawn in this same year.

With regard to health, there is evidence of a positive impact on the availability of health insurance for children in both age cohorts at the end of kindergarten and at the end of 1st grade for the 4-year-old cohort (although the strength of the evidence varies by year and cohort). There also is suggestive evidence at the end of kindergarten, for the 4-year-old cohort, on parent-reported overall health status of the child.

Finally, with regard to parenting practices, there is again a story that is concentrated in the younger cohort. For the 3-year-old cohort, there is (1) moderate evidence of less use of time out as a discipline strategy at the end of kindergarten, (2) suggestive evidence of less use of both spanking at the end of kindergarten and time out at the end of 1st grade, and again, (3) moderate evidence of less use of the authoritarian parenting style at the end of 1st grade. These impacts on

parenting are particularly relevant, given the pattern of impacts on social emotional outcomes for the 3-year-old cohort.

Variation in Impacts

In addition to looking at Head Start's average impact on the diverse set of children and families who participate in the program, it is important to understand how impacts vary among different types of participants. In particular, there is evidence that Head Start has a differential impact on some outcomes for some subgroups of children over others. The subgroup findings are exploratory and use a modified level of evidence compared to the main impact findings.

Many subgroups experienced sustained positive benefits from Head Start in one or all of the four domains (cognitive, social-emotional, health, and parenting practices) through 1st grade. Among the 4-year-olds, these subgroups include children of parents with mild depressive symptoms, and Dual Language Learners. Among the 3-year-old cohort, they include children with special needs, children of parents with no depressive symptoms, children from higher risk households, and children in non-urban settings. Black children in the 4-year-old cohort and children in the lowest academic quartile at baseline and Dual Language Learners in the 3-year-old cohort also experienced impacts, though they were not sustained through 1st grade. We also identified several subgroups that experienced a mixture of favorable and unfavorable impacts, including 4-year-olds in the lowest academic quartile at baseline and 3-year-olds from households with a moderate number of risk factors.

Finally, we identified a few subgroups of concern, that is, subgroups that experienced solely—or primarily—negative impacts of Head Start. The group that showed the most widespread negative impacts was 3-year-olds whose parents had moderate depressive symptoms. These children experienced consistent, sustained negative impacts throughout the life of the study on those variables for which there were differential impacts across subgroups. These impacts were found across domains and across reporters.

Discussion

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 examined the impacts of Head Start on these four domains.

The study shows that providing access to Head Start led to improvements in the quality of the early childhood settings and programs children experienced. On nearly every measure of quality traditionally used in early childhood research, the Head Start group had more positive experiences than those in the control group.

These impacts on children's experiences translated into favorable impacts of moderate size at the end of one year in the domains of children's cognitive development and health as well as in parenting practices. There were more significant findings across the measures within these domains for 3-year-olds in that first year, and the 3-year-old cohort also experienced a decrease in problematic behavior in the social-emotional domain. Yet, by the end of 1st grade, there were few significant differences between the Head Start group as a whole and the control group as a whole for either cohort.

The differences at the end of 1st grade included a favorable impact on receipt of health insurance for the 4-year-old cohort, consistent with earlier impacts on health insurance for both cohorts. Further, there are longer-term favorable impacts related to children's social emotional development and relationships with their parents for the 3-year-old cohort. According to parent report, this younger group experienced favorable impacts on behavioral and social emotional outcomes during the early years of the program and into kindergarten. By 1st grade, these impacts were limited to outcomes related to parent-child relationships and parenting practices. It is possible that these benefits in the parent-child relationship are both related to earlier improvements in behavior and may lead to longer term benefits for children. However, this is only one hypothesis, and the issue requires further analysis.

This study evaluated the Head Start program 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 did receive 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. Consequently, to achieve measurable impacts, Head Start (as noted above) has

to outperform what children would receive in the absence of the program's availability. Improved child care and pre-K standards across the nation may have contributed to the lack of significant differences found between the Head Start and control group children.

Second, although the quality is high on average, Head Start programs vary 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 are not trying their best to improve children's development in these areas, but rather on average the current program may not be 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 variation may have contributed to the lack of statistically significant differences in the cognitive domain.

Head Start has always had a particular emphasis on young children with special needs, and indeed we see secondary impacts through 1st grade in the two subgroups of relevance: children whose parents have been told their child has special needs or disabilities (for 3-year-olds) and children with the lowest cognitive skills upon entering Head Start (for the 4-year-olds). Thus, Head Start has benefits for these groups of children that last into the early elementary school

Similarly, the Head Start performance standards emphasize the importance of respecting and individualizing services for children as needed based on their cultural and linguistic backgrounds. The findings from this study demonstrate that Black children (in the 4-year-old cohort) and Dual Language Learners are among the groups that benefited from access to Head Start. However, most of these impacts only lasted through kindergarten.

The subgroup findings do not present a consistent picture of favorable impacts for groups that have traditionally been emphasized as higher risk. While the children from higher risk households benefited in the 3-year-old group, there were no patterns of differences in impacts by household risk for the 4-year-olds. Further, it was the children of caregivers with less severe depressive symptoms that experienced favorable impacts through 1st grade in both cohorts. There also appears to be a pattern in which, for the 3-year-old cohort, the children from families in the middle of the risk categories (neither highest nor lowest) actually experienced some negative impacts. This pattern is particularly strong, and concerning, for children of caregivers

with moderate levels of depressive symptoms. The explanation for these patterns is unclear and warrants more attention.

Finally, this study also found that, in the 3-year-old cohort, Head Start had benefits through 1st grade for children from non-urban communities. It is possible that this finding represents the difficulties that children and families in non-urban communities have in getting comprehensive services and in finding quality early care and education for their children absent Head Start. Indeed, children were more likely to participate in a second year of Head Start if there was less competition from other preschools in the area. These are questions that should be pursued in future research.

In sum, this report finds that providing access to Head Start has benefits for both 3-year-olds and 4-year-olds in the cognitive, health, and parenting domains, as well as in the social emotional domain for 3-year-olds only. However, averaging across all children, the benefits of access to Head Start at age four are largely absent by 1st grade. For 3-year-olds, there are few sustained benefits, although access to the program may lead to improved parent-child relationships through first grade, a potentially important finding for children's longer term development. Moreover, several subgroups of children in this study experience benefits of Head Start into 1st grade. It will be important in future research to examine whether the positive parent-child relationships for the 3-year-old cohort translate into improved outcomes as children get older, as well as whether the findings for subgroups of children persist over the longer term.

To that end, the study children have been followed through 3rd grade. The 3rd grade report will examine the extent to which impacts of Head Start on initial school readiness are altered or maintained as children enter pre-adolescence. Further, that report will provide a greater focus on how children's later experiences in the school and community affect their outcomes at 1st and 3rd grade. 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 types of programs, center, classrooms, and other experiences relate to more positive impacts for children and families? What accounts for the subgroup patterns observed in this report? Are there some later experiences that help to sustain impacts through the early elementary grades? Will the sustained improvements in children's social-emotional outcomes and parenting lead to longer term benefits for children?

Hopefully, researchers will take advantage of the data from this study, which will be made available through a data archive, to further the understanding of the role Head Start plays in the well-being of children and families.

REFERENCES

Abbott-Shim, M., Lambert, R., & McCarty, F. (2003). A comparison of school readiness outcomes for children randomly assigned to a Head Start program and the program's wait list. *Journal of Education for Students Placed at Risk, 8*(2), 191-214.

Achenbach, T. M., Edelbrock, C., & Howell, C. T. (1987). Empirically based assessment of the behavioral/emotional problems of 2-3-year old children. *Journal of Abnormal Child Psychology, 15,* 629-650.

Advisory Committee on Head Start Research and Evaluation. (1999). *Evaluating Head Start: A recommended framework for studying the impact of the Head Start program.* Washington, DC: U.S. Department of Health and Human Services.

Amato, P. R., & Fowler, F. (2002). Parenting practices, child adjustment, and family diversity. *Journal of Marriage and Family, 64*(3), 703-717.

American Academy of Pediatrics, American Public Health Association, and National Resource Center for Health and Safety in Child Care and Early Education. (2002). *Caring for our children: National health and safety performance standards: Guidelines for out-of-home child care programs, 2nd edition*. Elk Grove Village, IL: American academy of pediatrics and Washington, DC: American Public Health Association.

Aos, S., Lieb, R., Mayfield, J., Miller, M., & Pennucci, A. (2004). *Benefits and costs of prevention and early intervention programs for youth*. Document #: 04-07-3901: Olympia, WA: Washington State Institute for Public Policy.

Arnett, J. (1989). Caregivers in day-care centers: Does training matter? *Journal of Applied Developmental Psychology, 10,* 541-552.

Barnett, W. S. (2002). *The battle over Head Start: What the research shows.* Presentation at a Science and Public Policy Briefing sponsored by the Federation of Behavioral, Psychological, and Cognitive Sciences, September 13, 2002.

Barnett, W. S., Lamy, C., & Jung, K. (2005). *The effects of state prekindergarten programs on young children's school readiness in five states.* New Brunswick, NJ: Rutgers University, The National Institute for Early Education Research\.

Barnett, W. S., Hustedt, J., Robin, K., & Schulman, K. (2005). *The state of preschool: 2005 state preschool yearbook.* New Brunswick, NJ: Rutgers University, National Institute for Early Education Research.

Baumrind, D. (1971). Cultural patterns of paternal authority. *Developmental Psychology Monograph, 4*.

Benjamini, Y., & Hochberg, Y. (1995). Controlling for the false discovery rate: A practical and powerful approach to multiple testing. *Journal of the Royal Statistical Society, Series B (Methodological), 57*(1), 289-300.

Biemiller, A. (2006). Vocabulary development and instruction: A prerequisite for school learning. In D. K. Dickinson & S. Neuman (Eds.), *Handbook of early literacy research: Vol 2, Chapter 3* (pp.41-51). New York: The Guilford Press.

Blair, C. (2002). School readiness: Integrating cognition and emotion in a neurobiological conceptualization of children's functioning at school entry. *American Psychologist, 57*, 111-127.

Block, J. H. (1965). *The child-rearing practices report (CRPR): A set of Q items for the description of parental socialization attitudes and values.* Berkley, CA: University of California at Berkley, Institute of Human Development.

Bloom, H. S. (1984). Accounting for no-shows in experimental evaluation designs. *Evaluation Review, 8*(2), 225-246.

Campbell, D. T., & Erlebacher, A. (1970). How regression artifacts in quasi-experimental evaluations can mistakenly make compensatory education look harmful. In J. Hellmuth (Ed.), *Disadvantaged Child*. New York: Brunner/Mazel.

Cappella, E., & Larner, M. B. (1999). America's schoolchildren: Past, present and future. *The Future of Children, 9*(2), 21-29.

Carnegie Corporation of New York. (1994). *Starting points: Meeting the needs of our youngest children.* New York: Author.

Charlesworth, R., Hart, C. H., Burts, D. C., Mosley, J. & Fleege, P. O. (1993). Measuring the developmental appropriateness of kindergarten teachers' beliefs and practices. *Early Childhood Research Quarterly, 8*, 255-276.

Cicirelli, V. G. (1969). *The impact of Head Start: An evaluation of the effects of Head Start on children's cognitive and affective development*. Athens, OH: Westinghouse Learning Corporation.

Cohen, J. (1987). *Statistical power analysis for the behavioral sciences* (2nd ed.). Hillsdale, NJ: Erlbaum.

Coolahan, K., McWayne, C., Fantuzzo, J., & Grim, S. (2002). Validation of a multidimensional assessment of parenting styles for low-income African-American families with preschool children. *Early Childhood Research Quarterly, 17*, 356-373.

Cunningham, A. E., & Stanovich, K. E. (1997). Early reading acquisition and its relation to reading experience and ability 10 years later. *Developmental Psychology,* 33, 934-945.

Deater-Deckard, K., Dodge, K. A., Bates, J. E., & Pettit, G. S. (1996). Physical discipline among African American and European American mothers: Links to children's externalizing behaviors. *Developmental Psychology, 32,* 1065-1072.

Denton, K., Reaney, L., & West, J. (2001). *Home educational activities, literacy resources and kindergartners' reading knowledge and skills*. Presentation at the biennial meeting of the Society for Research in Child Development, Minneapolis, MN.

Developing skills checklist—home inventory. (1990). Monterey, CA: CTB/McGraw-Hill.

Duncan, G. J., Dowsett, C. J., Claessens, A., Magnuson, K., Huston, A. C., Klebanov, P., Pagani, L. S., Feinstein, L., Engel, M., Brooks-Gunn, J., Sexton, H., Duckworth, K., & Japel, C. (2007). School readiness and later achievement. *Developmental Psychology, 43*(6), 1428-1446.

Dunn, L. M., Dunn, L. L., & Dunn, D. M. (1997). *Peabody picture and vocabulary test, third edition (PPVT).* Circle Pines, MN: American Guidance Services.

Dunn, L. M., Padilla, E. R., Lugo, D. E., & Dunn, L. L. (1986). *Test de vocabulario en imágenes peabody*. Circle Pines, MN: American Guidance Service.

Eamon, M. K. (2005). Social-demographic, school, neighborhood, and parenting influences on the academic achievement of Latino young adolescents. *Journal of Youth and Adolescence, 34*, 163-174.

Early, D. M., Maxwell, K. L., Burchinal, M., Alva, S., Bender, R. H., Bryant, D., Cai, K., Clifford, R. M., Ebanks, C., Griffin, J. A., Henry, G. T., Howes, C., Iriondo-Perez, J., Jeon, H. J., Mashburn, A. J., Peisner-Feinberg, E., Pianta, R. C., Vandergrift, N., & Zill, N. (2007). Teachers' education, classroom quality, and young children's academic skills: Results from seven studies of preschool programs. *Child Development, 78*(2), 558-580.

Ehri, L. C., & Roberts, T. (2006). The roots of learning to read and write: Acquisition of letters and phonemic awareness. In D. K. Dickinson & S. Neuman (Eds.), *Handbook of early literacy research: Vol. 2, Chapter 9* (pp. 113-131). New York: The Guilford Press.

FACES Research Team. Color names and counting. Modified from the Color concepts and number concepts tasks in J. M. Mason & J. Stewart (1989), *The CAP early childhood diagnostic instrument* (prepublication edition), American Testronics.

FACES Research Team. Letter naming task. Modified from a test used in the Head Start Quality Research Center's curricular intervention studies.

FACES Research Team. Story and print concepts. Modified from the Story and print concepts in J. M. Mason & J. Stewart (1989), *The CAP early childhood diagnostic instrument* (prepublication edition), American Testronics.

FACES Research Team. Writing sample. Modified from the Name writing tasks in J. M. Mason & J. Stewart (1989), *The CAP early childhood diagnostic instrument* (prepublication edition), American Testronics.

Finn, J. D., & Achilles, C. M. (1990). Answers and questions about class size: A statewide experiment. *American Educational Research Journal, 27*(3), 557-577.

Gennetian, L. A., Morris, P. A., Bos, J. M., & Bloom, H. S. (2005). Constructing instrumental variables from experimental data to explore how treatments produce effects. In H. S. Bloom (Ed.), *Learning more from social experiments: Evolving analytic approaches.* New York: Russell Sage Foundation.

Glass, G. V., McGaw, B., & Smith, M. L. (1981). *Meta-analysis in social research.* London: Sage.

Goal One Technical Planning Group. (1991). The Goal One Technical Planning Subgroup report on school readiness. In National Education Goals Panel (Ed.), *Potential strategies for long-term indicator development: Reports of the technical planning subgroups* (Report No. 91-0, pp. 1-18). Washington, DC: National Education Goals Panel.

Goal One Technical Planning Group. (1993). *Reconsidering children's early development and learning: Toward shared beliefs and vocabulary. Draft report to the National Education Goals Panel.* Washington, DC: National Education Goals Panel.

Gormley, W. T. Jr. (2007). Early childhood care and education: Lessons and puzzles. *Journal of Policy Analysis and Management, 26,* (3), 633-671.

Gormley, W. T. Jr., Gayer, T., Phillips, D., & Dawson, B. (2005). The effects of universal pre-K on cognitive development. *Developmental Psychology, 41*(6), 872-884.

Graham, S., & Perin, D. (2007). *Writing next: Effective strategies to improve writing of adolescents in middle and high schools.* New York, NY: The Carnegie Corporation of New York.

Harms, T., & Clifford, R. M. (1989). *Family day care rating scale (FDCRS).* New York, NY: Teachers College Press.

Harms, T., Clifford, R. M., & Cryer, D. (1998). *Early childhood environment rating scale-revised edition (ECERS-R).* New York, NY: Teachers College Press.

Hubbs-Tait, L., Culp, A. M., Huey, E., Culp, R., & Miller, C. (2002). Relation of maternal cognitive stimulation, emotional support, and intrusive behavior during Head Start to children's kindergarten cognitive abilities. *Child Development, 73,* 110-131.

Hubbs-Tait, L., Culp, A. M., Huey, E., Culp, R. E., Starost, H., & Hare, C. (2002). Relation of Head Start attendance to children's cognitive and social outcomes: Moderation by family risk. *Early Childhood Research Quarterly, 17*(4), 539-558.

Institute for Policy Research News, Northwestern University. (2006). Evidence of children's achievement: IPR experts hold policy briefing on Capitol Hill. *Fall, 28*(1). Retrieved from http://www.northwestern.edu/ipr/publications/newsletter/iprn0608/DCbriefing.html

Kagan, S. L., Moore, E., & Bredekamp, S. (Eds.). (1995). *Considering children's early development and learning: Toward shared belief and vocabulary.* Washington, DC: National Education Goals Panel.

Kreisman, M. B. (2003). Evaluating academic outcomes of Head Start: An application of general growth mixture modeling. *Early Childhood Research Quarterly, 18*(2), 238-254.

Johnson, H. P., O'Brien-Strain, M., & Acumen LLC. (2000). *Getting to know the future customers of the Office of Child Support: Projections report for 2004 and 2009.* Washington, DC: U.S. Department of Health and Human Services, Administration for Children and Families.

Lareau, A. (2002). Invisible inequality: Social class and child rearing in black and white families. *American Sociological Review, 67* (October), 747-776.

Linver, M. R., Brooks-Gunn, J., & Kohen, D. E. (2002). Family processes as pathways from income to young children's development. *Developmental Psychology, 38*, 719-734.

Lonigan, C. J., Wagner, R. K., Torgesen, J. K., & Rashotte, C. (2002). *Preschool comprehensive test of phonological & print processing*. (Subtests: Print Awareness and Elision).

Lutz, M. N., Fantuzzo, J., & McDermott, P. (2000). *Adjustment scales for preschool intervention (ASPI).* Philadelphia, PA: University of Pennsylvania.

Lutz, M. N., Fantuzzo, J., & McDermott, P. (2002). Multidimensional assessment of emotional and behavioral adjustment problems of low-income preschool children: Development and initial validation. *Early Childhood Research Quarterly*, *17*(3), 338-355.

Martin, J. A., Hamilton, B. E., Sutton, P. D., Ventura, S. J., Menacker, F., Kirmeyer, S., & Mathews, T. J. (2009). Births: Final data for 2006. *National Vital Statistics Reports, 57*(7). U.S. Department of Health and Human Services, Centers for Disease Control and Prevention, National Center for Health Statistics.

Mason, J. M., and Stewart, J. (1989). *The CAP early childhood diagnostic instrument* (prepublication edition). American Testronics.

McCarthy, D. (1970, 1972). *McCarthy scales of children's abilities*. San Antonio, TX: The Psychological Corporation. (Subtest: Draw-a-Design Task).

McCartney, K., & Rosenthal, R. (2000). Effect size, practical importance, and social policy for children. *Child Development, 71*(1), 173-180.

McLoyd, V. C. (1990). The impact of economic hardship on black families and children: Psychological distress, parenting, and socioemotional development. *Child Development, 61*, 311-346.

Meier, J. (2003, June 20). *Kindergarten readiness study: Head Start success. Interim report.* San Bernardino, CA: Preschool Services Department of San Bernardino County.

National Council of Teachers of Mathematics. (1989). *Curriculum and evaluation standards*. Reston, VA: Author.

National Council of Teachers of Mathematics. (2000). *Principles and standards for school mathematics executive summary*. Reston, VA: Author.

National Institute of Child Health and Human Development, NIH, DHHS. (2006). *The NICHD Study of Early Child Care and Youth Development (SECCYD): Findings for Children up to Age 4 ½ Years*. Washington, DC: U.S. Government Printing Office.

NICHD Early Child Care Research Network & Duncan, G. (2003). Modeling the impacts of child care quality on children's preschool cognitive development. *Child Development, 74*(5), 1454-1475.

Onatsu-Arvilommi, T., Nurmi, J., & Aunola, K. (1998). Mothers' and fathers' well-being, parenting styles, and their children's cognitive and behavioral strategies at primary school. *European Journal of Psychology of Education, 13*, 536-543.

Phillips, D., Mekos, M., Scarr, S., McCartney, K., Abbott-Shim, M. (2001). Within and beyond the classroom door: Assessing quality in child care centers. *Early Childhood Research Quarterly, 15*, 475-496.

Pianta, R. C. (1992). *Child-parent relationship scale*. Charlottesville, VA: University of Virginia.

Pianta, R. C. (1996). *Student-teacher relationship scale*. Charlottesville, VA: University of Virginia.

Pinderhughes, E. E., Dodge, K. E., Bates, J. E., Pettis, G. S., & Zelli, A. (2000). Discipline responses influences of parents' socioeconomic status, ethnicity, beliefs about parenting, stress, and cognitive-emotional processes. *Journal of Family Psychology, 14*(3), 380-400.

Pong, S., Hao, L., & Gardner, E. (2005). The roles of parenting styles and school performance of Immigrant Asian and Hispanic adolescents. *Social Science Quarterly, 86*, 928-950.

Ramey, C. T., & Ramey, S. L. (2004). Early education interventions and intelligence. Implications for Head Start. I n E. Zigler, & S. Styfco (Eds.), *The Head Start debates* (p. 3-17). Baltimore, MD: Paul H. Brookes Publishing Co.

Reynolds, A. J., & Temple, J. A. (1995). Quasi-experimental estimates of the effects of a preschool intervention. *Evaluation Review, 19*, 347-373.

Reynolds, A., Temple, J., Ou, S-R., Robertson, D. L., Mersky, J. P., Topitzes, J. W., & Niles, M. D. (2007). Effects of a school-based, early childhood intervention on adult health and well-being: A 19-year follow up of low-income families. *Archives of Pediatric and Adolescent Medicine, 161*(8), 730-733.

Ritblatt, S. N., Brassert, S. M., Johnson, R., & Gomez, F. (2001). Are two better than one? The impact of years in Head Start on child outcomes, family environment and reading at home. *Early Childhood Research Quarterly, 16*, 525-537.

Rock, D. A., & Pollack, J. M. (2002). *Early childhood longitudinal study-kindergarten class of 1998-99 (ECLS-K): Psychometric report for kindergarten through first grade.* NCES 2002-05. Washington, DC: U.S. Department of Education: National Center for Education Statistics.

Schweinhart, L. J. (2005). *The High/Scope Perry preschool study through age 40.* Ypsilanti, MI: High/Scope Press.

Seligman, M. E. P. (1993). *What you can change…and what you can't.* New York: Ballantine Books.

Slavin, R. E. (1990). *Cooperative learning: Theory, research, and practice*. Englewood Cliffs, NJ: Prentice Hall.

Snow, K., Burns, M. S., & Griffin, P., (Eds.). (1998). *Preventing reading difficulties in young children*. Washington, DC: National Academy Press.

Strizek, G. A., Pittonsberger, T. L., Riordan, K. E., Lyter, D. M., & Orlofsky, G. F. (2006). *Characteristics of schools, districts, teachers, principals, and school libraries in the United States: 2003-04 schools and staffing survey*. Washington, DC: National Center for Education Statistics.

The Urban Institute, Health Insurance Trends. (2007). Retrieved June 10, 2008 from the Urban Institute Web Site: http://www.urban.org/toolkit/issues/healthinsurance.cfm.

University of North Carolina, Frank Porter Graham Child Development Institute (ND). *The Abecedarian project.* Retrieved April 3, 2008, from http://www.fpg.unc.edu/~abc/

U.S. Department of Education, National Center for Education Statistics. (2002). *Children's reading and mathematics achievement in kindergarten and first grade*. Washington, DC: Author.

U.S. Department of Education, National Center for Education Statistics. (2000). *America's Kindergartners*. Washington, DC: Author.

U.S. Department of Education, National Center for Education Statistics. (2007). *The condition of education*. Washington, DC: Author.

U.S. Department of Health and Human Services, Administration for Children, Youth, and Families. (1999). *Head Start program performance standards and other regulations*. Washington, DC: Author.

U.S. Department of Health and Human Services. (2000). *Healthy people 2010: 21 – Oral health*. Retrieved from: www.healthypeople.gov/data/midcourse/html/focusareas/FA21Objectives.htm.

U.S. Department of Health and Human Services, Administration for Children and Families. (2001). *National evaluation of family support programs: Final report*. Washington, DC: Author.

U.S. Department of Health and Human Services, Administration for Children & Families, Administration on Children, Youth & Families, Commissioner's Office of Research & Evaluation. (January, 2001). *Head Start FACES: Longitudinal findings on program performance. Third progress report.* Washington, DC: Author.

U.S. Department of Health and Human Services. Administration for Children, Youth and Families. (2002). *Making a difference in the lives of infants and toddlers and their families: The impact of Early Head Start*. Washington, DC: Author.

U.S. Department of Health and Human Services, Administration for Children, Youth and Families. (2003). *Head Start FACES 2000: A whole-child perspective on program performance. Fourth progress report.* Washington, DC: Author.

U.S. Department of Health and Human Services, Administration for Children and Families. (2004). Head Start Program Fact Sheet. Retrived September 14, 2009 from the U.S. Department of Health and Human Services Website: http://www.acf.hhs.gov/programs/ohs/about/fy2004.html.

U.S. Department of Health and Human Services, Administration for Children and Families. (June 2005). *Head Start Impact Study: First year findings.* Washington, DC: Author.

U.S. Department of Health and Human Services, National Institute of Child Health and Human Development, Administration on Children, Youth, and Families, and the Office of the Assistant Secretary for Planning and Evaluation. (2002). *Early childhood education and school readiness workshop: Conceptual models, constructs, and measures*. Draft proceedings, November 2002.

U.S. General Accounting Office. (April 2000). *Oral health: Dental disease is a chronic problem among low-income populations.* GAO/HEHS-00-72. Washington, DC: Author.

U.S. General Accounting Office. (1997). *Head Start: Research provides little information on impact of current program* (p. 8). Report to the Chairman, Committee on the Budget, House of Representatives. Washington, DC: Author.

U.S. General Accounting Office. (1998*). Head Start: Challenges in monitoring program quality and demonstrating results* (p. 22). Report to Congressional Requestors. Washington, DC: Author.

U.S. General Accounting Office. (2003). *Head Start: Curriculum use and individual child assessment in cognitive and language development.* Report to Congressional Requesters. Washington, DC: Author.

Whitehurst, G. J., & Lonigan, C. J. (1998). Child development and emergent literacy. *Child Development,* 68, 848-872.

Whitehurst G., & Lonigan, C. J. (2001). Emergent literacy: Development from prereaders to readers. In S. B. Neuman & D. K. Dickinson (Eds.), *Handbook of early literacy research* (pp. 11-29). New York: Guilford Press.

Woodcock, R. W., & Muñoz-Sandoval, A. F. (1996). *Bateriá Woodcock-Muñoz pruebas de aprovechamiento-Revisada*. Itasca, IL: Riverside Publishing. (Subtests: Identificación de letras y palabras, Dictado, and Problemas aplicados.)

Woodcock, R. W., McGrew, K. S., & Mather, N. (2001). *Woodcock-Johnson III tests of achievement*. Itasca, IL: Riverside Publishing. (Subtests: Letter-Word Recognition, Spelling, Oral Comprehension, and Applied Problems.)

Wolf, F. M. (1986). *Meta-analysis: Quantitative methods for research synthesis*. Newbury Park, CA: Sage.

Wong, V. C., Cook, T. D., Barnett, W. S., & Jung, K. (2008). An effectiveness-based evaluation of five state pre-kindergarten programs (p. 32). *Journal of Policy Analysis and Management*, *27*(1), 122-154.

Zill, N., Resnick, G., Kim, K., O'Donnell, K., Sorongon, A., McKey, R., Pai-Samant, S., et al. (2003). *Head Start FACES 2000: A whole-child perspective on program performance.* Report prepared for the Administration for Children and Families, U.S. Department of Health and Human Services (DHHS) under contract HHS-105-96-1912, Head Start Quality Research Consortium's Performance Measure Center.

Appendix A

Section 649(g) of the Head Start Act, 1988
(PL 105-285)

Appendix A: Section 649(g) of the Head Start Act, 1998 (PL 105-285)

(g) NATIONAL HEAD START IMPACT STUDY. --

(1) EXPERT PANEL. --

(A) IN GENERAL. --The Secretary shall appoint an independent panel consisting of experts in program evaluation and research, education, and early childhood programs--

(i) to review, and make recommendations on, the design and plan for the research (whether conducted as a single assessment or as a series of assessments) described in paragraph (2), within 1 year after the date of enactment of the Coats Human Services Reauthorization Act of 1998;

(ii) to maintain and advise the Secretary regarding the progress of the research; and

(iii) to comment, if the panel so desires, on the interim and final research reports submitted under paragraph (7).

(B) TRAVEL EXPENSES. --The members of the panel shall not receive compensation for the performance of services for the panel, but shall be allowed travel expenses, including per diem in lieu of subsistence, at rates authorized for employees of agencies under subchapter I of chapter 57 of title 5, United States Code, while away from their homes or regular places of business in the performance of services for the panel. Notwithstanding section 1342 of title 31, United States Code, the Secretary may accept the voluntary and uncompensated services of members of the panel.

(2) GENERAL AUTHORITY: After reviewing the recommendations of the expert panel, the Secretary shall make a grant to, or enter into a contract or cooperative agreement with an organization to conduct independent research that provides a national analysis of the impact of Head Start programs. The Secretary shall ensure that the organization shall have expertise in program evaluation, and research, education, and early childhood programs.

(3) DESIGNS AND TECHNIQUES. --The Secretary shall ensure that the research uses rigorous methodological designs and techniques, (based on the recommendations of the expert panel) including longitudinal designs, control groups, nationally recognized standardized measures, and random selection and assignment, as appropriate. The Secretary may provide that the research shall be conducted as a single comprehensive assessment or as a group of coordinated assessments designed to provide, when taken together, a national analysis of the impact of Head Start programs.

(4) PROGRAMS. --The Secretary shall ensure that the study focuses primarily on Head Start programs that operate in the 50 States, the Commonwealth of Puerto Rico or the District of Columbia and that do not specifically target special populations.

(5) ANALYSIS. --The Secretary shall ensure that the organization conducting the research--

(A)(i) determines if, overall, the Head Start programs have impacts consistent with their primary goal of increasing the social competence of children, by increasing the everyday effectiveness of the children in dealing with their present environments and future responsibilities, and increasing their school readiness;

(ii) considers whether the Head Start programs--

(I) enhance the growth and development of children in cognitive, emotional, and physical health areas;

(II) strengthen families as the primary nurturers of their children; and

(III) ensure that children attain school readiness; and

(iii) examines--

(I) the impact of the Head Start programs on increasing access of children to such services as educational, health, and nutritional services, and linking children and families to needed community services; and

(II) how receipt of services described in subclause (I) enriches the lives of children and families participating in Head Start programs;

(B) examines the impact of Head Start programs on participants on the date the participants leave Head Start programs, at the end of kindergarten, and at the end of first grade (whether in public or private school), by examining a variety of factors, including educational achievement, referrals for special education or remedial course work, and absenteeism;

(C) makes use of random selection from the population of all Head Start programs described in paragraph (4) in selecting programs for inclusion in the research; and

(D) includes comparisons of individuals who participate in Head Start programs with control groups (including control groups) composed of--

(i) individuals who participate in other early childhood programs (such as public or private preschool programs and day care); and

(ii) individuals who do not participate in any other early childhood program; and

(6) CONSIDERATION OF SOURCES OF VARIATION. --In designing the research, the Secretary shall, to the extent practicable, consider addressing possible sources of variation in impact of Head Start programs, including variations in impact related to such factors as—

(A) Head Start program operations;

(B) Head Start program quality;

(C) the length of time a child attends a Head Start program;

(D) the age of the child on entering the Head Start program;

(E) the type of organization (such as a local educational agency or a community action agency) providing services for the Head Start program;

(F) the number of hours and days of program operation of the Head Start program (such as whether the program is a full-working-day, full calendar year program, a part-day program, or a part-year program); and

(G) other characteristics and features of the Head Start program (such as geographic location, location in an urban or a rural service area, or participant characteristics), as appropriate.

(7) REPORTS. --

(A) SUBMISSION OF INTERIM REPORTS. --The organization shall prepare and submit to the Secretary two interim reports on the research. The first interim report shall describe the design of the research, and the rationale for the design, including a description of how potential sources of variation in impact of Head Start programs have been considered in designing the research. The second interim report shall describe the status of the study and preliminary findings of the study, as appropriate.

(B) SUBMISSION OF FINAL REPORT. --The organization shall prepare and submit to the Secretary a final report containing the findings of the research.

(C) TRANSMITTAL OF REPORTS TO CONGRESS. --

(i) IN GENERAL. --The Secretary shall transmit, to the committees described in clause (ii), the first interim report by September 30, 1999, the second interim report by September 30, 2001, and the final report by September 30, 2003.

(ii) COMMITTEES. --The committees referred to in clause (i) are the Committee on Education and the Workforce of the House of Representatives and the Committee on Labor and Human Resources of the Senate.

(8) DEFINITION. --In this subsection, the term 'impact', used with respect to a Head Start program, means a difference in an outcome for a participant in a program that would not have occurred without the participation in the program.

Appendix B

Advisory Committee on Head Start Research and Evaluation

Advisory Committee on Head Start Research and Evaluation

Membership*

Wade Horn: CHAIR
Assistant Secretary of ACF
U.S. Department of Health and Human Services
Administration for Children and Families

Mark Appelbaum
Professor
University of California – San Diego
Department of Psychology

Margaret Burchinal
Research Associate Professor
University of North Carolina, Chapel Hill
Frank Porter Graham Child Development Institute

Gayle Cunningham
Executive Director
Jefferson County Committee for Economic Opportunity

Naomi Goldstein: Designated Federal Official
Director, Office of Planning, Research and Evaluation
U.S. Department of Health and Human Services
Administration for Children and Families

Barbara Broman
Acting Deputy Assistant Secretary for Human Services Policy
U.S. Department of Health and Human Services
Office of the Assistant Secretary for Planning and Evaluation

Nell Carvell
Director of LEAP and Head Start Initiatives
Southern Methodist University
Learning Therapist Program

Frank Fuentes
Acting Associate Commissioner for the Head Start Bureau
U.S. Department of Health and Human Services
Administration for Children and Families

Sarah Greene
President and Chief Executive Officer
National Head Start Association

Mary Louise Hemmeter
Associate Professor
Vanderbilt University
Department of Special Education

Marci Kanstoroom
Executive Editor, Education Next
Research Fellow, Hoover Institution

Joan Ohl
Commissioner of ACYF
U.S. Department of Health and Human Services
Administration on Children, Youth and Families

Craig Ramey
Professor and Co-Director
Georgetown School of Nursing and Health Studies
Center for Health Equity, Research, Implementation, and Training (CHERITH)

Ron Haskins
Senior Fellow
The Brookings Institute

Nicholas Ialongo
Associate Professor
Johns Hopkins University
Bloomberg School of Public Health
Department of Mental Hygiene

Kristin Moore
President
Child Trends

David Olds
Director
University of Colorado Health Sciences Center
Prevention Research Center for Family and Child Health

William Rosenberg
Director
Drexel University Survey Research Center

George Smith
Executive Director and CEO
Management Planning Institute, Inc.

William Douglas Tynan
Staff Psychologist
Nemours – AI DuPont Hospital for Children
Behavior Health

Edward Zigler
Sterling Professor of Psychology, Emeritus
Yale University
Center in Child Development and Social Policy

Dorothy Strickland
Samuel DeWitt Proctor Professor of Education
Rutgers University
Graduate School of Education

Betty Ward-Zukerman
Assistant Director
Education, Workforce, and Income Security Team
U.S. Government Accountability Office

*Affiliation of the members reflects their position at the time of appointment to the Advisory Committee.

Appendix C

Language Decision Form

Appendix C: Language Decision Form

To the best of your knowledge,

1. What language does the child speak most often at home?

ENGLISH 01
SPANISH 02
OTHER (SPECIFY)..... 03
_____

2. What language does the child speak most often at this child care setting?

ENGLISH 01
SPANISH 02
OTHER (SPECIFY)..... 03
_____

3. What language does it appear this child prefers to speak?

ENGLISH 01
SPANISH 02
OTHER (SPECIFY)..... 03
_____

Language in which at least two of three responses are the same:

_____
LANGUAGE

4. If language is other than English or Spanish, ask main care provider: Can child understand and answer questions in English? (IF YES, PROCEED WITH ENGLISH TESTING. OTHERWISE FOLLOW INSTRUCTIONS FOR CHILDREN BEING TESTED IN OTHER LANGUAGE)

YES..... 1
NO..... 2

5. Language child will be tested in:

_____
LANGUAGE

Appendix D

Back-up Tables for Chapter 3

Exhibit D.1: Percentage of Children by Their Settings' ECERS-R Total and Subscale Scores, and Mean Scores, 4-Year-Old Cohort in Head Start or Center-Based Care, 2003

ECERS-R Scale	1	2	3	4	5	6	7	Missing	Total	Mean
Total:										
Head Start Group	0.0	0.8	4.2	14.6	34.9	40.5	4.5	0.5	100%	5.25
Control Group	0.1	3.3	10.6	19.1	38.4	25.0	2.5	1.0	100%	4.75
Space and Furnishings:										
Head Start Group	0.0	1.0	2.9	16.1	32.9	37.3	9.3	0.5	100%	5.25
Control Group	0.0	2.1	8.0	25.8	29.4	28.0	5.9	0.8	100%	4.83
Personal Care Routines:										
Head Start Group	0.1	3.7	5.9	11.6	19.2	29.1	30.0	0.5	100%	5.50
Control Group	2.7	3.8	12.3	12.2	14.2	32.4	21.6	0.8	100%	5.12
Language-Reasoning:										
Head Start Group	0.1	2.0	5.8	16.6	28.1	28.2	18.8	0.5	100%	5.19
Control Group	2.2	3.1	17.1	13.3	27.7	25.1	10.7	0.8	100%	4.69
Activities:										
Head Start Group	0.1	3.3	11.1	23.8	37.9	19.5	3.9	0.5	100%	4.67
Control Group	2.4	10.7	18.4	25.8	20.5	18.6	2.8	0.8	100%	4.11
Interaction:										
Head Start Group	0.4	1.8	2.2	6.1	14.6	43.1	31.3	0.5	100%	5.83
Control Group	1.0	4.1	3.4	7.6	27.7	31.7	23.5	1.0	100%	5.41
Program Structure:										
Head Start Group	1.5	1.8	4.3	8.9	12.9	25.7	44.6	0.5	100%	5.80
Control Group	3.0	6.6	9.1	16.5	24.8	10.6	28.4	1.0	100%	5.00

Exhibit D.2: Percentage of Children by Their Settings' FDCRS Total and Subscale Scores, and Mean Scores, 4-Year-Old Cohort in Child Care Homes, 2003

FDCRS Scale	1	2	3	4	5	6	7	Missing	Total	Mean
Total:										
Head Start Group	0.0	0.0	21.0	0.0	79.0	0.0	0.0	0.0	100%	4.67
Control Group	2.9	13.6	64.2	2.0	17.3	0.0	0.0	0.0	100%	4.34
Space and Furnishings:										
Head Start Group	0.0	0.0	21.0	79.0	0.0	0.0	0.0	0.0	100%	4.00
Control Group	2.9	24.2	53.7	14.5	4.8	0.0	0.0	0.0	100%	4.21
Basic Care:										
Head Start Group	0.0	0.0	0.0	21.0	79.0	0.0	0.0	0.0	100%	4.19
Control Group	7.9	9.3	0.0	61.7	21.0	0.0	0.0	0.0	100%	4.53
Language and Reasoning:										
Head Start Group	0.0	0.0	21.0	0.0	0.0	79.0	0.0	0.0	100%	4.65
Control Group	2.9	24.2	44.4	11.3	2.0	15.3	0.0	0.0	100%	4.28
Learning Activities:										
Head Start Group	0.0	21.0	0.0	0.0	79.0	0.0	0.0	0.0	100%	4.94
Control Group	16.6	19.4	14.1	46.0	2.0	2.0	0.0	0.0	100%	4.23
Social Development:										
Head Start Group	0.0	0.0	0.0	0.0	21.0	79.0	0.0	0.0	100%	4.17
Control Group	2.9	0.0	9.3	55.0	26.0	6.8	0.0	0.0	100%	4.79

Exhibit D.3: Percentage of Children by Their Settings' Arnett Total and Subscale Scores, and Mean Scores, 4-Year-Old Cohort in Out-of-Own-Home Care, 2003

Arnett Scale	1	2	3	4	Missing*	Total	Mean
Total:							
Head Start Group	0.0	1.3	30.6	67.4	0.7	100%	3.56
Control Group	0.0	5.0	34.5	59.8	0.6	100%	3.46
Sensitivity:							
Head Start Group	0.1	6.3	44.5	45.7	3.5	100%	3.41
Control Group	1.2	13.0	42.9	42.6	0.4	100%	3.27
Harshness:							
Head Start Group	0.0	0.5	8.8	89.0	1.7	100%	3.90
Control Group	0.0	2.2	8.9	88.0	1.0	100%	3.87
Detachment:							
Head Start Group	0.0	0.7	9.2	88.7	1.5	100%	3.89
Control Group	0.1	1.4	15.4	82.2	1.0	100%	3.81
Permissiveness:							
Head Start Group	0.2	1.0	23.1	74.2	1.5	100%	3.74
Control Group	0.0	2.2	21.3	75.8	0.7	100%	3.74
Independence:							
Head Start Group	0.0	4.8	49.3	43.3	2.6	100%	3.40
Control Group	0.1	7.2	60.8	30.9	1.0	100%	3.24

Exhibit D.4: Percentage of Children by Their Settings' ECERS-R Total and Subscale Scores, and Mean Scores, 3-Year-Old Cohort in Head Start or Center-Based Care, 2003

ECERS-R Scale	1	2	3	4	5	6	7	Missing	Total	Mean
Total:										
Head Start Group	0.0	1.4	4.3	18.9	31.9	37.1	5.9	0.7	100%	5.17
Control Group	0.0	4.1	13.0	21.7	33.9	23.3	2.7	1.3	100%	4.70
Space and Furnishings:										
Head Start Group	0.2	1.0	4.7	17.1	32.5	35.1	8.8	0.6	100%	5.16
Control Group	0.0	0.9	9.6	24.1	34.6	23.4	6.5	1.0	100%	4.83
Personal Care Routines:										
Head Start Group	0.4	5.2	4.8	10.0	20.6	30.7	27.5	1.0	100%	5.46
Control Group	0.2	7.2	12.9	16.3	19.5	23.6	19.1	1.2	100%	4.96
Language-Reasoning:										
Head Start Group	0.2	2.3	6.7	18.6	25.8	25.0	20.8	0.6	100%	5.15
Control Group	1.0	7.0	16.7	22.7	16.4	19.9	15.3	1.0	100%	4.59
Activities:										
Head Start Group	0.2	3.0	8.9	27.0	38.5	16.0	5.7	0.6	100%	4.68
Control Group	2.8	9.5	19.3	28.7	24.3	11.6	2.8	1.0	100%	4.03
Interaction:										
Head Start Group	1.0	4.3	3.5	7.6	14.9	39.4	28.6	0.7	100%	5.64
Control Group	1.1	5.4	8.7	10.6	14.6	34.6	24.0	1.3	100%	5.37
Program Structure:										
Head Start Group	3.0	3.6	5.5	9.6	14.5	24.7	38.6	0.7	100%	5.54
Control Group	5.6	2.3	8.3	14.1	19.3	22.0	27.0	1.3	100%	5.11

Exhibit D.5: Percentage of Children by Their Settings' FDCRS Total and Subscale Scores, and Mean Scores, 3-Year-Old Cohort in Child Care Homes, 2003

FDCRS Scale	1	2	3	4	5	6	7	Missing	Total	Mean
Total:										
Head Start Group	0.0	20.7	0.0	52.6	15.6	11.2	0.0	0.0	100%	3.72
Control Group	5.4	14.4	14.4	39.7	17.7	8.6	0.0	0.0	100%	3.71
Space and Furnishings:										
Head Start Group	0.0	20.7	0.0	52.6	15.6	11.2	0.0	0.0	100%	3.41
Control Group	10.7	5.5	19.5	38.0	20.4	5.8	0.0	0.0	100%	3.44
Basic Care:										
Head Start Group	0.0	20.7	0.0	52.6	0.0	15.6	11.2	0.0	100%	3.40
Control Group	3.5	2.3	15.5	12.7	34.4	18.8	0.0	12.8	100%	4.17
Language and Reasoning:										
Head Start Group	7.2	0.0	13.4	46.3	21.8	11.2	0.0	0.0	100%	3.99
Control Group	7.7	8.6	28.0	29.4	10.2	10.5	5.6	0.0	100%	3.76
Learning Activities:										
Head Start Group	7.2	13.4	6.2	46.3	15.6	11.2	0.0	0.0	100%	3.76
Control Group	7.7	19.0	23.5	9.1	35.9	4.8	0.0	0.0	100%	3.57
Social Development:										
Head Start Group	0.0	0.0	13.4	7.2	52.6	15.6	11.2	0.0	100%	4.60
Control Group	0.0	3.5	3.2	37.3	24.6	20.7	0.0	10.7	100%	4.62

Exhibit D.6: Percentage of Children by Their Settings' Arnett Total and Subscale Scores, and Mean Scores, 3-Year-Old Cohort in Child Care Homes, 2003

Arnett Scale	1	2	3	4	Missing	Total	Mean
Total:							
Head Start Group	0.0	2.5	33.5	63.2	0.8	100%	3.52
Control Group	0.0	6.6	46.1	46.5	0.8	100%	3.39
Sensitivity:							
Head Start Group	0.4	7.9	42.0	46.2	3.5	100%	3.39
Control Group	1.8	9.5	52.0	32.2	4.5	100%	3.20
Harshness:							
Head Start Group	0.0	1.9	7.9	87.3	2.9	100%	3.88
Control Group	0.0	2.3	11.7	84.8	1.2	100%	3.84
Detachment:							
Head Start Group	0.0	1.6	12.8	84.8	0.8	100%	3.84
Control Group	0.0	2.9	16.6	79.2	1.3	100%	3.77
Permissiveness:							
Head Start Group	0.4	2.8	23.7	70.6	2.5	100%	3.69
Control Group	0.5	5.2	24.1	67.3	2.9	100%	3.63
Independence:							
Head Start Group	0.6	8.5	44.6	42.4	3.9	100%	3.34
Control Group	2.0	17.6	53.5	24.8	2.2	100%	3.03

Exhibit D.7: Percentage of Children by Their Settings' ECERS-R Total and Subscale Scores, and Mean Scores, 3-Year-Old Cohort in Head Start or Center-Based Care, 2004

ECERS-R Scale	1	2	3	4	5	6	7	Missing	Total	Mean
Total:										
Head Start Group	0.0	0.8	7.7	22.1	38.2	26.1	4.8	0.3	100%	4.96
Control Group	0.0	1.2	7.8	23.1	35.9	27.9	3.6	0.5	100%	4.93
Space and Furnishings:										
Head Start Group	0.1	0.3	7.6	20.4	38.0	27.9	5.8	0.0	100%	4.98
Control Group	0.2	0.7	4.7	24.0	33.5	29.3	7.5	0.1	100%	5.00
Personal Care Routines:										
Head Start Group	1.6	8.9	16.5	17.0	17.8	22.9	15.1	0.1	100%	4.66
Control Group	0.9	9.9	21.3	14.9	19.4	20.7	12.4	0.5	100%	4.53
Language-Reasoning:										
Head Start Group	0.5	3.5	7.3	19.5	25.3	25.6	18.2	0.1	100%	5.04
Control Group	0.5	1.9	6.6	20.5	25.7	24.0	20.4	0.5	100%	5.11
Activities:										
Head Start Group	0.3	4.3	12.2	34.1	29.8	12.8	6.2	0.1	100%	4.47
Control Group	0.9	3.4	14.2	33.2	25.9	16.7	5.3	0.5	100%	4.45
Interaction:										
Head Start Group	0.4	1.7	3.9	8.3	20.8	33.2	31.1	0.5	100%	5.70
Control Group	2.2	1.4	3.5	6.0	24.5	28.8	33.1	0.5	100%	5.68
Program Structure:										
Head Start Group	2.2	2.8	6.4	8.3	19.3	19.4	41.4	0.1	100%	5.57
Control Group	2.7	2.7	6.5	15.5	15.0	22.1	35.1	0.6	100%	5.40

Exhibit D.8: Percentage of Children by Their Settings' FDCRS Total and Subscale Scores, and Mean Scores, 3-Year-Old Cohort in Child Care Homes, 2004

FDCRS Scale	1	2	3	4	5	6	7	Missing	Total	Mean
Total:										
Head Start Group	0.0	28.5	21.2	0.0	50.3	0.0	0.0	0.0	100%	3.75
Control Group	0.0	29.5	0.0	62.5	0.0	0.0	0.0	8.0	100%	3.26
Space and Furnishings:										
Head Start Group	0.0	28.5	21.2	36.7	0.0	13.6	0.0	0.0	100%	3.59
Control Group	0.0	57.3	0.0	0.0	34.7	0.0	0.0	8.0	100%	3.19
Basic Care:										
Head Start Group	0.0	0.0	49.7	13.6	12.0	24.7	0.0	0.0	100%	4.06
Control Group	0.0	29.5	38.4	24.1	0.0	0.0	0.0	8.0	100%	2.92
Language and Reasoning:										
Head Start Group	0.0	0.0	49.7	13.6	24.7	12.0	0.0	0.0	100%	3.83
Control Group	29.5	0.0	34.7	0.0	0.0	27.8	0.0	8.0	100%	3.29
Learning Activities:										
Head Start Group	28.5	0.0	21.2	24.7	13.6	12.0	0.0	0.0	100%	3.38
Control Group	29.5	0.0	0.0	62.5	0.0	0.0	0.0	8.0	100%	3.25
Social Development:										
Head Start Group	0.0	0.0	28.5	50.3	0.0	21.2	0.0	0.0	100%	4.22
Control Group	0.0	29.5	0.0	0.0	62.5	0.0	0.0	8.0	100%	3.95

Exhibit D.9: Percentage of Children by Their Settings' Arnett Total and Subscale Scores, and Mean Scores, 3-Year-Old Cohort in Out-of-Own-Home Care, 2004

Arnett Scale	1	2	3	4	Missing	Total	Mean
Total:							
Head Start Group	0.2	1.3	37.3	60.5	0.6	100%	3.59
Control Group	0.0	2.7	32.2	63.7	1.3	100%	3.62
Sensitivity:							
Head Start Group	0.7	11.9	43.8	42.2	1.4	100%	3.29
Control Group	0.2	12.3	38.9	47.0	1.7	100%	3.35
Harshness:							
Head Start Group	0.2	0.4	8.7	89.9	0.8	100%	3.90
Control Group	0.0	1.6	8.0	88.5	1.9	100%	3.89
Detachment:							
Head Start Group	0.0	1.7	11.4	85.7	1.2	100%	3.85
Control Group	0.3	1.8	12.4	84.0	1.4	100%	3.83
Permissiveness:							
Head Start Group	0.2	1.5	26.3	70.9	1.1	100%	3.70
Control Group	0.0	2.1	24.7	71.9	1.3	100%	3.71
Independence:							
Head Start Group	0.0	7.4	51.4	40.2	0.9	100%	3.33
Control Group	0.0	7.3	50.7	39.8	2.3	100%	3.33

Appendix E

IOT Tables

Exhibit E4.2: Estimated IOT Impacts on Cognitive Outcomes, by Year: 4-Year-Old Cohort

Outcome	ITT Regression Adjusted Impact	IOT		
	Impact	Impact	p-value	Effect Size
Head Start Year (Spring 2003)				
Language and Literacy Measures				
PPVT (Adapted)	3.55	**5.31****	0.028	0.13
WJ-III Letter-Word Identification	5.98	**8.94****	0.017	0.34
WJ-III Spelling	3.77	**5.64****	0.029	0.22
WJ-III Oral Comprehension	-0.94	-1.41	0.395	-0.08
CTOPPP Elision	2.45	3.66	0.444	0.07
Color Identification	0.08	**0.12*****	0.010	0.25
Letter Naming	2.36	**3.53*****	0.002	0.37
WJ-III Pre-Academic Skills	4.23	**6.32****	0.022	0.29
Spanish Language and Literacy Measures				
TVIP (Adapted)	9.04	10.65	0.106	0.25
WM Letter-Word Identification	1.91	2.25	0.180	0.16
Math Skills Measures				
WJ-III Applied Problems	3.18	4.75	0.139	0.18
Counting Bears	0.04	0.06	0.181	0.12
Fine Motor Skills Measures				
McCarthy Draw-a-Design	0.20	0.30	0.110	0.15
Parent-Reported Literacy Measure				
Emergent Literacy Scale	0.43	**0.64*****	0.000	0.46
Kindergarten Year (Spring 2004)				
Language and Literacy Measures				
PPVT (Adapted)	1.78	2.66	0.328	0.06
WJ-III Letter-Word Identification	-0.19	-0.28	0.918	-0.01
WJ-III Spelling	-0.52	-0.78	0.764	-0.03
WJ-III Oral Comprehension	-0.91	-1.36	0.327	-0.08
CTOPPP Elision	-2.85	-4.26	0.374	-0.09
Letter Naming	0.40	0.60	0.274	0.09
WJ-III Pre-Academic Skills	-0.47	-0.70	0.745	-0.03
WJ-III Word Attack	-1.13	-1.69	0.639	-0.05
WJ-III Basic Reading Skills	-0.71	-1.06	0.728	-0.03
Spanish Language and Literacy Measures				
TVIP (Adapted)	-1.03	-1.21	0.868	-0.02
WM Letter-Word Identification	-4.28	-5.04	0.130	-0.19

Exhibit E4.2: Estimated IOT Impacts on Cognitive Outcomes, by Year: 4-Year-Old Cohort (continued)

Outcome	ITT Regression Adjusted Impact	IOT		
	Impact	Impact	p-value	Effect Size
Math Skills Measures				
WJ-III Applied Problems	0.12	0.18	0.936	0.01
WJ-III Quantitative Concepts	-0.13	-0.19	0.920	-0.01
WJ-III Math Reasoning	-0.08	-0.12	0.951	-0.01
School Performance Assessment Measures				
School Accomplishments	0.00	0.00	0.997	0.00
Promotion	0.00	0.00	0.888	0.00
Language and Literacy Ability	0.04	0.06	0.424	0.13
Math Ability	0.05	0.07	0.191	0.17
Social Studies and Science Ability	0.03	0.07	0.501	0.19
1st Grade Year (Spring 2005)				
Language and Literacy Measures				
PPVT (Adapted)	2.95	4.41*	0.072	0.14
WJ-III Letter-Word Identification	0.83	1.24	0.705	0.03
WJ-III Spelling	1.55	2.32	0.347	0.09
WJ-III Oral Comprehension	0.34	0.51	0.717	0.03
WJ-III Pre-Academic Skills	0.95	1.42	0.510	0.06
WJ-III Word Attack	1.71	2.56	0.324	0.08
WJ-III Basic Reading Skills	1.08	1.61	0.550	0.05
WJ-III Academic Applications	0.38	0.57	0.730	0.03
WJ-III Academic Skills	1.11	1.66	0.446	0.07
WJ-III Passage Comprehension	0.17	0.25	0.922	0.01
WJ-III Writing Sample	0.15	0.22	0.824	0.02
Spanish Language and Literacy Measures				
TVIP (Adapted)	5.25	6.18	0.240	0.15
WM Letter-Word Identification	-4.30	-5.06	0.397	-0.11
Math Skills Measures				
WJ-III Applied Problems	0.82	1.23	0.523	0.06
WJ-III Quantitative Concepts	0.32	0.48	0.819	0.03
WJ-III Math Reasoning	0.47	0.70	0.705	0.04
WJ-III Calculation	1.41	2.11	0.255	0.11

Exhibit E4.2: Estimated IOT Impacts on Cognitive Outcomes, by Year: 4-Year-Old Cohort (continued)

	ITT Regression Adjusted Impact	IOT		
Outcome	**Impact**	**Impact**	**p-value**	**Effect Size**
School Performance Assessment Measures				
School Accomplishments	-0.59	-0.88	0.500	-0.09
Promotion	0.01	0.01	0.376	0.06
Language and Literacy Ability	-0.02	-0.03	0.433	-0.07
Math Ability	-0.05	-0.07	0.148	-0.19
Social Studies and Science Ability	-0.02	-0.03	0.362	-0.08

Notes: The four-year old cohort compliance rate is 0.66897.
Sample sizes, standard errors, and confidence intervals for all estimates can be found on the Administration for Children and Families, Office of Planning, Research and Evaluation Website at: http://www.acf.hhs.gov/programs/opre/hs/impact_study/index.html.

Key:

*** p<0.01
** p<0.05
* p<0.10

Bold IOT impact indicates the outcome passes the Benjamini-Hochberg test for multiple comparisons with a 10-percent false discovery rate.

Exhibit E4.5: Estimated IOT Impacts on Cognitive Outcomes, by Year: 3-Year-Old Cohort

Outcome	ITT Regression Adjusted Impact	IOT		
	Impact	Impact	p-value	Effect Size
Head Start Year (Spring 2003)				
Language and Literacy Measures				
PPVT (Adapted)	6.53	**9.36*** **	0.000	0.25
WJ-III Letter-Word Identification	6.14	**8.80*** **	0.000	0.37
WJ-III Spelling	2.28	3.05	0.130	0.13
WJ-III Oral Comprehension	0.28	0.40	0.698	0.03
CTOPPP Elision	5.01	**7.18***	0.061	0.14
Color Identification	0.04	0.06	0.179	0.12
Letter Naming	1.56	**2.24*** **	0.005	0.34
WJ-III Pre-Academic Skills	4.25	**6.09*** **	0.004	0.31
Spanish Language and Literacy Measures				
TVIP (Adapted)	5.21	6.09	0.365	0.15
WM Letter-Word Identification	1.59	1.86	0.380	0.15
Pre-Writing Measure				
McCarthy Draw-a-Design	0.16	**0.22*** **	0.007	0.19
Parent-Reported Literacy Measure				
Emergent Literacy Scale	0.48	**0.69*** **	0.000	0.50
Math Skills Measures				
WJ-III Applied Problems	4.35	**6.24****	0.012	0.21
Counting Bears	0.03	0.04	0.241	0.10
Age 4 Year (Spring 2004)				
Language and Literacy Measures				
PPVT (Adapted)	2.03	2.91	0.251	0.07
WJ-III Letter-Word Identification	2.56	3.67	0.112	0.13
WJ-III Spelling	0.28	0.40	0.875	0.02
WJ-III Oral Comprehension	0.25	0.36	0.743	0.02
CTOPPP Elision	8.26	**11.84*** **	0.002	0.21
Color Identification	0.01	0.01	0.466	0.04
Letter Naming	0.85	1.22	0.155	0.13
WJ-III Pre-Academic Skills	1.24	1.78	0.378	0.08
Spanish Language and Literacy Measures				
TVIP (Adapted)	-1.33	-1.55	0.803	-0.03
WM Letter-Word Identification	3.05	3.56	0.334	0.19
Pre-Writing Measure				
McCarthy Draw-a-Design	-0.09	-0.13	0.482	-0.06

Exhibit E4.5: Estimated IOT Impacts on Cognitive Outcomes, by Year: 3-Year-Old Cohort (continued)

Outcome	ITT Regression Adjusted Impact	IOT		
	Impact	Impact	p-value	Effect Size
Parent-Reported Literacy Measure				
Emergent Literacy Scale	0.20	**0.29*** **	0.002	0.22
Math Skills Measures				
WJ-III Applied Problems	0.75	1.08	0.551	0.05
Counting Bears	0.01	0.01	0.777	0.03
Kindergarten Year (Spring 2005)				
Language and Literacy Measures				
PPVT (Adapted)	0.26	0.37	0.851	0.01
WJ-III Letter-Word Identification	0.24	0.34	0.899	0.01
WJ-III Spelling	0.45	0.65	0.774	0.03
WJ-III Oral Comprehension	0.50	0.72	0.633	0.04
CTOPPP Elision	-3.52	-5.05	0.241	-0.11
Letter Naming	-0.32	-0.46	0.340	-0.08
WJ-III Pre-Academic Skills	-0.02	-0.03	0.988	-0.00
WJ-III Word Attack	-1.37	-1.96	0.563	-0.06
WJ-III Basic Reading Skills Composite	-0.54	-0.77	0.801	-0.03
Spanish Language and Literacy Measures				
TVIP (Adapted)	-7.51	-8.78	0.117	-0.22
WM Letter-Word Identification	8.73*	10.20*	0.053	0.30
Math Skills Measures				
WJ-III Applied Problems	-0.94	-1.35	0.519	-0.06
WJ-III Quantitative Concepts	-0.88	-1.26	0.310	-0.08
WJ-III Math Reasoning	-0.91	-1.30	0.408	-0.08
School Performance Assessment Measures				
School Accomplishments	-0.65	-0.93	0.203	-0.13
Promotion	-0.01	-0.01	0.709	-0.05
Language and Literacy Ability	-0.04	-0.06	0.127	-0.14
Math Ability	-0.07	**-0.10*** **	0.003	-0.29
Social Studies and Science Ability	-0.03	-0.04	0.121	-0.13
1st Grade Year (Spring 2006)				
Language and Literacy Measures				
PPVT (Adapted)	2.32	3.33	0.151	0.11
WJ-III Letter-Word Identification	0.37	0.53	0.848	0.02
WJ-III Spelling	-1.20	-1.72	0.438	-0.07
WJ-III Oral Comprehension	1.35	1.94*	0.051	0.12

Exhibit E4.5: Estimated IOT Impacts on Cognitive Outcomes, by Year: 3-Year-Old Cohort (continued)

Outcome	ITT Regression Adjusted Impact	IOT		
	Impact	Impact	p-value	Effect Size
Language and Literacy Measures (cont'd)				
WJ-III Pre-Academic Skills	0.24	0.34	0.869	0.01
WJ-III Word Attack	-0.60	-0.86	0.759	-0.03
WJ-III Basic Reading Skills	-0.08	-0.11	0.966	-0.00
WJ-III Academic Applications Composite	0.73	1.05	0.489	0.06
WJ-III Academic Skills Composite	-0.60	-0.86	0.633	-0.04
WJ-III Passage Comprehension	0.76	1.09	0.580	0.05
WJ-III Writing Sample+	-0.09	-0.13	0.928	-0.01
Spanish Language and Literacy Measures				
TVIP (Adapted)	0.04	0.05	0.993	0.00
WM Letter-Word Identification	-0.54	-0.63	0.910	-0.01
Math Skills Measures				
WJ-III Applied Problems	1.58	2.27	0.163	0.11
WJ-III Quantitative Concepts	0.78	1.12	0.450	0.06
WJ-III Math Reasoning	1.20	1.72	0.231	0.10
WJ-III Calculations	-0.03	-0.04	0.977	-0.00
School Performance Assessment Measures				
School Accomplishments	-0.29	-0.42	0.705	-0.04
Promotion	-0.02	-0.03	0.248	-0.12
Language and Literacy Ability	0.00	0.00	0.888	0.00
Math Ability	-0.02	-0.03	0.448	-0.07
Social Studies and Science Ability	-0.03	-0.04	0.286	-0.12

Notes: The three-year old cohort compliance rate is 0.69739.
Sample sizes, standard errors, and confidence intervals for all estimates can be found on the Administration for Children and Families, Office of Planning, Research and Evaluation Website at: http://www.acf.hhs.gov/programs/opre/hs/impact_study/index.html.

Key:

*** p<0.01
** p<0.05
* p<0.10

+ Indicates the reliability (Cronbach alpha) for the measure is <0.6.

Bold IOT impact indicates the outcome passes the Benjamini-Hochberg test for multiple comparisons with a 10-percent false discovery rate.

Exhibit E5.1: Estimated IOT Impacts on Social-Emotional Outcomes, by Year: 4-Year-Old Cohort

Outcome	ITT Regression Adjusted Impact	IOT		
	Impact	Impact	p-value	Effect Size
Head Start Year (Spring 2003)				
Parent-Reported Measures				
Aggressive Behavior+	-0.16	-0.24	0.164	-0.15
Hyperactive Behavior+	-0.09	-0.13	0.324	-0.09
Withdrawn Behavior+	-0.04	-0.06	0.575	-0.07
Total Problem Behavior	-0.27	-0.40	0.289	-0.12
Social Competencies+	-0.04	-0.06	0.566	-0.05
Social Skills and Positive Approaches to Learning	-0.06	-0.09	0.682	-0.05
Closeness	0.25	0.37	0.146	0.14
Conflict	-0.23	-0.34	0.698	-0.05
Positive Relationships+	0.56	0.84	0.419	0.11
Kindergarten Year (Spring 2004)				
Parent-Reported Measures				
Aggressive Behavior+	-0.08	-0.12	0.477	-0.08
Hyperactive Behavior	0.11	0.16	0.273	0.11
Withdrawn Behavior+	0.00	0.00	0.986	0.00
Total Problem Behavior	0.09	0.13	0.710	0.04
Social Competencies+	-0.03	-0.05	0.770	-0.04
Social Skills and Positive Approaches to Learning+	0.07	0.10	0.483	0.07
Closeness	-0.06	-0.09	0.793	-0.03
Conflict	-0.13	-0.19	0.788	-0.03
Positive Relationships+	0.03	0.04	0.956	0.01
Teacher-Reported Measures				
ASPI – Aggressive	-0.09	-0.13	0.893	-0.02
ASPI – Inattentive/Hyperactive	-0.69	-1.03	0.286	-0.12
ASPI – Low Energy	0.10	0.15	0.888	0.02
ASPI – Oppositional	0.13	0.19	0.819	0.03
ASPI – Peer Interactions	-0.89	-1.33	0.410	-0.12
ASPI – Shy/Socially Reticent	0.64	0.96	0.418	0.13
ASPI – Structured Learning	-0.67	-1.00	0.410	-0.10
ASPI – Teacher Interaction	0.20	0.30	0.811	0.03
Closeness	0.26	0.39	0.557	0.09
Conflict	-0.35	-0.52	0.558	-0.09
Positive Relationships	0.63	0.94	0.445	0.11

Exhibit E5.1: Estimated IOT Impacts on Social-Emotional Outcomes, by Year: 4-Year-Old Cohort (continued)

Outcome	ITT Regression Adjusted Impact	IOT		
	Impact	Impact	p-value	Effect Size
1st Grade Year (Spring 2005)				
Parent-Reported Measures				
Aggressive Behavior	-0.09	-0.13	0.483	-0.08
Hyperactive Behavior	0.00	0.00	0.972	0.00
Withdrawn Behavior+	-0.13	-0.19*	0.077	-0.19
Total Problem Behavior	-0.19	-0.28	0.453	-0.07
Social Competencies+	-0.02	-0.03	0.753	-0.03
Social Skills and Positive Approaches to Learning	0.02	0.03	0.764	0.02
Closeness	-0.01	-0.01	0.944	-0.01
Conflict	-0.50	-0.75	0.373	-0.11
Positive Relationships+	0.41	0.61	0.507	0.08
Teacher-Reported Measures				
ASPI – Aggressive	-0.72	-1.08	0.257	-0.14
ASPI – Inattentive/Hyperactive	-0.26	-0.39	0.731	-0.05
ASPI – Low Energy	0.75	1.12	0.169	0.16
ASPI – Oppositional	-0.36	-0.54	0.637	-0.07
ASPI – Peer Interactions	-0.38	-0.57	0.630	-0.05
ASPI – Shy/Socially Reticent	1.37	2.05**	0.019	0.28
ASPI – Structured Learning	0.74	1.11	0.306	0.10
ASPI – Teacher Interaction	1.29	1.93*	0.099	0.19
Closeness	0.22	0.33	0.465	0.07
Conflict	0.09	0.13	0.838	0.02
Positive Relationships	0.20	0.03	0.728	0.00

Notes: The four-year old cohort compliance rate is 0.66897.
Sample sizes, standard errors, and confidence intervals for all estimates can be found on the Administration for Children and Families, Office of Planning, Research and Evaluation Website at: http://www.acf.hhs.gov/programs/opre/hs/impact_study/index.html.

Key:

*** p<0.01
** p<0.05
* p<0.10

+ Indicates the reliability (Cronbach alpha) for the measure is <0.6.

Bold IOT impact indicates the outcome passes the Benjamini-Hochberg test for multiple comparisons with a 10-percent false discovery rate.

Exhibit E5.2: Estimated IOT Impacts on Social-Emotional Outcomes, by Year: 3-Year-Old Cohort

Outcome	ITT Regression Adjusted Impact	IOT		
	Impact	Impact	p-value	Effect Size
Head Start Year (Spring 2003)				
Parent-Reported Measures				
Aggressive Behavior	-0.10	-0.14	0.274	-0.08
Hyperactive Behavior	-0.33	**-0.47*** **	0.001	-0.30
Withdrawn Behavior+	-0.04	-0.06	0.510	-0.06
Total Problem Behavior	-0.52	**-0.75*** **	0.003	-0.20
Social Competencies+	-0.03	0.16	0.637	0.12
Social Skills and Positive Approaches to Learning	0.04	0.36	0.745	0.20
Closeness	0.18	0.19	0.220	0.07
Conflict	-0.05	-0.09	0.893	-0.01
Positive Relationships+	0.28	0.40	0.517	0.05
Age 4 Year (Spring 2004)				
Parent-Reported Measures				
Aggressive Behavior	-0.12	-0.17	0.203	-0.10
Hyperactive Behavior+	-0.13	-0.19	0.242	-0.13
Withdrawn Behavior+	-0.08	-0.11	0.248	-0.12
Total Problem Behavior	-0.39	-0.56	0.115	-0.15
Social Competencies	-0.01	0.16	0.868	0.12
Social Skills and Positive Approaches to Learning	0.19	0.27*	0.055	0.16
Closeness	0.22	0.32	0.232	0.12
Conflict	-0.39	-0.56	0.461	-0.08
Positive Relationships+	0.72	1.03	0.261	0.13
Kindergarten Year (Spring 2005)				
Parent-Reported Measures				
Aggressive Behavior	-0.08	-0.11	0.382	-0.06
Hyperactive Behavior	-0.18	-0.26**	0.048	-0.17
Withdrawn Behavior+	-0.03	-0.04	0.563	-0.04
Total Problem Behavior	-0.26	-0.37	0.246	-0.10
Social Competencies+	0.11	0.16	0.179	0.12
Social Skills and Positive Approaches to Learning	0.25	0.36*	0.075	0.20
Closeness	0.13	0.19	0.434	0.07
Conflict	-0.06	-0.09	0.888	-0.01
Positive Relationships	0.12	0.17	0.803	0.02

Exhibit E5.2: Estimated IOT Impacts on Social-Emotional Outcomes, by Year: 3-Year-Old Cohort (continued)

Outcome	ITT Regression Adjusted Impact	IOT		
	Impact	Impact	p-value	Effect Size
Teacher-Reported Measures				
ASPI – Aggressive	0.40	0.57	0.318	0.08
ASPI – Inattentive/Hyperactive	-0.02	-0.03	0.972	0.00
ASPI – Low Energy	0.52	0.75	0.241	0.11
ASPI – Oppositional	0.03	0.04	0.953	0.01
ASPI – Peer Interactions	0.64	0.92	0.335	0.08
ASPI – Shy/Socially Reticent	0.07	0.10	0.882	0.01
ASPI – Structured Learning	0.74	10.6	0.136	0.10
ASPI – Teacher Interaction	0.27	0.39	0.649	0.04
Closeness	-0.26	-0.37	0.295	-0.08
Conflict	0.03	0.04	0.959	0.01
Positive Relationships	-0.40	-0.57	0.483	-0.06
1st Grade Year (Spring 2006)				
Parent-Reported Measures				
Aggressive Behavior	-0.05	-0.07	0.624	-0.04
Hyperactive Behavior	-0.11	-0.16	0.127	-0.10
Withdrawn Behavior+	0.02	0.03	0.732	0.03
Total Problem Behavior	-0.15	-0.22	0.439	-0.05
Social Competencies	0.08	0.11	0.317	0.09
Social Skills and Positive Approaches to Learning	0.05	0.07	0.642	0.04
Closeness	0.29	0.42**	0.013	0.15
Conflict	-0.55	-0.79	0.210	-0.12
Positive Relationships	0.77	1.10*	0.098	0.14
Teacher-Reported Measures				
ASPI – Aggressive	-0.54	-0.77	0.266	-0.10
ASPI – Inattentive/Hyperactive	-0.45	-0.65	0.402	-0.08
ASPI – Low Energy	0.44	0.63	0.383	0.09
ASPI – Oppositional	0.04	0.06	0.944	0.01
ASPI – Peer Interactions	-0.43	-0.62	0.584	-0.05
ASPI – Shy/Socially Reticent	0.21	0.30	0.623	0.04
ASPI – Structured Learning	-0.12	-0.17	0.855	-0.02
ASPI – Teacher Interaction	-0.15	-0.22	0.819	-0.02
Closeness	0.28	0.40	0.301	0.09
Conflict	-0.16	-0.23	0.722	-0.03

Exhibit E5.2: Estimated IOT Impacts on Social-Emotional Outcomes, by Year: 3-Year-Old Cohort (continued)

Outcome	ITT Regression Adjusted Impact	IOT		
	Impact	Impact	p-value	Effect Size
Teacher-Reported Measures (cont'd)				
Positive Relationships	0.44	0.63	0.461	0.07

Notes: The three-year old cohort compliance rate is 0.69739.
Sample sizes, standard errors, and confidence intervals for all estimates can be found on the Administration for Children and Families, Office of Planning, Research and Evaluation Website at: http://www.acf.hhs.gov/programs/opre/hs/impact_study/index.html.

Key:

*** p<0.01
** p<0.05
* p<0.10

+ Indicates the reliability (Cronbach alpha) for the measure is <0.6.

Bold IOT impact indicates the outcome passes the Benjamini-Hochberg test for multiple comparisons with a 10-percent false discovery rate.

Exhibit E6.1: Estimated IOT Impacts on Parent-Reported Health Outcomes, by Year: 4-Year-Old Cohort

Outcome	ITT Regression Adjusted Impact	IOT		
	Impact	Impact	p-value	Effect Size
Head Start Year (Spring 2003)				
Parent-Reported Measures				
Child Received Dental Care	0.15	**0.22*** **	0.000	0.45
Child Has Health Insurance Coverage	0.01	0.01	0.733	0.05
Child's Overall Health Status is Excellent/Good	-0.03	-0.04	0.244	-0.12
Child Needs Ongoing Care	0.01	0.01	0.422	0.05
Child Had Care for Injury Last Month	-0.02	-0.03	0.409	-0.09
Kindergarten Year (Spring 2004)				
Parent-Reported Measures				
Child Received Dental Care	0.03	0.04	0.435	0.09
Child Has Health Insurance Coverage	0.04	0.06*	0.056	0.17
Child's Overall Health Status is Excellent/GOod	0.05	0.07*	0.098	0.18
Child Needs Ongoing Care	-0.02	-0.03	0.432	-0.09
Child Had Care for Injury Last Month	0.02	0.03	0.547	0.10
1st Grade Year (Spring 2005)				
Parent-Reported Measures				
Child Received Dental Care	0.02	0.03	0.550	0.06
Child Has Health Insurance Coverage	0.04	0.06**	0.044	0.17
Child's Overall Health Status is Excellent/Good	-0.01	-0.01	0.858	-0.04
Child Needs Ongoing Care	0.02	0.03	0.291	0.09
Child Had Care for Injury Last Month	0.02	0.03	0.303	0.09

Notes: The four-year old cohort compliance rate is 0.66897.
Sample sizes, standard errors, and confidence intervals for all estimates can be found on the Administration for Children and Families, Office of Planning, Research and Evaluation Website at: http://www.acf.hhs.gov/programs/opre/hs/impact_study/index.html.

Key:
*** p<0.01
** p<0.05
* p<0.10
Bold IOT impact indicates the outcome passes the Benjamini-Hochberg test for multiple comparisons with a 10-percent false discovery rate.

Exhibit E6.2: Estimated IOT Impacts on Parent-Reported Health Outcomes, by Year: 3-Year-Old Cohort

Outcome	ITT Regression Adjusted Impact	IOT		
	Impact	Impact	p-value	Effect Size
Head Start Year (Spring 2003)				
Parent-Reported Measures				
Child Received Dental Care	0.17	**0.24*** **	0.000	0.49
Child Has Health Insurance Coverage	0.00	0.00	0.803	0.00
Child's Overall Health Status is Excellent/Good	0.05	0.07**	0.045	0.17
Child Needs Ongoing Care	0.00	0.00	0.988	0.00
Child Had Care for Injury Last Month	-0.01	-0.01	0.699	-0.05
Age 4 Year (Spring 2004)				
Parent-Reported Measures				
Child Received Dental Care	0.10	**0.14*** **	0.001	0.30
Child Has Health Insurance Coverage	0.00	0.00	0.935	0.00
Child's Overall Health Status is Excellent/Good	0.00	0.00	0.851	0.00
Child Needs Ongoing Care	0.01	0.01	0.739	0.04
Child Had Care for Injury Last Month	0.03	0.04	0.089	0.15
Kindergarten Year (Spring 2005)				
Parent-Reported Measures				
Child Received Dental Care	0.03	0.04	0.270	0.10
Child Has Health Insurance Coverage	0.04	0.06**	0.044	0.19
Child's Overall Health Status is Excellent/Good	0.00	0.00	0.889	0.00
Child Needs Ongoing Care	-0.03	-0.04	0.114	-0.11
Child Had Care for Injury Last Month	0.00	0.00	0.985	0.00
1st Grade Year (Spring 2006)				
Parent-Reported Measures				
Child Received Dental Care	0.01	0.01	0.786	0.03
Child Has Health Insurance Coverage	0.02	0.03	0.252	0.10
Child's Overall Health Status is Excellent/Good	0.02	0.03	0.434	0.08

Exhibit E6.2: Estimated IOT Impacts on Parent-Reported Health Outcomes, by Year: 3-Year-Old Cohort (continued)

Outcome	ITT Regression Adjusted Impact	IOT		
	Impact	Impact	p-value	Effect Size
Parent-Reported Measures (cont'd)				
Child Needs Ongoing Care	-0.01	-0.01	0.578	-0.04
Child Had Care for Injury Last Month	0.02	0.03	0.294	0.11

Notes: The three-year old cohort compliance rate is 0.69739.
Sample sizes, standard errors, and confidence intervals for all estimates can be found on the Administration for Children and Families, Office of Planning, Research and Evaluation Website at: http://www.acf.hhs.gov/programs/opre/hs/impact_study/index.html.

Key:

*** p<0.01
** p<0.05
* p<0.10

Bold IOT impact indicates the outcome passes the Benjamini-Hochberg test for multiple comparisons with a 10-percent false discovery rate.

Exhibit E7.1: Estimated IOT Impacts on Parenting Outcomes, by Year: 4-Year-Old Cohort

Outcome	ITT Regression Adjusted Impact	IOT		
	Impact	Impact	p-value	Effect Size
Head Start Year (Spring 2003)				
Parent-Reported Measures				
Parent Spanked Child in Last Week	-0.01	-0.01	0.750	-0.03
Parent Used Time Out in Last Week	-0.08	-0.12**	0.025	-0.26
Parent Read to Child in Last Week	0.03	0.04	0.396	0.10
Parental Safety Practices Scale+	0.03	0.04	0.382	0.00
Family Cultural Enrichment Scale+	0.08	0.12	0.368	0.08
Kindergarten Year (Spring 2004)				
Parent-Reported Measures				
Parent Spanked Child in Last Week	0.00	0.00	0.869	0.00
Parent Used Time Out in Last Week	0.01	0.01	0.689	0.03
Parent Read to Child in Last Week	-0.03	-0.04	0.385	-0.09
Parental Safety Practices Scale+	0.04	0.06	0.156	0.16
Family Cultural Enrichment Scale+	0.14	0.21	0.142	0.15
Parent Style: Authoritarian	0.00	0.00	0.981	0.00
Parent Style: Authoritative	0.05	0.07	0.164	0.15
Parent Style: Neglectful	-0.03	-0.04	0.211	-0.14
Parent Style: Permissive	-0.02	-0.03	0.447	-0.08
Teacher-Reported Measures				
School Contact and Communication	-0.01	-0.01	0.845	-0.04
Parent Participation	-0.01	-0.01	0.841	-0.05
1st Grade Year (Spring 2005)				
Parent-Reported Measures				
Parent Spanked Child in Last Week	0.00	0.00	0.976	0.00
Parent Used Time Out in Last Week	-0.04	-0.06	0.322	-0.12
Parent Read to Child in Last Week	-0.01	-0.01	0.733	-0.03
Family Cultural Enrichment Scale+	0.04	0.06	0.612	0.04
Parent Style: Authoritarian	-0.03	-0.04	0.199	-0.15
Parent Style: Authoritative	0.04	0.06	0.158	0.12
Parent Style: Neglectful	-0.02	-0.03	0.327	-0.11
Parent Style: Permissive	0.00	0.00	0.936	0.00

Exhibit E7.1: Estimated IOT Impacts on Parenting Outcomes, by Year: 4-Year-Old Cohort (continued)

Outcome	ITT Regression Adjusted Impact	IOT		
	Impact	Impact	p-value	Effect Size
Teacher-Reported Measures				
School Contact and Communication	-0.02	-0.03	0.570	-0.08
Parent Participation+	-0.01	-0.01	0.817	-0.04

Notes: The four-year old cohort compliance rate is 0.66897.
Sample sizes, standard errors, and confidence intervals for all estimates can be found on the Administration for Children and Families, Office of Planning, Research and Evaluation Website at: http://www.acf.hhs.gov/programs/opre/hs/impact_study/index.html.

Key:

*** p<0.01
** p<0.05
* p<0.10

+ Indicates the reliability (Cronbach alpha) for the measure is <0.6.

Bold IOT impact indicates the outcome passes the Benjamini-Hochberg test for multiple comparisons with a 10-percent false discovery rate.

Exhibit E7.2: Estimated IOT Impacts on Parenting Outcomes, by Year: 3-Year-Old Cohort

Outcome	ITT Regression Adjusted Impact	IOT		
	Impact	Impact	p-value	Effect Size
Head Start Year (Spring 2003)				
Parent-Reported Measurers				
Parent Spanked Child in Last Week	-0.07	**-0.10****	0.025	-0.20
Parent Used Time Out in Last Week	-0.04	-0.06	0.205	-0.12
Parent Read to Child in Last Week	0.07	**0.10****	0.030	0.22
Parental Safety Practices Scale+	0.03	0.04	0.146	0.13
Family Cultural Enrichment Scale+	0.25	**0.36*****	0.000	0.26
Age 4 Year (Spring 2004)				
Parent-Reported Measurers				
Parent Spanked Child in Last Week	0.01	0.01	0.635	0.03
Parent Used Time Out in Last Week	-0.02	-0.03	0.355	-0.06
Parent Read to Child in Last Week	-0.01	-0.01	0.827	-0.03
Parental Safety Practices Scale+	0.02	0.03	0.313	0.08
Family Cultural Enrichment Scale+	0.04	0.06	0.593	0.04
Parent Style: Authoritarian	-0.04	**-0.06*****	0.005	-0.21
Parent Style: Authoritative	0.04	0.06	0.186	0.12
Parent Style: Neglectful	0.00	0.00	0.826	0.00
Parent Style: Permissive	0.00	0.00	0.893	0.00
Kindergarten Year (Spring 2005)				
Parent-Reported Measures				
Parent Spanked Child in Last Week	-0.04	-0.06*	0.070	-0.12
Parent Used Time Out in Last Week	-0.07	-0.10**	0.013	-0.20
Parent Read to Child in Last Week	0.03	0.04	0.305	0.09
Parental Safety Practices Scale+	0.01	0.01	0.714	0.04
Family Cultural Enrichment Scale+	0.00	0.00	0.968	0.00
Parent Style: Authoritarian	0.00	0.00	0.950	0.00
Parent Style: Authoritative	0.00	0.00	0.905	0.00
Parent Style: Neglectful	-0.02	-0.03	0.202	-0.11
Parent Style: Permissive	0.03	0.04	0.310	0.11
Teacher-Reported Measures				
School Contact and Communication	0.00	0.00	0.879	0.00
Parent Participation+	0.00	0.00	0.886	0.00

Exhibit E7.2: Estimated IOT Impacts on Parenting Outcomes, by Year: 3-Year-Old Cohort (continued)

Outcome	ITT Regression Adjusted Impact	IOT		
	Impact	Impact	p-value	Effect Size
1st Grade Year (Spring 2006)				
Parent-Reported Measures				
Parent Spanked Child in Last Week	-0.03	-0.04	0.183	-0.10
Parent Used Time Out in Last Week	-0.05	-0.07*	0.075	-0.14
Parent Read to Child in Last Week	0.01	0.01	0.746	0.03
Family Cultural Enrichment Scale+	0.01	0.01	0.879	0.01
Parent Style: Authoritarian	-0.03	-0.04**	0.046	-0.16
Parent Style: Authoritative	0.00	0.00	0.877	0.00
Parent Style: Neglectful	0.00	0.00	0.974	0.00
Parent Style: Permissive	0.02	0.03	0.320	0.08
Teacher-Reported Measures				
School Contact and Communication	0.02	0.03	0.453	0.07
Parent Participation+	0.01	0.01	0.626	0.04

Notes: The three-year old cohort compliance rate is 0.69739.
Sample sizes, standard errors, and confidence intervals for all estimates can be found on the Administration for Children and Families, Office of Planning, Research and Evaluation Website at: http://www.acf.hhs.gov/programs/opre/hs/impact_study/index.html.

Key:

*** p<0.01
** p<0.05
* p<0.10

+ Indicates the reliability (Cronbach alpha) for the measure is <0.6.

Bold IOT impact indicates the outcome passes the Benjamini-Hochberg test for multiple comparisons with a 10-percent false discovery rate.

Appendix F

The Impact of Head Start on Children and Head Start, Child Care, and Early School Experiences in Puerto Rico

Appendix F: The Impact of Head Start on Children and Head Start, Child Care, and Early School Experiences in Puerto Rico

Introduction

This chapter serves a dual role. The chapter provides the estimated impacts for Head Start programs operating on the island of Puerto Rico and it provides a description of the services and experiences of children during their Head Start years. The estimated impacts are reported for the cognitive, social-emotional, health and parenting practices domains as reported for the mainland study sample in earlier chapters. The chapter also provides a look at what access to Head Start means for children--what if any difference it makes in the type of care giving arrangement, whether children attend preschool, where they attend preschool, characteristics of their early childhood care and education, qualifications of their care providers and teachers, and the quality of their experiences in these settings. Results are presented separately for the 3- and 4-year-old cohorts.

Readers must, however, use these results with caution because the National Head Start Impact Study was not designed to produce reliable impact estimates for specific geographic areas such as Puerto Rico. The sample in Puerto Rico includes only three Head Start programs consisting of 22 centers. There are fewer than 200 children with completed assessments and a parent interview each spring. These sample sizes are too small to reliably estimate regression models containing baseline child covariates. Moreover, they make standard errors large and unstable.

Impacts

The first section of this chapter reports estimated impacts for Head Start programs operating on the island of Puerto Rico. Child assessment and parent interviews were conducted in Spanish at each data collection point for the Puerto Rico sample. The Spanish parent interview is a direct translation of the English parent interview. The child assessment[1] was designed to measure the child's skills in language and literacy, math, and prewriting. The components of the direct child assessment included the Test de Vocabulario en Imágenes

[1] See Chapter 2 and the Technical Report for the Head Start Impact Study for discussion of the measures.

Peabody (TVIP-adapted), the Spanish versión of the PPVT; the letter-word identification(Identificación de letras y palabras), math (Problemas aplicados), and spelling (Dictado) tests from the Batería Woodcock-Muñoz, the Spanish versión of the Woodcock-Johnson Achievement Tests; and the Spanish Versión of the Preschool Comprehensive Test of Phonological and Print Processing (CTOPPP). Children were also administered the direct translation for 3 measures—Counting Bears, Color Identification, and the McCarthy Draw-a Design Task. Similar to the earlier chapters, the impacts are presented by child cohort for each domain (i.e., cognitive, social-emotional, health, and parenting practices).

As noted above, it is important for readers to understand that these estimates must be used with caution. The National Head Start Impact Study was designed to produce national Head Start impact estimates by child characteristics such as age cohort (3- and 4-year-olds), gender, race/ethnicity, and language (English vs. Spanish) for children in the 50 United States and Puerto Rico, combined. It was not designed to produce estimates for specific geographic areas, including Puerto Rico, separately. The sample in Puerto Rico includes only three Head Start programs, consisting of 22 centers. There are fewer than 200 children with completed assessments and a parent interview each spring from Puerto Rico. From this sample it is generally possible to produce simple differences in means between the treatment and non-Head Start groups. However, the sample sizes in Puerto Rico are too small to reliably estimate regression models containing baseline child covariates as was done in the preceding chapters. Furthermore, because of the small sample sizes, the estimated standard errors – which are used in the tests of statistical significance -- are large and unstable and create unreliable effect sizes.

Due to the different tests administered to the children in Puerto Rico and the language difference used in all the data collection instruments, it is not appropriate to include Puerto Rico in the analysis with the mainland. Yet, because the Head Start program in Puerto Rico and the mainland Head Start program have the same focus—to foster the cognitive and social-emotional development of young children from low income families, it is important to examine the differences for children in the Head Start group and the control group in Puerto Rico along the same domains discussed in the earlier chapters (i.e., cognitive, social-emotional, health, and parenting practices). Within these limitations, the statistically significant mean differences ($p \leq 0.10$) are indicated in Exhibits F.1 through F.8.

Impacts for the 4-Year-Old Cohort

As shown in Exhibits F.1 through F.4, there were four statistically significant impacts for children in the 4-year-old cohort from Puerto Rico. At the end of the Head Start/preschool year, children in the Head Start group scored higher on the CTOPPP than children in the control group. Likewise, at the end of 1st grade, children in the Head Start group performed better on the Woodcock-Muñoz Identificación de letras y palabras than their counterparts in the control group. At the end of kindergarten, parents of children in the Head Start group reported having more conflict with their children than parents of children in the control group. While at the end of 1st grade, parents of the children in the Head Start group reported less closeness in the relationship with their children than did parents of children in the control group. These findings are favorable for the cognitive outcomes, but unfavorable for the social-emotional outcomes. No other impacts on children in the 4-year-old cohort from Puerto Rico were found for any other outcomes over the years of the study.

Impacts for the 3-Year-Old Cohort

Also shown in Exhibits F.5 through F.8, several statistically significant impacts were found for children in the 3-year-old cohort from Puerto Rico. Nine cognitive outcomes were significant across the Head Start and kindergarten years. Children in the Head Start group scored higher on the Woodcock-Muñoz Problemas aplicados than did their peers in the control group at the end of three consecutive years--at the end of the Head Start year, the age 4 year, and the end of kindergarten. At the end of the Head Start year, children in the Head Start group were also better able to identify their colors than those in the control group. At the end of the age 4 year, children in the Head Start group continued to outperform their peers in the control group on the TVIP (adapted), a measure of receptive vocabulary and the parents of children in the Head Start group reported a higher level of emerging literacy skills for their children than the parents of the control group children. Finally, children in the Head Start group outperformed their peers in the control group on three language and literacy assessments at the end of kindergarten, with stronger performances on the TVIP (adapted), the Woodcock-Muñoz Identificación de letras y palabras, and CTOPPP Elision. The Head Start group children in the 3-year-old cohort demonstrate some evidence of sustained cognitive impacts compared to their counterparts in the control group.

Exhibit F.1: Estimated Impacts on Cognitive Outcomes by Year: 4-Year-Old Cohort

Outcome	Mean Estimates			
	Head Start Group	Control Group	Head Start – Control	p-value
Head Start Baseline (Fall 2002)				
Language and Literacy Measures				
TVIP (Adapted)	264.40	266.64		
WM Identificación de letras y palabras	347.81	350.85		
WM Dictado	352.50	359.18		
CTOPPP Elision	243.31	247.36		
Color Identification	0.30	0.26		
Pre-writing Measure+				
McCarthy Draw-a-Design	3.24	3.38		
Parent-Reported Literacy Measure				
Emergent Literacy Scale	2.07	2.10		
Math Skills Measures+				
WM Problemas aplicados	395.30	395.86		
Counting Bears	0.27	0.24		
Head Start Year (Spring 2003)				
Language and Literacy Measures				
TVIP (Adapted)	308.84	303.92	4.92	0.714
WM Identificación de letras y palabras	354.59	354.08	0.51	0.842
WM Dictado	360.65	359.72	0.93	0.859
CTOPPP Elision	250.12	232.36	17.77*	0.076
Color Identification	0.38	0.39	-0.01	0.926
Pre-writing Measure				
McCarthy Draw-a-Design	3.56	3.63	-0.07	0.775
Parent-Reported Literacy Measure				
Emergent Literacy Scale	3.09	2.64	0.45	0.104
Math Skills Measures				
WM Problemas aplicados	402.52	401.52	1.00	0.813
Counting Bears	0.40	0.26	0.14	0.231
Kindergarten Year (Spring 2004)				
Language and Literacy Measures				
TVIP (Adapted)	370.29	378.81	-8.52	0.392
WM Identificación de letras y palabras	373.80	370.27	3.53	0.636
WM Dictado	400.43	401.06	-0.64	0.922
CTOPPP Elision	253.56	245.11	8.45	0.388
Math Skills Measures				
WM Problemas aplicados	425.88	422.52	3.36	0.448

Exhibit F.1: Estimated Impacts on Cognitive Outcomes by Year: 4-Year-Old Cohort (continued)

Outcome	Mean Estimates			
	Head Start Group	Control Group	Head Start – Control	p-value
1st Grade Year (Spring 2005)				
Language and Literacy Measures				
TVIP (Adapted)	412.81	413.28	-0.48	0.975
WM Identificación de letras y palabras	449.75	408.20	41.55**	0.047
WM Dictado	433.00	424.15	8.85	0.237
Math Skills Measures				
WM Problemas aplicados	443.72	437.35	6.37	0.220

Key:
*** $p < 0.01$
** $p < 0.05$
* $p < 0.10$

Exhibit F.2: Estimated Impacts on Social-Emotional Outcomes by Year: 4-Year-Old Cohort[a]

Outcome	Mean Estimates			
	Head Start Group	Control Group	Head Start – Control	p-value
Head Start Baseline (Fall 2002)				
Parent-Reported Measures				
Aggressive Behavior	3.01	4.27		
Hyperactive Behavior	2.52	2.93		
Withdrawn Behavior	0.85	0.74		
Total Problem Behavior	6.87	8.51		
Social Competencies Checklist	11.10	11.05		
Social Skills and Positive Approaches to Learning	12.59	12.49		
Head Start Year (Spring 2003)				
Parent-Reported Measures				
Aggressive Behavior	3.33	3.86	-0.53	0.254
Hyperactive Behavior	2.49	2.74	-0.25	0.625
Withdrawn Behavior	0.87	0.71	0.16	0.449
Total Problem Behavior	7.17	7.82	-0.65	0.499
Social Competencies Checklist	11.18	11.10	0.08	0.826
Social Skills and Positive Approaches to Learning	12.87	13.04	-0.17	0.634
Closeness	33.76	32.99	0.78	0.429
Conflict	21.61	20.85	0.77	0.676
Positive Relationships	60.15	60.08	0.07	0.976
Kindergarten Year (Spring 2004)				
Parent-Reported Measures				
Aggressive Behavior	2.72	2.73	-0.01	0.990
Hyperactive Behavior	2.24	2.34	-0.10	0.771
Withdrawn Behavior	0.99	0.90	0.09	0.726
Total Problem Behavior	6.37	6.45	-0.08	0.911
Social Competencies Checklist	11.07	11.27	-0.20	0.312
Social Skills and Positive Approaches to Learning	12.60	13.10	-0.51	0.129
Closeness	33.65	33.86	-0.22	0.658
Conflict	22.36	19.79	2.57**	0.047
Positive Relationships	59.29	61.83	-2.54	0.120
1st Grade Year (Spring 2005)				
Parent-Reported Measures				
Aggressive Behavior	2.48	2.99	-0.50	0.342
Hyperactive Behavior	1.76	2.42	-0.66	0.108
Withdrawn Behavior	0.84	0.85	-0.01	0.973
Total Problem Behavior	5.61	7.07	-1.47	0.160
Social Competencies Checklist	11.11	11.20	-0.09	0.788
Social Skills and Positive Approaches to Learning	12.70	12.88	-0.19	0.576

Exhibit F.2: Estimated Impacts on Social-Emotional Outcomes by Year: 4-Year-Old Cohort[a] (continued)

Outcome	Mean Estimates			
	Head Start Group	Control Group	Head Start – Control	p-value
Parent-Reported Measures (cont'd)				
Closeness	33.46	34.26	-0.80*	0.074
Conflict	18.74	18.69	0.05	0.981
Positive Relationships	62.72	63.35	-0.63	0.785

Key:
******* $p < 0.01$
** $p < 0.05$
* $p < 0.10$

[a]As noted in Chapter 5, for the following social-emotional measures lower values are better (and "negative" impacts are "good"): (1) parent reports of Aggressive, Hyperactive, Withdrawn and Total Problem Behaviors; and (2) the Pianta parent measure of Conflict.

Exhibit F.3: Estimated Impacts on Health Outcomes by Year: 4-Year-Old Cohort

Outcome	Mean Estimates			
	Head Start Group	Control Group	Head Start – Control	p-value
Head Start Baseline (Fall 2002)				
Parent-Reported Measures				
Child Received Dental Care	0.89	0.68		
Child Has Health Insurance Coverage	0.96	0.93		
Child's Overall Health Status is Excellent/Good	0.53	0.82		
Child Needs Ongoing Care	0.35	0.28		
Child Had Care for Injury Last Month	0.08	0.05		
Head Start Year (Spring 2003)				
Parent-Reported Measures				
Child Received Dental Care	0.67	0.54	0.12	0.412
Child Has Health Insurance Coverage	0.96	0.93	0.02	0.535
Child's Overall Health Status is Excellent/Good	0.62	0.72	-0.10	0.284
Child Needs Ongoing Care	0.29	0.24	0.05	0.688
Child Had Care for Injury Last Month	0.32	0.18	0.14	0.358
Kindergarten Year (Spring 2004)				
Parent-Reported Measures				
Child Received Dental Care	0.78	0.63	0.16	0.276
Child Has Health Insurance Coverage	0.93	0.87	0.07	0.495
Child's Overall Health Status is Excellent/Good	0.67	0.69	-0.02	0.879
Child Needs Ongoing Care	0.37	0.46	-0.10	0.533
Child Had Care for Injury Last Month	0.28	0.20	0.09	0.399
1st Grade Year (Spring 2005)				
Parent-Reported Measures				
Child Received Dental Care	0.74	0.69	0.05	0.478
Child Has Health Insurance Coverage	0.98	0.94	0.04	--
Child's Overall Health Status is Excellent/Good	0.76	0.71	0.05	0.587
Child Needs Ongoing Care	0.29	0.24	0.05	0.626
Child Had Care for Injury Last Month	0.14	0.13	0.01	0.793

Key:
*** $p < 0.01$
** $p < 0.05$
* $p < 0.10$

Note: p-value was not calculated for Child Has Health Insurance Coverage in 2005 due to high response rates which resulted in a standard deviation of 0 for one or both groups (Sudaan warning).

Exhibit F.4: Estimated Impacts on Parenting Outcomes by Year: 4-Year-Old Cohort

Outcome	Mean Estimates			
	Head Start Group	Control Group	Head Start – Control	p-value
Head Start Baseline (Fall 2002)				
Parent-Reported Measures				
Parent Spanked Child in Last Week	0.55	0.63		
Parent Used Time Out in Last Week	0.60	0.48		
Parent Read to Child in Last Week	0.26	0.25		
Parental Safety Practices Scale	3.42	3.52		
Family Cultural Enrichment Scale	3.70	3.35		
Head Start Year (Spring 2003)				
Parent-Reported Measures				
Parent Spanked Child in Last Week	0.64	0.80	-0.16	0.194
Parent Used Time Out in Last Week	0.62	0.61	0.01	0.951
Parent Read to Child in Last Week	0.17	0.20	-0.03	0.815
Parental Safety Practices Scale	3.45	3.47	-0.02	0.789
Family Cultural Enrichment Scale	3.97	3.32	0.66	0.174
Kindergarten Year (Spring 2004)				
Parent-Reported Measures				
Parent Spanked Child in Last Week	0.49	0.41	0.08	0.462
Parent Used Time Out in Last Week	0.57	0.38	0.18	0.170
Parent Read to Child in Last Week	0.19	0.14	0.05	0.525
Parental Safety Practices Scale	3.35	3.46	-0.11	0.118
Family Cultural Enrichment Scale	4.39	4.01	0.38	0.107
Parenting Style: Authoritarian	0.05	0.14	-0.09	0.447
Parenting Style: Authoritative	0.76	0.56	0.20	0.170
Parenting Style: Neglectful	0.07	0.08	-0.01	-0.853
Parenting Style: Permissive	0.12	0.25	-0.13	0.379
1st Grade Year (Spring 2005)				
Parent-Reported Measures				
Parent Spanked Child in Last Week	0.52	0.59	-0.08	0.450
Parent Used Time Out in Last Week	0.54	0.57	-0.03	0.826
Parent Read to Child in Last Week	0.18	0.28	-0.10	0.283
Family Cultural Enrichment Scale	3.75	3.69	0.05	0.867
Parenting Style: Authoritarian	0.08	0.04	0.04	0.471
Parenting Style: Authoritative	0.83	0.86	-0.03	0.692
Parenting Style: Neglectful	0.01	0.00	0.01	--
Parenting Style: Permissive	0.08	0.09	-0.01	0.858

Key:
******* $p < 0.01$
** $p < 0.05$
* $p < 0.10$

Note: p-value was not calculated for Parenting Style: Neglectful in 2005 due to low response rates which resulted in a standard deviation of 0 for one or both groups (Sudaan warning).

Exhibit F.5: Estimated Impacts on Cognitive Outcomes by Year: 3-Year-Old Cohort

Outcome	Mean Estimates			
	Head Start Group	Control Group	Head Start – Control	p-value
Head Start Baseline (Fall 2002)				
Language and Literacy Measures				
TVIP (Adapted)	228.73	237.69		
WM Identificación de letras y palabras	347.49	346.39		
WM Dictado	326.04	328.44		
CTOPPP Elision	231.63	233.18		
Color Identification	0.12	0.12		
Pre-writing Measure+				
McCarthy Draw-a-Design	2.55	2.42		
Parent-Reported Literacy Measure				
Emergent Literacy Scale	1.73	1.63		
Math Skills Measures				
WM Problemas aplicados	385.48	384.72		
Counting Bears	0.09	0.11		
Head Start Year (Spring 2003)				
Language and Literacy Measures				
TVIP (Adapted)	266.69	268.77	-2.08	0.761
WM Identificación de letras y palabras	348.66	348.46	0.20	0.939
WM Dictado	345.48	339.34	6.15	0.276
CTOPPP Elision	235.83	246.90	-11.07	0.264
Color Identification	0.36	0.14	0.22**	0.017
Pre-writing Measure				
McCarthy Draw-a-Design	2.98	2.90	0.08	0.768
Parent-Reported Literacy Measure				
Emergent Literacy Scale	2.37	1.92	0.45	0.111
Math Skills Measures				
WM Problemas aplicados	394.96	382.44	12.52***	0.005
Counting Bears	0.20	0.08	0.12	0.120
Age 4 Year (Spring 2004)				
Language and Literacy Measures				
TVIP (Adapted)	354.75	336.46	18.28*	0.066
WM Identificación de letras y palabras	362.79	362.21	0.59	0.733
WM Dictado	383.77	382.52	1.24	0.639
CTOPPP Elision	240.00	228.82	11.18	0.333
Color Identification	0.71	0.71	0.00	0.968
Pre-writing Measure				
McCarthy Draw-a-Design	4.86	4.37	0.49	0.130
Parent-Reported Literacy Measure				
Emergent Literacy Scale	3.67	3.38	0.29*	0.082

Exhibit F.5: Estimated Impacts on Cognitive Outcomes by Year: 3-Year-Old Cohort (continued)

Outcome	Mean Estimates			
	Head Start Group	Control Group	Head Start – Control	p-value
Math Skills Measures				
WM Problemas aplicados	419.88	410.46	9.42**	0.028
Counting Bears	0.55	0.65	-0.10	0.496
Kindergarten Year (Spring 2005)				
Language and Literacy Measures				
TVIP (Adapted)	407.96	386.98	20.97***	0.006
WM Identificación de letras y palabras	389.83	378.31	11.52**	0.019
WM Dictado	412.64	408.34	4.31	0.301
CTOPPP Elision	266.39	251.72	14.67***	0.000
Math Skills Measures				
WM Problemas aplicados	438.23	430.20	8.03***	0.010
1st Grade Year (Spring 2006)				
Language and Literacy Measures				
TVIP (Adapted)	437.55	428.18	9.37	0.339
WM Identificación de letras y palabras	490.28	468.12	22.16	0.409
WM Dictado	446.28	442.21	4.07	0.434
Math Skills Measures				
WM Problemas aplicados	454.30	447.52	6.78	0.107

Key:
*** $p < 0.01$
** $p < 0.05$
* $p < 0.10$

Exhibit F.6: Estimated Impacts on Social-Emotional Outcomes by Year: 3-Year-Old Cohort[a]

Outcome	Mean Estimates			
	Head Start Group	Control Group	Head Start – Control	p-value
Head Start Baseline (Fall 2002)				
Parent-Reported Measures				
Aggressive Behavior	3.60	3.47		
Hyperactive Behavior	2.55	2.81		
Withdrawn Behavior	0.62	0.81		
Total Problem Behavior	7.20	7.42		
Social Competencies Checklist	10.85	11.10		
Social Skills and Positive Approaches to Learning	12.33	11.81		
Head Start Year (Spring 2003)				
Parent-Reported Measures				
Aggressive Behavior	3.83	3.40	0.43	0.303
Hyperactive Behavior	2.57	3.22	-0.65***	0.009
Withdrawn Behavior	1.08	0.94	0.14	0.553
Total Problem Behavior	7.83	8.05	-0.22	0.763
Social Competencies Checklist	10.96	10.93	0.03	0.833
Social Skills and Positive Approaches to Learning	12.54	12.36	0.18	0.605
Closeness	33.74	33.70	0.04	0.937
Conflict	21.40	22.75	-1.36	0.451
Positive Relationships	60.12	58.80	1.32	0.542
Age 4 Year (Spring 2004)				
Parent-Reported Measures				
Aggressive Behavior	2.97	3.24	-0.27	0.661
Hyperactive Behavior	2.33	2.66	-0.33	0.488
Withdrawn Behavior	0.90	1.05	-0.14	0.589
Total Problem Behavior	6.38	7.43	-1.04	0.444
Social Competencies Checklist	11.15	10.94	0.21	0.231
Social Skills and Positive Approaches to Learning	13.06	12.46	0.60	0.161
Closeness	33.42	33.90	-0.48	0.265
Conflict	20.71	22.89	-2.19	0.266
Positive Relationships	60.61	59.00	1.61	0.489
Kindergarten Year (Spring 2005)				
Parent-Reported Measures				
Aggressive Behavior	2.94	2.79	0.15	0.685
Hyperactive Behavior	2.05	1.98	0.07	0.874
Withdrawn Behavior	0.47	0.94	-0.46**	0.030
Total Problem Behavior	5.75	6.25	-0.51	0.537
Social Competencies Checklist	11.46	11.27	0.19*	0.074
Social Skills and Positive Approaches to Learning	13.14	12.36	0.78***	0.003

Exhibit F.6: Estimated Impacts on Social-Emotional Outcomes by Year: 3-Year-Old Cohort[a] (continued)

Outcome	Mean Estimates			
	Head Start Group	Control Group	Head Start – Control	p-value
Parent-Reported Measures (cont'd)				
Closeness	34.42	34.36	0.06	0.821
Conflict	17.15	19.83	-2.68	0.164
Positive Relationships	65.27	62.22	3.05	0.144
1st Grade Year (Spring 2006)				
Parent-Reported Measures				
Aggressive Behavior	2.59	3.15	-0.56	0.100
Hyperactive Behavior	1.89	2.24	-0.36	0.443
Withdrawn Behavior	0.97	1.08	-0.11	0.615
Total Problem Behavior	6.02	6.81	-0.79	0.369
Social Competencies Checklist	11.28	11.27	0.01	0.971
Social Skills and Positive Approaches to Learning	12.99	12.60	0.39	0.188
Closeness	33.71	34.33	-0.62	0.301
Conflict	19.84	19.35	0.49	0.825
Positive Relationships	61.51	62.89	-1.38	0.573

Key:
******* $p < 0.01$
** $p < 0.05$
* $p < 0.10$

[a]As noted in Chapter 5, for the following social-emotional measures lower values are better (and "negative" impacts are "good"): (1) parent reports of Aggressive, Hyperactive, Withdrawn and Total Problem Behaviors; and (2) the Pianta parent measure of Conflict.

Exhibit F.7: Estimated Impacts on Health Outcomes by Year: 3-Year-Old Cohort

Outcome	Mean Estimates			
	Head Start Group	Control Group	Head Start – Control	p-value
Head Start Baseline (Fall 2002)				
Parent-Reported Measures				
Child Received Dental Care	0.87	0.63		
Child Has Health Insurance Coverage	0.94	0.92		
Child's Overall Health Status is Excellent/Good	0.69	0.75		
Child Needs Ongoing Care	0.23	0.14		
Child Had Care for Injury Last Month	0.10	0.11		
Head Start Year (Spring 2003)				
Parent-Reported Measures				
Child Received Dental Care	0.65	0.51	0.14	0.362
Child Has Health Insurance Coverage	0.89	0.95	-0.06	0.178
Child's Overall Health Status is Excellent/Good	0.72	0.72	0.00	0.984
Child Needs Ongoing Care	0.24	0.26	-0.02	0.844
Child Had Care for Injury Last Month	0.44	0.46	-0.01	0.923
Age 4 Year (Spring 2004)				
Parent-Reported Measures				
Child Received Dental Care	0.75	0.72	0.03	0.753
Child Has Health Insurance Coverage	0.97	1.00	-0.03	--
Child's Overall Health Status is Excellent/Good	0.65	0.67	-0.02	0.832
Child Needs Ongoing Care	0.28	0.38	-0.10	0.324
Child Had Care for Injury Last Month	0.18	0.39	-0.21*	0.080
Kindergarten Year (Spring 2005)				
Parent-Reported Measures				
Child Received Dental Care	0.78	0.75	0.03	0.727
Child Has Health Insurance Coverage	1.00	1.00	0.00	--
Child's Overall Health Status is Excellent/Good	0.80	0.81	-0.01	0.929
Child Needs Ongoing Care	0.22	0.28	-0.06	0.421
Child Had Care for Injury Last Month	0.06	0.27	-0.21**	0.016

Exhibit F.7: Estimated Impacts on Health Outcomes by Year: 3-Year-Old Cohort (continued)

Outcome	Mean Estimates			
	Head Start Group	Control Group	Head Start – Control	p-value
1st Grade Year (Spring 2006)				
Parent-Reported Measures				
Child Received Dental Care	0.78	0.94	-0.16**	0.020
Child Has Health Insurance Coverage	0.99	0.92	0.07	--
Child's Overall Health Status is Excellent/Good	0.82	0.81	0.01	0.947
Child Needs Ongoing Care	0.21	0.22	-0.02	0.900
Child Had Care for Injury Last Month	0.22	0.12	0.10	0.298

Key:
*** $p < 0.01$
** $p < 0.05$
* $p < 0.10$

Note: p-values were not calculated for Child Has Health Insurance Coverage in 2004, 2005, and 2006 due to high response rates which resulted in a standard deviation of 0 for one or both groups (Sudaan Warning).

Exhibit F.8: Estimated Impacts on Parenting Outcomes by Year: 3-Year-Old Cohort

Outcome	Mean Estimates			
	Head Start Group	Control Group	Head Start – Control	p-value
Head Start Baseline (Fall 2002)				
Parent-Reported Measures				
Parent Spanked Child in Last Week	0.58	0.80		
Parent Used Time Out in Last Week	0.48	0.50		
Parent Read to Child in Last Week	0.16	0.18		
Parental Safety Practices Scale	3.46	3.47		
Family Cultural Enrichment Scale	3.38	3.95		
Head Start Year (Spring 2003)				
Parent-Reported Measures				
Parent Spanked Child in Last Week	0.60	0.81	-0.20**	0.037
Parent Used Time Out in Last Week	0.60	0.57	0.03	0.825
Parent Read to Child in Last Week	0.11	0.09	0.02	0.789
Parental Safety Practices Scale	3.49	3.46	0.03	0.670
Family Cultural Enrichment Scale	4.12	3.40	0.73***	0.009
Age 4 Year (Spring 2004)				
Parent-Reported Measures				
Parent Spanked Child in Last Week	0.56	0.66	-0.10	0.275
Parent Used Time Out in Last Week	0.53	0.51	0.02	0.810
Parent Read to Child in Last Week	0.21	0.20	0.01	0.851
Parental Safety Practices Scale	3.53	3.52	0.00	0.950
Family Cultural Enrichment Scale	4.36	4.73	-0.37	0.196
Parenting Style: Authoritarian	0.14	0.10	0.04	0.607
Parenting Style: Authoritative	0.62	0.74	-0.12	0.176
Parenting Style: Neglectful	0.07	0.00	0.07	--
Parenting Style: Permissive	0.20	0.16	0.04	0.628
Kindergarten Year (Spring 2005)				
Parent-Reported Measures				
Parent Spanked Child in Last Week	0.48	0.56	-0.08	0.509
Parent Used Time Out in Last Week	0.56	0.52	0.04	0.708
Parent Read to Child in Last Week	0.22	0.19	0.03	0.795
Parental Safety Practices Scale	3.50	3.50	0.00	0.991
Family Cultural Enrichment Scale	3.74	4.29	-0.55**	0.043
Parenting Style: Authoritarian	0.11	0.05	0.06	0.367
Parenting Style: Authoritative	0.68	0.73	-0.05	0.671
Parenting Style: Neglectful	0.02	0.05	-0.03	--
Parenting Style: Permissive	0.20	0.17	0.03	0.685

Exhibit F.8: Estimated Impacts on Parenting Outcomes by Year: 3-Year-Old Cohort (continued)

Outcome	Mean Estimates			
	Head Start Group	Control Group	Head Start – Control	p-value
1st Grade Year (Spring 2006)				
Parent-Reported Measures				
Parent Spanked Child in Last Week	0.52	0.61	-0.09	0.460
Parent Used Time Out in Last Week	0.61	0.66	-0.05	0.720
Parent Read to Child in Last Week	0.25	0.30	-0.05	0.680
Family Cultural Enrichment Scale	4.35	4.64	-0.28	0.361
Parenting Style: Authoritarian	0.24	0.03	0.20	--
Parenting Style: Authoritative	0.51	0.85	-0.34***	0.003
Parenting Style: Neglectful	0.04	0.00	0.04	--
Parenting Style: Permissive	0.22	0.12	0.10	0.252

Key:
*** $p < 0.01$
** $p < 0.05$
* $p < 0.10$

Note: p-values were not calculated for Parenting Style: Neglectful in 2004, 2005, and 2006 and Parenting Style: Authoritarian in 2006 due to low response rates which resulted in a standard deviation of 0 for one or both groups (Sudaan Warning).

With regard to the social-emotional outcomes, several impacts were observed for children in the 3-year-old cohort from Puerto Rico: at the end of the Head Start year, parents of children in the Head Start group reported that their children displayed less hyperactive behavior than parents of children in the control group. There were several favorable impacts on the Head Start group children at the end of kindergarten. Parents of children in the Head Start group reported that their children showed less withdrawn behavior, better social competencies, and social skills and positive approaches to learning than parents of children in the control group. In the health domain, children in the Head Start group were less likely to have received care for an injury in the last month than children in the control group at the end of the age 4 year and at the end of kindergarten. Whereas 39 percent of the control group received care for an injury, only 18 percent of the Head Start group received such care at the end of the age 4 year. At the end of kindergarten, the percent dropped to 27 percent for the control group children and 6 percent for the Head Start group children. This finding is unclear and can be interpreted as the Head Start group children have fewer injuries that require care or their parents do not seek care for their injuries. Likewise, the parents of the control group children may be less likely to seek care for

injuries or the control group children may have more injuries. Surprisingly, there was one unfavorable health finding for the 3-year-old cohort: at the end of 1st grade, children who were in the Head Start group were less likely to have received dental care since September than children in the control group. While 94 percent of children in the control group had received dental care since September, only 78 percent of children in the Head Start group did.

Finally, there were several impacts for the parenting outcomes for the 3-year-old cohort children living in Puerto Rico but the findings are mixed. At the end of the Head Start year, parents in the Head Start group reported that they were less likely to spank their children and had engaged in more cultural enrichment activities with their children than did parents in the control group. However, there were two unfavorable impacts on parenting behavior for children in the 3-year-old cohort in Puerto Rico. At the end of kindergarten, parents of children in the Head Start group were less likely to expose their children to cultural enrichment activities (a reversal from the Head Start year) than their peers in the control group. Likewise, at the end of 1st grade, parents in the Head Start group reported that they were less likely to use an authoritative parenting style than parents of children in the control group.

Summary of Main Impacts

Overall, the findings from the estimated impacts for the selected Head Start programs operating on the island of Puerto Rico provide limited evidence of any effect of Head Start on children's cognitive, social-emotional, or health outcomes or the parenting practices or styles of the parents of these children in the 4-year-old cohort. The most notable finding is the favorable impacts found for the children in the 4-year-old cohort at the end of the Head Start year on the CTOPPP and at the end of 1st grade on the Woodcock-Muñoz Identificación de letras y palabras.

However, Head Start does appear to make a difference for the 3-year-old cohort particularly in the cognitive and social-emotional domains. For these children, there is a significant difference between the Head Start group and the control group on the Woodcock-Muñoz Problemas aplicados at the end of the Head Start year, age 4 year, and kindergarten. For all years, the Head Start group scores are higher than the control group scores. Children in the Head Start group score higher than the control group children on a number of language and literacy measures including Color Identification at the end of the Head Start year, the TVIP (adapted), and parent's rating of emerging literacy skills at the end of the age 4 year, and the

TVIP (adapted), Woodcock-Muñoz Identificación de letras y palabras, and CTOPPP at the end of kindergarten. These cognitive findings suggest sustained impacts through the end of kindergarten for the 3-year-old cohort in the Head Start group.

The impact of access to Head Start is also evident in the social-emotional domain. Parents of children in the Head Start group report less hyperactive behavior at the end of the Head Start year and less withdrawn behavior for their children at the end of kindergarten than parents of children in the control group. At the end of kindergarten, parents of the Head Start group children also report better social competencies and social skills and positive approaches to learning than parents in the control group. Although there are findings in the health and parenting practices domains, the results are mixed with both favorable and unfavorable results for the Head Start group children.

Children's Experiences

This section describes the preschool experiences of study children in Puerto Rico. It parallels the chapter on children's experiences on the mainland (i.e., Chapter 3), except that it omits discussion of kindergarten and first grade experiences[2]. As was the case with the mainland children, two cohorts of children were followed ─ children who applied to Head Start for the first time at age four and children who applied for the first time at age three, both in the fall of 2002. Two years of preschool experience are described for the 3-year-old cohort. The analysis addresses questions that were also addressed in the mainland chapters: What difference did random assignment make in whether children actually attended Head Start? How different were their experiences from those of children in the non-Head Start group as a result?

As in the mainland chapter, the data for these analyses come from center director interviews, teacher surveys, care provider interviews (for family child care providers), classroom/child care home observations, and parent interviews. See Chapters 2 and 3 for a discussion of the methodology and measures used.

[2] Kindergarten and 1st grade teacher data was not used. During one data collection period, there was a teacher strike and the data was not collected. During another data collection period the data collected was considered unreliable.

The Impact of Head Start on Children's Preschool Settings in Puerto Rico

Children's focal settings in Puerto Rico were defined the same way as they were for children on the mainland. Exhibit F.9 below is the same as an exhibit in Chapter 3 but is repeated here for convenience in interpreting the Puerto Rico settings.

Exhibit F.9: Definitions of Children's Focal Settings

1. *Head Start*: center-based, home-based, and combination programs funded with Federal Head Start dollars.
2. *Non-Head Start Center*: center-based program as differentiated from child care that takes place in someone's home or in federally funded Head Start classrooms.
3. *Non-Relative's Home*: non-parental care that takes place in a non-relative's home that is not the child's own home. This category includes regulated family child care providers as well as home-based child care providers who are exempt from licensing requirements.
4. *Relative's Home:* non-parental care that takes place in a relative's home that is not the child's own home. This category includes regulated family child care providers who are relatives of the child, as well as home-based relative care providers who are exempt from licensing requirements.
5. *Child's Own Home with a Non-Relative*: non-parental care that takes place in the child's own home by a non-relative of the child. Providers in this category generally are exempt from licensing requirements.
6. *Child's Own Home with a Relative*: non-parental care that takes place in the child's own home by a relative of the child. Providers in this category generally are exempt from licensing requirements.
7. *Parent Care*: care by the child's parent or guardian, typically in the child's own home.

Impact on Preschool Setting: 4-Year-Old Cohort

As shown in Exhibit F.10, having access to Head Start had a highly significant impact on children's focal settings in Puerto Rico. For the 4-year-old cohort in the Head Start group, more than 90 percent were in Head Start during the spring of their Head Start year. Among children in the control group, the largest proportion, almost 75 percent, were in parent care, while another 4 percent were cared for in the home of a relative. A small proportion of children in both the Head Start group and the control group (6% and 7% respectively) attended non-Head Start centers.

Exhibit F.10: Percentage of Children by Focal Setting, 4-Year-Old Cohort, Spring 2003

Focal Setting	Head Start Group	Control Group	Magnitude of Impact***
Head Start	90.2	14.9	75.3
Non-Head Start center	5.5	6.5	-1.0
Home of non-relative	0.0	0.0	0.0
Home of relative	0.0	4.1	-4.1
Own home with relative	0.0	0.0	0.0
Own home with non-relative	0.0	0.0	0.0
Parent Care	4.3	74.5	-70.2
Total	100%	100%	

*p≤0.05, **p≤0.01, ***p≤0.001

The parent interviews also obtained information on how long children had been in the Head Start programs. As shown in Exhibit F.11, almost 100 percent of the Head Start group children in Puerto Rico who were enrolled in Head Start attended the program for at least eight months, indicating that they had participated for the entire school year by the time their parents were interviewed.

Exhibit F.11: Percentage of Head Start Group Children Who Enrolled in Head Start by Months of Participation, 4-Year-Old Cohort, Spring 2003

Months in Head Start as of Spring 2003	Percentage
≤4	0.0
5	0.0
6	0.0
7	9.0
8	0.53
9	46.18
10	48.77
Total (all treatment children in Head Start)	100%
Mean	9.43 months

Impact on Preschool Setting: 3-Year-Old Cohort

The 3-year-old cohort in Puerto Rico, as on the mainland, had 2 years in which to attend preschool or child care before entering kindergarten. As shown in Exhibit F.12, having access to Head Start had a significant impact on focal setting when this cohort was 3 years old: 84 percent were in Head Start during that year (2002-03). Among children in the control group, the largest proportion was in parent care (55%), followed by non-Head Start centers (19%).

Exhibit F.12: Percentage of Children by Focal Setting, 3-Year-Old Cohort, Spring 2003

Focal Setting	Head Start Group	Control Group	Magnitude of Impact**
Head Start	83.6	11.6	72.0
Non-Head Start center	5.8	19.3	-13.5
Home of non-relative	0.0	0.0	0.0
Home of relative	2.8	11.0	-8.2
Own home with relative	0.0	3.1	-3.1
Own home with non-relative	0.0	0.0	0.0
Parent Care	7.8	55.1	-47.3
Total	100%	100%	

*p≤0.05, **p≤0.01, ***p≤0.001

All the children in the 3-year-old Head Start group who enrolled in Head Start attended the program for at least eight months, as shown in Exhibit F.13.

Exhibit F.13: Percentage of Head Start Group Children by Months in Head Start, 3-Year-Old Cohort, Spring 2003

Months in Head Start as of Spring 2003	Percentage
≤3	0.0
4	0.0
5	0.0
6	0.0
7	0.0
8	4.50
9	64.88
10	30.62
Total (all treatment children in Head Start)	100%
Mean	9.26 months

The 3-year-old cohort's second year (2003-2004) of preschool experiences was very different from their first year. Control group children were no longer prohibited from enrolling in Head Start during this second year, and, indeed, a large proportion of the children left parent care and went into Head Start. As shown in Exhibit F.14, there were no significant differences between the Head Start group and the control group in their enrollment for the second year of Head Start. Over 70% of children in both the Head Start and the control groups enrolled in Head Start.

Exhibit F.14: Percentage of Children by Focal Setting, 3-Year-Old Cohort, Spring 2004

Focal Setting	Program	Control	Magnitude of Impact
Head Start	76.5	71.5	5.0
Non-Head Start center	17.7	14.9	2.8
Home of non-relative	0.0	0.0	0.0
Home of relative	0.0	0.0	0.0
Own home with relative	0.0	0.0	0.0
Own home with non-relative	0.0	0.0	0.0
Parent Care	2.1	8.5	-6.4
Kindergarten	3.7	5.2	-1.5
Missing/not ascertained	0.0	0.0	0.0
Total	100%	100%	

*p≤0.05, **p≤0.01, ***p≤0.001

As in the first year, the children enrolled in Head Start attended for eight months or more on average. The Head Start group averaged 9.3 months of Head Start and the non-Head Start group averaged 8.5 months of Head Start.

Characteristics of Early Childhood Settings in Puerto Rico

This section describes the preschool environments of the children in Puerto Rico. It begins with a description of the experiences of children in the 4-year-old cohort (who had a single Head Start year) and then describes the experiences of children in the 3-year-old cohort (who typically had two Head Start years). The exhibits in this section follow the same format as those for the mainland children's experiences, with two approaches to presenting the findings: (1) a measure of the impact of Head Start, which includes all children regardless of their focal

setting, and (2) a description of what the children experienced, which excludes children who could not have had the experiences because of their focal setting. Because the second approach incorporates non-randomized sets of children in both study groups, the observed differences do not represent the impact of access to Head Start and are provided only for descriptive purposes. Unlike the analysis of focal settings in the previous section, no significance testing was conducted on the Puerto Rico data presented in this section because the necessary replicate weights were not developed.

Characteristics of Early Childhood Settings: 4-Year-Old Cohort

As shown in the "Impact" columns of Exhibit F.15, there are differences between the Head Start group and the control group on almost every measure of children's preschool experiences, and in most cases, the magnitude of the differences is quite large (e.g., 50-70 percentage points). This finding is driven primarily by the large proportion of children in parent care; when those children are excluded from the analysis (shown in the shaded columns), many of the differences are far smaller. But for some of the quality indicators, as discussed below, the Head Start group generally had higher scores, even when children in parent or own home care are excluded.

Some of the more important differences in preschool experiences between the Head Start group and the control group are summarized below. These highlight areas where there were relatively large differences in both the impact columns and the descriptive columns, although the differences could not be tested for statistical significance. The findings in Exhibit F.15 show that, compared to the control group, children in the Head Start group were in settings:

- Where they had a variety of language/literacy and math instructional activities at least three times a week;
- Where parents participated in setting activities, the center provided monthly teacher training, and the family received a home visit from their children's care settings;
- That offered a variety of services for children including health services, hearing and vision screening, mental health services, and nutrition;
- That offered a variety of services for families, including job assistance, adult education, counseling, and other types of services;
- Where their teacher had a bachelor's degree and had early childhood education courses in college;

Exhibit F.15: Percentage of Children by Their Settings' Characteristics, 4-Year-Old Cohort, Spring 2003

Characteristic	Impact (All children)			Description (Only children in non-parental care)	
	Head Start Group (% YES)	Control Group (% YES)	Magnitude of Impact	Head Start Group (% YES)	Control Group (% YES)
Parent or own home care	4.2	77.2	-73.0		
Center Environment and Characteristics					
Qualifications: Director has at least a bachelor's degree	92.2	22.1	70.1	96.2	96.9
Qualifications: Director has been in current position at least 4 years	28.8	11.0	17.8	30.1	48.3
Training: Center provides teacher training at least monthly	82.4	12.6	69.8	86.1	55.4
Training: Center provides teacher mentoring	34.3	7.7	26.6	35.9	34.0
Turnover: Center has low proportion of new lead teachers ($\leq$ 20%)	73.5	19.9	53.6	76.8	87.2
Center size > 50	1.8	2.4	-0.6	1.8	10.5
Competition from other preschools:					
Lots	1.8	7.7	-5.9	1.9	33.7
Some	5.5	0.0	5.5	5.8	0.0
Little or none	88.4	15.1	-73.3	92.4	66.3
Center always is filled to capacity	75.0	10.2	64.8	78.3	44.9
Center is affiliated with a school	5.0	0.0	5.0	5.2	0.0
Center uses curriculum	91.9	22.8	69.1	96.0	100.0
Services Available for Children:					
Hearing/vision screening/ referrals	94.5	12.7	81.8	98.7	53.2
Mental health services	88.4	12.7	75.7	92.3	53.2
Health Services	88.4	10.3	78.1	92.3	43.1
Nutrition services	92.6	12.7	79.9	96.6	53.2
Center provides transportation	23.6	3.5	20.1	24.7	15.5
Services Available for Families:					
Job training/employment assistance	79.7	12.7	67.0	83.2	53.2
Adult education/literacy	82.1	12.7	69.4	85.7	53.2
Family counseling or mental health services	86.4	12.7	73.7	90.2	53.2
Help with dealing with family violence	89.7	12.7	77.0	93.6	53.2
Help with housing	81.1	12.0	69.1	84.7	50.2
Help with utilities	21.1	2.4	18.7	22.1	10.1

Exhibit F.15: Percentage of Children by Their Settings' Characteristics, 4-Year-Old Cohort, Spring 2003 (continued)

Characteristic	Impact (All children)			Description (Only children in non-parental care)	
	Head Start Group (% YES)	Control Group (% YES)	Magnitude of Impact	Head Start Group (% YES)	Control Group (% YES)
Services Available for Families: (cont'd)					
Help with medical care	82.4	12.7	69.7	86.0	53.2
Family received home visit from focal setting	33.7	0.7	33.0	35.2	2.6
Alcohol/drug abuse treatment or counseling	85.1	12.7	72.4	88.9	53.2
Food and nutrition assistance	64.1	9.0	55.1	66.9	37.9
Income assistance	58.4	8.4	50.0	60.9	35.4
Foster care program	12.5	2.4	10.1	13	10.1
Other	8.3	0.0	8.3	8.7	0.0
Teacher/Care Provider Qualifications and Training					
Had college ECE courses or obtained CDA	89.8	12.7	77.1	93.8	48.7
Obtained CDA (with/without college ECE courses)	15.9	0.7	15.2	16.6	2.7
Highest educational attainment was associate's degree	3.0	3.3	-0.3	3.1	12.7
Highest educational attainment was bachelor's degree or higher	91.9	16.8	75.1	95.9	64.2
Attained bachelor's degree or higher in ECE *(subset of previous row)*	70.9	12.0	58.9	74.0	46.0
Received at least 25 hours of training in past year	45.1	9.5		47.1	36.5
Receives mentoring at least once/ month	24.9	2.9	22.0	26.0	12.8
Classroom Environment					
Had average ECERS-R/FDCRS rating of at least 5 out of 7	35.2	3.2	32.0	36.5	10.5
Had average ECERS-R/FDCRS rating of at least 6 out of 7	0.0	0.0	0.0	0.0	0.0
Had highest average ECERS-R/ FDCRS rating (7)	0.0	0.0	0.0	0.0	0.0
Had average Arnett rating of at least 3 out of 4	96.2	30.6	65.6	100.0	100.0
Had highest average Arnett rating (4)	0.0	0.0	0.0	0.0	0.0
Met child/staff ratio standard	62.7	10.5	52.2	65.2	34.4

Exhibit F.15: Percentage of Children by Their Settings' Characteristics, 4-Year-Old Cohort, Spring 2003 (continued)

Characteristic	Impact (All children in each group)			Description (Only children in non-parental care)	
	Head Start Group (% YES)	Control (% YES)	Magnitude of Impact	Head Start Group (% YES)	Control (% YES)
Classroom Environment (cont'd)					
Parent participated in setting activity at least once	95.7	19.2	76.5	100.0	75.2
Classroom Activities					
Provides at least 3 of 4 activities other than literacy and math at least 3 times/week	81.4	21.4	60.0	85.0	81.8
Provides at least 7 of 12 literacy activities at least 3 times/week	75.3	16.8	58.5	78.7	64.3
Provides at least 5 of 8 math activities at least 3 times/week	86.6	17.5	69.1	90.5	66.9
Overall Quality Composite					
At or above overall mean	70.0	9.5	60.5	73.1	36.5

- Where the center had little or no competition from other preschools and was always filled to capacity;
- That met the child/staff ratio standard; and
- That had indicators of higher quality, including higher ECERS-R and quality composite scores.

Characteristics of Early Childhood Settings: 3-Year-Old Cohort

The chapter now turns to the experiences of the 3-year-old cohort in two years of preschool (2002-2003 and 2003-2004) in Puerto Rico. The impacts estimated for the 3-year-old cohort represent the effects of access to Head Start for two years prior to entering kindergarten. The children who were randomized into the non-Head Start group (and not eligible for Head Start during the first year of the study) were allowed to enroll in Head Start the following year.

The Head Start Year: 3-Year-Old Cohort

The findings for the 3-year-old cohort's Head Start year are very similar to those discussed above for the 4-year-old cohort. As shown in the "Magnitude of Impact" columns of

Exhibit F.16, there are differences between the Head Start group and the control group on every measure of children's preschool experiences, and in most cases, the magnitude of the differences is quite large (e.g., 20-60 percentage points). These differences are driven by the large difference between the Head Start group and the control group in the percentages of children in exclusively parental care, although this difference is not as large as for the 4-year-old cohort.

Some of the more important differences in preschool experiences between the Head Start group and the control group are summarized below. These highlight areas where there were relatively large differences in both the impact columns and the descriptive columns, although the differences could not be tested for statistical significance. The findings in Exhibit F.16 show that, compared to the control group, children in the Head Start group were more likely to be in settings:

- Where math and literacy activities were provided, parents participated in setting activities, transportation was provided, teachers received monthly training, and families received a home visit from the child care setting;
- That offered a variety of services for children, including health services, hearing and vision screening, mental health services, and nutrition;
- That offered a variety of services for families, including job assistance, adult education, counseling, and other types of services;
- Where their teachers had bachelor's degrees and college ECE courses;
- Where the center had little or no competition from other preschools and was always filled to capacity; and
- That had indicators of higher quality, including higher ECERS-R, Arnett, and quality composite scores.

The Age 4 Year: 3-Year-Old Cohort

Like their counterparts on the mainland, the Puerto Rican 3-year-old cohort's preschool experiences in the second year were very different from their experiences in the first year. As shown in Exhibit F.14, most of these children were in some type of center-based care in the second year (94.2% for children in the Head Start group and 86.4% for children in the control group), and there was much less difference between the Head Start group and the control group on the characteristics of their care settings. Control group children were allowed to go to Head Start for their age 4 year, and about 72 percent opted to do so. This rate of Head

Exhibit F.16: Percentage of Children by Their Settings' Characteristics, 3-Year-Old Cohort, Head Start Year, Spring 2003

Characteristic	Impact (All children			Description (Only children in non-parental care)	
	Head Start Group (% YES)	Control Group (% YES)	Magnitude of Impact	Head Start Group (% YES)	Control Group (% YES)
Parent or own home care	8.6	54.3	-45.7		
Center Environment and Characteristics					
Qualifications: Director has at least a bachelor's degree	72.7	28.1	44.6	82.2	81.1
Qualifications: Director has been in current position at least 4 years	25.2	8.9	16.3	28.6	25.7
Training: Center provides teacher training at least monthly	63.0	12.6	50.4	71.2	36.3
Training: Center provides teacher mentoring	39.8	16.3	23.5	45.0	46.8
Turnover: Center has low proportion of new lead teachers ($\leq$ 20%)	66.1	25.0	41.1	74.7	72.0
Center size > 50	5.5	12.9	-7.4	6.2	37.1
Competition from other preschools:					
Lots	3.6	8.2	-4.6	4.1	23.6
Some	4.3	0.0	4.3	4.8	0.0
Little or none	77.6	26.5	51.1	87.8	76.4
Center always is filled to capacity	72.2	8.9	63.3	81.7	25.7
Center is affiliated with a school	0.0	10.0	-10.0	0.0	28.9
Center uses curriculum	85.5	31.5	54.0	96.7	90.9
Services Available for Children:					
Hearing/vision screening/ referrals	82.5	18.6	63.9	90.4	43.2
Mental health services	80.0	15.1	64.9	87.7	35.1
Health services	80.0	18.6	61.4	87.7	43.2
Nutrition services	80.0	18.6	61.4	87.7	43.2
Center provides transportation	21.0	0.0	21.0	23.7	0.0
Services Available for Families:					
Job training/employment assistance	74.6	5.8	68.8	81.7	13.6
Adult education/literacy	72.4	5.8	66.6	79.3	13.6
Family counseling or mental health services	74.6	13.7	60.9	81.7	31.9
Help with dealing with family violence	80.0	12.1	67.9	87.7	28.1
Help with housing	66.4	10.1	56.3	72.8	23.5
Help with utilities	37.6	0.0	37.6	41.1	0.0

Exhibit F.16: Percentage of Children by Their Settings' Characteristics, 3-Year-Old Cohort, Head Start Year, Spring 2003 (continued)

Characteristic	Impact (All children)			Description (Only children in non-parental care)	
	Head Start Group (% YES)	Control Group (% YES)	Magnitude of Impact	Head Start Group (% YES)	Control Group (% YES)
Services Available for Families: (cont'd)					
Help with medical care	67.4	5.8	61.6	73.8	13.6
Family received home visit from focal setting	31.7	4.9	26.8	34.4	11.7
Alcohol/drug abuse treatment or counseling	74.6	8.8	65.8	81.7	20.5
Food and nutrition assistance	61.9	11.5	50.4	67.8	26.8
Income assistance	55.1	10.4	44.7	60.4	24.0
Foster care program	2.0	0.0	2.0	2.2	0.0
Other	6.8	15.1	-8.3	7.4	35.1
Teacher/Care Provider Qualifications and Training					
Had college ECE courses or obtained CDA	72.7	18.9	53.8	79.3	48.0
Obtained CDA (with/without college ECE courses)	12.0	6.2	5.8	13.1	13.5
Highest educational attainment was associate's degree	15.7	10.0	5.7	17.2	22.0
Highest educational attainment was bachelor's degree or higher	68.7	15.1	53.6	75.0	33.0
Attained bachelor's degree or higher in ECE *(subset of previous row)*	45.4	5.7	39.7	49.5	12.5
Received at least 25 hours of training in past year	24.9	3.8	21.1	27.3	8.3
Receives mentoring at least once/ month	45.4	16.2	29.2	51.3	46.6
Classroom Environment					
Had average ECERS-R/FDCRS rating of at least 5 out of 7	35.3	5.8	29.5	38.7	12.0
Had average ECERS-R/FDCRS rating of at least 6 out of 7	0.0	0.0	0.0	0.0	0.0
Had highest average ECERS-R/ FDCRS rating (7)	0.0	0.0	0.0	0.0	0.0
Had average Arnett rating of at least 3 out of 4	85.3	36.0	49.3	93.5	74.9
Had highest average Arnett rating (4)	0.0	0.0	0.0	0.0	0.0
Met child/staff ratio standard	58.5	30.9	27.6	64.2	64.2

Exhibit F.16: Percentage of Children by Their Settings' Characteristics, 3-Year-Old Cohort, Head Start Year, Spring 2003 (continued)

Characteristic	Impact (All children in each group)			Description (Only children in non-parental care)	
	Head Start Group (% YES)	Control Group (% YES)	Magnitude of Impact	Head Start Group (% YES)	Control Group (% YES)
Classroom Environment (cont'd)					
Parent participated in setting activity at least once	84.3	29.6	54.7	91.5	70.7
Classroom Activities					
Provides at least 3 of 4 activities other than literacy and math at least 3 times/week	80.3	39.1	41.2	87.6	85.6
Provides at least 7 of 12 literacy activities at least 3 times/week	58.5	14.9	43.6	63.8	32.7
Provides at least 5 of 8 math activities at least 3 times/week	70.3	41.8	28.5	76.7	62.3
Overall Quality Composite					
At or above overall mean	75.2	18.8	56.4	82.1	41.0

Start enrollment is close to the 76.5 percent rate for the Head Start group. With a larger number of control group children in Head Start or other center-based care for their second year, the control group children had much more access to services than they did in their first year, as can be seen in Exhibit F.17. Also, the measures of the quality and characteristics of care received by the Head Start group and the control group show fewer differences in the second year as compared to the first year (which was presented in Exhibit F.16).

As shown in Exhibit F.17, differences between the Head Start group and the control group were small, and in some cases favored children in the control group. For example, children in the control group were more likely to attend centers that offered services to families such as job assistance, adult education, and counseling, and that had a director who had attained at least a bachelor's degree and had been the director for at least 4 years. The control group's settings also were somewhat more likely to score above the mean on the quality composite.

Exhibit F.17: Percentage of Children by Their Settings' Characteristics, 3-Year-Old Cohort, Age 4 Year, Spring 2004

Characteristic	Impact (All children in each group)			Description (Only children in non-parental care)	
	Head Start Group (% YES)	Control Group (% YES)	Magnitude of Impact	Head Start Group (% YES)	Control Group (% YES)
Parent care or own home	0.7	4.4	-3.7		
Center Environment and Characteristics					
Qualifications: Director has at least bachelor's degree	71.2	49.8	21.4	71.7	52.1
Qualifications: Director has been in current position at least 4 years	56.8	47.6	9.2	57.2	49.8
Training: Center provides teacher training at least monthly	2.2	0.0	2.2	2.3	0.0
Turnover: Center has low proportion of new lead teachers ($\leq$ 20%)	30.2	21.9	8.3	30.4	22.9
Center size > 50	23.6	29.7	-6.1	23.8	31.1
Competition from other preschools:					
Lots	7.0	0.0	7.0	7.0	0.0
Some	5.3	0.0	5.3	5.4	0.0
Little or none	68.9	62.7	6.2	69.4	65.7
Center always is filled to capacity	64.4	53.2	11.2	64.9	55.6
Center is affiliated with a school	0.0	0.0	0.0	0.0	0.0
Center uses curriculum	77.0	62.7	14.3	77.5	65.7
Center provides teacher mentoring	58.8	47.3	11.5	59.2	49.5
Services Available for Children:					
Hearing/vision screening/referrals	86.3	93.3	-7.0	87.6	100.0
Mental health services	86.3	93.3	-7.0	87.6	100.0
Health services	81.8	93.3	-11.5	83.0	100.0
Nutrition services	89.7	93.3	-3.6	91.0	100.0
Center provides transportation	4.6	0.0	4.6	4.7	0.0
Services Available for Families					
Job training/employment assistance	58.5	81.6	-23.1	59.4	87.4
Adult education/literacy	65.4	89.2	-23.8	66.4	95.6
Family counseling or mental health services	64.9	76.4	-11.5	65.9	81.8
Help dealing with family violence	69.8	84.9	-15.1	70.8	91.0
Help with housing	56.8	62.6	-5.8	57.6	67.1
Help with utilities	5.0	6.0	-1.0	5.0	6.5
Help with medical care	59.4	71.2	-11.8	60.3	76.3

Exhibit F.17: Percentage of Children by Their Settings' Characteristics, 3-Year-Old Cohort, Age 4 Year, Spring 2004 (continued)

Characteristic	Impact (All children in each group)			Description (Only children in non-parental care)	
	Head Start Group (% YES)	Control Group (% YES)	Magnitude of Impact	Head Start Group (% YES)	Control Group (% YES)
Services Available for Families: (cont'd)					
Alcohol/drug abuse treatment or counseling	54.7	79.0	-24.3	55.6	84.7
Food and nutrition assistance	47.8	60.2	-12.4	48.5	64.5
Income assistance	15.8	12.1	3.7	16	12.9
Foster care payments	6.7	0.0	6.7	6.8	0.0
Other	0.0	0.0	0.0	0.0	0.0
Teacher/Care Provider Qualifications and Training					
Had college ECE courses or obtained CDA	85.6	92.6	-7.0	86.2	96.9
Obtained CDA (with or without college ECE courses)	20.1	7.8	12.3	20.3	8.2
Received at least 25 hours of training in past year	21.9	18.2	3.7	22.1	19.0
Highest educational attainment was associate's degree	76.1	83.5	-7.4	76.6	87.4
Receives mentoring at least once/ month	35.5	31.5	4.0	35.7	32.9
Highest educational attainment was bachelor's degree or higher	0.0	0.0	0.0	0.0	0.0
Attained bachelor's degree or higher in ECE *(subset of previous row)*	0.0	0.0	0.0	0.0	0.0
Classroom Environment					
Had average ECERS-R/FDCRS rating of at least 5 out of 7	66.8	73.4	-6.6	70.5	84.7
Had average ECERS-R/FDCRS rating of at least 6 out of 7	42.4	55.6	-13.2	44.7	64.2
Had highest average ECERS-R/ FDCRS rating (7)	6.6	12.4	-5.8	7.0	14.3
Had average Arnett rating of at least 3 out of 4	90.6	83.6	7.0	95.6	96.5
Had highest average Arnett rating (4)	15.7	30.7	-15.0	16.6	35.4
Met child/staff ratio standard	59.7	65.3	-5.6	63.0	75.4
Parent participated in setting activity at least once	87.0	83.4	3.6	88.9	91.1

Exhibit F.17: Percentage of Children by Their Settings' Characteristics, 3-Year-Old Cohort, Age 4 Year, Spring 2004 (continued)

Characteristic	Impact (All children in each group)			Description (Only children in non-parental care)	
	Head Start Group (% YES)	Control Group (% YES)	Magnitude of Impact	Head Start Group (% YES)	Control Group (% YES)
Classroom Activities					
Provides at least 3 of 4 activities other than literacy and math at least 3 times/week	85.0	80.0	5.0	85.7	83.8
Provides at least 5 of 8 math activities at least 3 times/week	78.0	82.0	-4.0	78.5	85.8
Provides at least 7 of 12 literacy activities at least 3 times/week	80.3	91.6	-11.3	80.9	95.8
Overall Quality Composite:					
At or above overall mean	60.0	68.6	-8.6	60.4	71.8

Summary

The following is a summary of the findings for Puerto Rico:

4-Year-Old Cohort

There were very few statistically significant impacts across any of the domains measured for the sample of 4-year-old children in Puerto Rico.

- **Cognitive, Social-Emotional, Health, and Parenting Practices Outcomes**
  - Children in the Head Start group scored higher on the CTOPPP Elision at the end of the Head Start year and the Woodcock-Muñoz Identificación de letras y palabras at the end of 1st grade than children in the control group.
  - Parents of children in the Head Start group reported more conflict between themselves and their children at the end of kindergarten than did parents of children in the control group. Additionally parents of children in the Head Start group reported less closeness in the relationship with their children at the end of 1st grade than did parents of children in the control group.
  - No other impacts were found on measures of cognitive development, social-emotional development, health outcomes and services, or parenting practices in any years.

- **Children's Experiences**
  - Children in the Head Start group were significantly more likely to be in Head Start (90%), while children in the control group were more likely to be in parent care (75%).
  - The care settings for children in the Head Start group were more likely to provide literacy and math activities, teacher training, home visits, and services for children and families.
  - The teachers of children in the Head Start group were more likely to have a bachelor's degree and early ECE courses in college.
  - The care settings for the children in the Head Start group were more likely to meet the child/staff ratio standard and score higher on quality indicators.

3-Year-Old Cohort

Impacts on the 3-year-old cohort were more common than for the 4-year-old cohort, and there is some evidence of sustained impacts in kindergarten on this group in several domains.

- **Cognitive Outcomes**
  - Children in the Head Start group had higher scores on the Woodcock-Muñoz Problemas aplicados test, an assessment of early math skills, at the end of their Head Start year than did children in the control group. Children in the Head Start group also score higher on the Color Identification task at this time than did children in the control group.
  - At the end of the age 4 year, children in the Head Start group continued to show higher scores on the Woodcock-Muñoz Problemas aplicados than their counterparts in the control group. During the same time period, children in the Head Start group had higher scores on the TVIP, a measure of receptive vocabulary, than did children in the control group and the parents of children in the Head Start group reported higher emergent literacy scores.
  - Likewise, at the end of kindergarten, children in the Head Start group outperformed their peers in the control group on four different direct assessments of cognitive skills: TVIP (adapted), Woodcock-Muñoz Identificación de letras y palabras, CTOPP Elision , and the Woodcock-Muñoz Problemas aplicados.
- **Social-Emotional Outcomes**
  - At the end of the Head Start year, parents of children in the Head Start group reported less hyperactive behavior among their children than did parents of children in the control group.
  - Likewise, there were several favorable impacts on the Head Start group at the end of kindergarten. Parents of children in the Head Start group reported their children to be less withdrawn, have better social competencies, and social skills and positive approaches to learning.

- **Health Outcomes**
  - At the end of the age 4 year and kindergarten, parents of children in the Head Start group reported less treatment for their child's injuries in the last month than did parents of children in the non-Head Start group.
  - Surprisingly, at the end of 1st grade, parents of children in the Head Start group reported that their children were *less* likely to have received dental care since September than parents of children in the control group.
- **Parenting Outcomes**
  - At the end of the Head Start year, parents in the Head Start group reported being less likely to spank their children and more likely to have exposed them to cultural enrichment activities than parents of children in the control group.
  - However, in later years, the impacts on parenting practices outcomes were less favorable. By the end of kindergarten, parents in the Head Start group reported being *less* likely to have exposed their children to cultural enrichment activities than parents of children in the control group. Likewise, at the end of 1st grade, parents in the Head Start group were less likely to use an authoritative parenting style than parents in the control group.
- **Children's Experiences**
  - During the Head Start year, children in the Head Start group were significantly more likely to be in Head Start (84%), while children in the control group were more likely to be in parent care (55%).
  - During the age 4 year, when the control group was allowed to enroll in Head Start, there were no statistically significant differences in focal settings: 77 percent of the Head Start group and 72 percent of the control group enrolled in Head Start, and the proportion of the control group in parent care dropped to 9 percent.
  - During the Head Start year, children in the Head Start group were more likely to experience literacy and math activities, parent participation in care settings, and home visits.
  - During the Head Start year, the settings of children in the Head Start group were more likely to offer services for children and families, provide monthly teacher training, meet the child/staff ratio standard, and achieve indicators of higher quality.
  - During the age 4 year, differences between the Head Start group and the control group were small and, in some cases, favored the control group, such as for services to families.