Building Futures
The Head Start Impact Study
Interim Report

September 2003

Executive Summary

Introduction

The Head Start program, which provides comprehensive early childhood development services to low-income children and their families, has experienced significant growth over the last decade, especially as greater attention has been paid to the need for early intervention. During this period, the U.S. General Accounting Office (GAO) released two reports underlining the lack of rigorous research on Head Start's effectiveness, i.e., ". . .*the body of research on current Head Start is insufficient to draw conclusions about the impact of the national program,*"[1] and, as a consequence, "*. . .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*."[2]

Based upon the GAO recommendation, and the testimony of research methodologists and early childhood experts, Congress mandated through the 1998 reauthorization of Head Start[3] that the Department of Health and Human Services determine, on a national level, the impact of Head Start on the children it serves.[4] In October 2000, DHHS awarded a contract to Westat in collaboration with The Urban Institute, the American Institutes for Research, and Decision Information Resources to conduct the Head Start Impact Study.

The design of the study is built upon the report of the Advisory Committee on Head Start Research, *Evaluating Head Start: A Recommended Framework for Studying the Impact of the Head Start Program*. This report set forth a framework for research on the impact of Head Start that is both scientifically credible and feasible. The report also noted that the legislative mandate clearly requires that the impact study must address two main questions. First, what difference does Head Start make to the key outcomes of development and learning (and, in particular, the multiple domains of school readiness) of the nation's low-income children? Second, under what circumstances does Head Start work best and for which children? In addition, the Committee acknowledged that the legislative language recommended the use of a rigorous methodology,

---

[1] U.S. General Accounting Office (1997). ***Head Start: Research Provides Little Information on Impact of Current Program***. Washington DC: Author.
[2] U.S. General Accounting Office (1998). ***Head Start: Challenges in Monitoring Program Quality and Demonstrating Results***. Washington DC: Author.
[3] Section 649(g) of the Head Start Act, as amended by the Coats Human Services Reauthorization Act of 1998 (P.L. 105-285)

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. The end result is that the Head Start Impact Study is one of a small handful of randomized, nationally representative studies ever conducted.

Study Overview

The Head Start Impact Study is a longitudinal study involving approximately 5,000 3- and 4-year-old newly entering children who were eligible for Head Start and who applied in fall 2002 to 383 centers across 84 nationally representative grantee/delegate agencies.[5] The participating children have been randomly assigned to either a Head Start group (that receives Head Start services) or to a control group (that does not receive Head Start services but may enroll in other available services selected by their parents or may be cared for at home). Every effort was made to minimize the burden on individual programs and not to significantly change typical enrollment and recruitment procedures.

Children enrolled in Early Head Start, Migrant Head Start and programs operated by Tribal Organizations, those considered extremely new (i.e., in operation less than 2 years), and those considered severely out of compliance were not included in the study.

Great care was taken to include only programs that were not currently serving all of the eligible children in their community. This was to ensure the availability of a sufficient number of unserved, eligible children available who could be randomly assigned to a control group, without causing the program to serve any fewer children than would otherwise be the case. These "saturation" determinations were based on grantee/delegate agencies' own reports of enrollment levels in fall 2001, along with other available information.

Data Collection

Data collection began in fall 2002 and is scheduled to continue through 2006, following children through the spring of their 1st grade year. It includes in-person interviews with parents,

---

[4] See Appendix A for the research-related amendments to the Head Start Act included in the 1998 reauthorization.

[5] The study sample includes both Head Start grantees and their delegate agencies. Grantees are organizations that have fiscal and administrative responsibility for programs in their jurisdiction. In some cases, they can subcontract with agencies to handle administrative oversight over some or all of these programs. Throughout this report we use the term grantee/delegate agencies to refer to both types of agencies.

in-person child assessments conducted twice during the first year of the study and annually afterwards, annual surveys with care providers and teachers, direct observations of the quality of different care settings, and teacher ratings of individual study children. The data to be collected include the following:

- Individual child data in areas related to school readiness, such as approaches to learning, language usage and emergent literacy, cognition and general knowledge, physical well-being and motor development, and social and emotional development.
- Information pertaining to: parenting practices, including parents' descriptions of the types of literacy activities they engage in with children at home; family resources and risk factors; and family demographic and socio-economic characteristics.
- Information on the structure, process, and quality of Head Start, child care, and school settings through first grade, including teachers' reports on their credentials and experience and structured observations of the quality of different care settings, including assessments of classroom resources and instructional practices.
- Community-level data relating to the availability of formal and informal family support services.

Status Update and Reports

The Head Start Impact Study began in October 2000 with initial activities focused on developing the study design, recruiting Head Start grantee/delegate agencies, and field testing study measures and procedures. The selection and recruitment of participating Head Start centers, families, and children continued through spring 2002, culminating with the collection of child, family, and program data in fall 2002 and spring 2003. Activities are on schedule for the planned final report of findings in December 2006. Accomplishments to date include:

- During the 2000-01 Head Start year, conducted an advance field test involving eight Head Start grantees, 21 Head Start centers, and approximately 400 children to test the feasibility of all study procedures.
- During the 2001-02 Head Start year, completed the selection and recruitment of a national probability sample of 84 grantee/delegate agencies and 378 Head Start centers to participate in the national study.
- During summer 2002, completed random assignment, in all recruited centers, resulting in 4,750 children being randomly assigned to either a Head Start treatment group or a non-Head Start control group.
- Completed fall 2002 data collection, with high overall response rates for both child assessments and parent interviews with a combined rate of 80 percent for fall 2002.
- Completion of data entry and checking for errors, as well as statistical weighting of the fall 2002 data. The initial assessment of the psychometric properties of the child assessments is currently underway.
- Completion of spring 2003 data collection with approximately 83 percent response rates for both child assessments and parent interviews.

- Have started data entry and checking for errors, as well as statistical weighting. Initial assessment of psychometric properties for the spring 2003 data collection will continue on an ongoing basis following the completion of this process for the fall 2002 data.
- A meeting was held of the Advisory Committee on Head Start Research on June 16 and 17, 2003, to update the committee on the progress of the study activities and to seek additional input and recommendations on the remaining work of the study. A summary of the Advisory Committee meeting can be found at http://www.acf.hhs.gov/programs/hsb/research/hsreac/index.htm

Overall, Advisory Committee members commented positively on the research design of the Impact Study and the success of implementation to date, especially given the complexity of the Congressional mandate and the ethical concerns of randomly assigning low-income children to a control group. Several members noted that when completed, the Impact Study will provide a rich array of data that will not only meet the Congressional requirements, but will also allow for an extensive range of secondary data analyses.

A final report of the study findings is scheduled for December 2006.

Table of Contents

Chapter 1: Introduction

Introduction

Section 649(g) of the Head Start Act, as amended by the Coats Human Services Reauthorization Act of 1998 (P.L. 105-285) required the Department of Health and Human Services (DHHS) to conduct a national analysis of the impact of Head Start. The legislation also charged the Secretary to appoint an independent panel of experts to review and make recommendations on the design of a plan for research on the impact of Head Start within one year after the date of enactment of P.L. 105-285, to advise the Secretary regarding the progress of the research, and to comment on the interim and final research reports**.** Three reports to Congress were mandated in the legislation. The first required interim report, summarizing the deliberations and recommendations of the Advisory Committee, was completed and transmitted to Congress in October 1999. The second report to Congress, describing the background and purposes of the study, the progress made to date in implementing the study, and the activities undertaken to conduct a field test and prepare for the full study, was submitted in June 2002. This third mandated report describes the status of the study activities to date, including preliminary information from the fall 2002 baseline data collection.

Background

The Head Start program provides comprehensive early childhood development services to low-income children and their families. Over the last decade the program has experienced significant growth, particularly as greater attention has been paid to the need for early intervention in the lives of low-income children. During this period of growth, the U.S. General Accounting Office (GAO) released two reports underlining the lack of rigorous research on Head Start's effectiveness, i.e., "*. . .the body of research on current Head Start is insufficient to draw conclusions about the impact of the national program*,"[6] and, as a consequence, "*. . .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*."[7]

[6] U.S. General Accounting Office (1997). ***Head Start: Research Provides Little Information on Impact of Current Program***. Washington DC: Author.
[7] U.S. General Accounting Office (1998). ***Head Start: Challenges in Monitoring Program Quality and Demonstrating Results***. Washington DC: Author.

Based upon the GAO recommendation, and the testimony of research methodologists and early childhood experts, Congress mandated through the 1998 reauthorization of Head Start that DHHS determine, on a national level, the impact of Head Start on the children it serves.[8] In October 2000, DHHS awarded a contract to Westat in collaboration with The Urban Institute, the American Institutes for Research, and Decision Information Resources to conduct this research study.

Given the very specific legislative mandate, along with input from the Advisory Committee on Head Start Research, DHHS has implemented the most rigorous and scientific study, given ethical constraints, which have been described in earlier reports to Congress. The design is built upon the Advisory Committee report, *Evaluating Head Start: A Recommended Framework for Studying the Impact of the Head Start Program*. This report set forth a framework for research on the impact of Head Start that is 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. The report also noted that the legislative mandate clearly requires that the impact study must address two main questions. These questions are discussed in the next section.

The Congressional Mandate

The goal of the national Head Start Impact Study (HSIS) is to answer the following critically important policy questions mandated in the 1998 legislation:[9]

1. *"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?"*
2. *"Under what circumstances does Head Start achieve the greatest impact? What works for which children? What Head Start services are most related to impact?"*

The first policy question consists of two parts: (1) the overall average effect of Head Start on the extent to which children enter school ready to learn, representing the **direct** impact of program participation on children's early development, and (2) the extent to which Head Start participation has an **indirect** effect by improving the ability of parents to support their children's learning and development, i.e., these factors are hypothesized to have a subsequent effect on the school readiness, and later school performance, of participating Head Start children. Though not specifically identified, it is also valuable to

[8] See Appendix A for the research-related amendments to the Head Start Act included in the 1998 reauthorization.

[9] 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.

understand the extent to which Head Start participation may affect the nature, duration, and quality of their early care and program experiences, which may, in turn, also lead to improvements in school readiness.

The second policy question recognizes the importance of also understanding how the impact of Head Start may vary: (1) it could vary for different types of children and their families; (2) it could vary according to the nature, duration, or quality of a child's early care and program experiences; or (3) it could vary among different communities.

These policy questions led to the specification of the following research questions that have guided the design and implementation of the Head Start Impact Study and that will form the basis for the eventual assessment of the efficacy of Head Start services:[10]

- ***Direct and indirect impacts:***
  - A. What impact does Head Start have on school readiness including children's approaches to learning; language development and emergent literacy; mathematical ability; physical well-being and motor development; and social and emotional development?
  - B. What impact does Head Start have on parental practices that contribute to children's school readiness (e.g., time spent reading to their child)? To what extent are these parenting practices related to child development outcomes?
  - C. What impact does Head Start have on the nature and quality of children's early care and program experiences (e.g., the intensity of reading instruction)? To what extent are these experiences related to child development outcomes?
- ***Variation in impacts related to external and pre-program characteristics:***
  - D. Do impacts vary according to children's characteristics at the time of entry into Head Start? These can include characteristics such as gender, race/ethnicity, age at program entry (3- vs. 4-year-olds), presence of disabilities, as well as the child's status on a number of developmental characteristics (e.g., emergent language ability) at the point of Head Start entry.
  - E. Do impacts vary by characteristics of the child's home environment at the time of entry into Head Start? These can include characteristics such as family structure (e.g., single parent, teen mother), household income, and parental practices related to school readiness before exposure to Head Start.
  - F. Do impacts vary by the characteristics of the communities in which Head Start programs operate? These include factors such as indicators of the economic and social environment (e.g., poverty, unemployment rates), and the policy environment related to the availability and quality of alternative services for low-income children (e.g., state and local government funding for preschool programs).

[10] For more details on the design of the Head Start Impact Study see: ***The Head Start Impact Study: Research Design and Preliminary Analysis Plan.(ACF, 2003)***

- ***Variation in impact related to characteristics that may be affected by Head Start participation:***

  G. Do impacts vary by parent's ability to support their children's development and/or characteristics of the home environment?
  H. Do impacts on children vary by the nature, duration, and quality of their early care and program experiences? For example, do impacts vary by the amount of language instruction they receive.

Contents of This Report

The Head Start Impact Study is one of only a handful of nationally representative, randomized design studies. In recommending a randomized design, the Advisory Committee acknowledged a number of challenges to implementing such a design in a well-established popular program such as Head Start. The remainder of this report presents the progress made toward meeting these challenges. It is divided into three chapters. Chapter 2 describes the design of the Head Start Impact Study, including the overall research design; the selection of participating Head Start programs, families, and children; and the characteristics of the selected study sample. Chapter 3 describes data collection procedures, measures, and sources. Preliminary information on the characteristics of selected programs and families based on data collected in fall 2002 are presented in Chapter 4. Chapter 5 provides a summary of study status to date. Six appendices provide additional technical details.

Chapter 2: The Design of the Head Start Impact Study

Overall Research Design

As discussed in Chapter 1, the primary purpose of this study is to determine whether Head Start has an impact on participating children and their parents and, if so, whether such effects vary among different types of children, families, communities, and configurations of children's early care and program experiences. 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*. This focus on impacts distinguishes this study from many others that seek primarily to examine relationships among participant outcomes and between participant outcomes and one or more individual or program characteristics (see, for example, the *Head Start Family and Child Experiences Study (FACES)*[11]). Instead, the present study uses information from participants *and statistically equivalent children who do not participate in Head Start* to determine whether Head Start **causes** the observed child and parent outcomes.

Given the goal of measuring program impacts, how do we determine what outcomes would have been observed if the children had not participated in Head Start? That is, how do we observe children having the same characteristics in two places at the same time—in Head Start and not in Head Start—and compare them? In many studies, researchers have addressed this problem by comparing program participants to a "participant-like" group of children who, in the ordinary course of events, do not participate in Head Start. However, 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 are likely to be different (on important factors that may lead to different outcomes independently of the effect of Head Start services) from those who do not. For example, parents who seek to enroll their children in Head Start may be more motivated to prepare them to start school than those parents who choose not to seek Head Start enrollment. Moreover, the reasons that these two types of parents make different decisions are both typically unobserved and likely to be related to the outcomes of interest in their own right. That is, the motivated parents do a host of things that may affect their children's development beyond enrolling them in Head Start. Because one cannot account for all of these underlying differences, one risks mis-attributing to program participation observed

[11] 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.

differences on a particular outcome measure (e.g., emergent literacy) that may, in fact, be a result of pre-existing differences between participants and non-participants.

To avoid this problem of selection bias, the Head Start Impact Study **randomly assigned** a sample of newly entering 3- and 4-year-old Head Start applicants[12] either to a ***treatment 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 may have received other 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 the treatment condition, or, in this case, the effect of Head Start participation on children's school readiness. The advantage of this research design is that if random assignment is properly implemented with a sufficient sample size, program participants should not differ in any systematic or unmeasured way from non-participants except through their participation in Head Start services.[13]

The remainder of this chapter describes how this randomized design was implemented and the characteristics of the study sample.

Selection of Head Start Programs

Most randomized studies are conducted in small demonstration programs or, if done in an ongoing program, in only 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 newly entering 3- and 4-year-old children. That is, both programs and children applying for entry into Head Start in fall 2002 were selected at random making results generalizable to the entire Head Start program,[14] not just the selected study sample. This approach responds to the Congressional mandate that the study provide *"a national analysis of the impact of Head Start"* based on the selection of Head Start grantee/delegate agencies[15] that *"operate in the 50 states, the Commonwealth of Puerto Rico, or the District of Columbia and that do not specifically target special populations."*

---

[12] 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.

[13] 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).

[14] As will be discussed below, some Head Start grantee/delegate agencies were intentionally excluded from the study. Study results will, therefore, be reported for both the restricted national sample and for the full Head Start population of newly entering children.

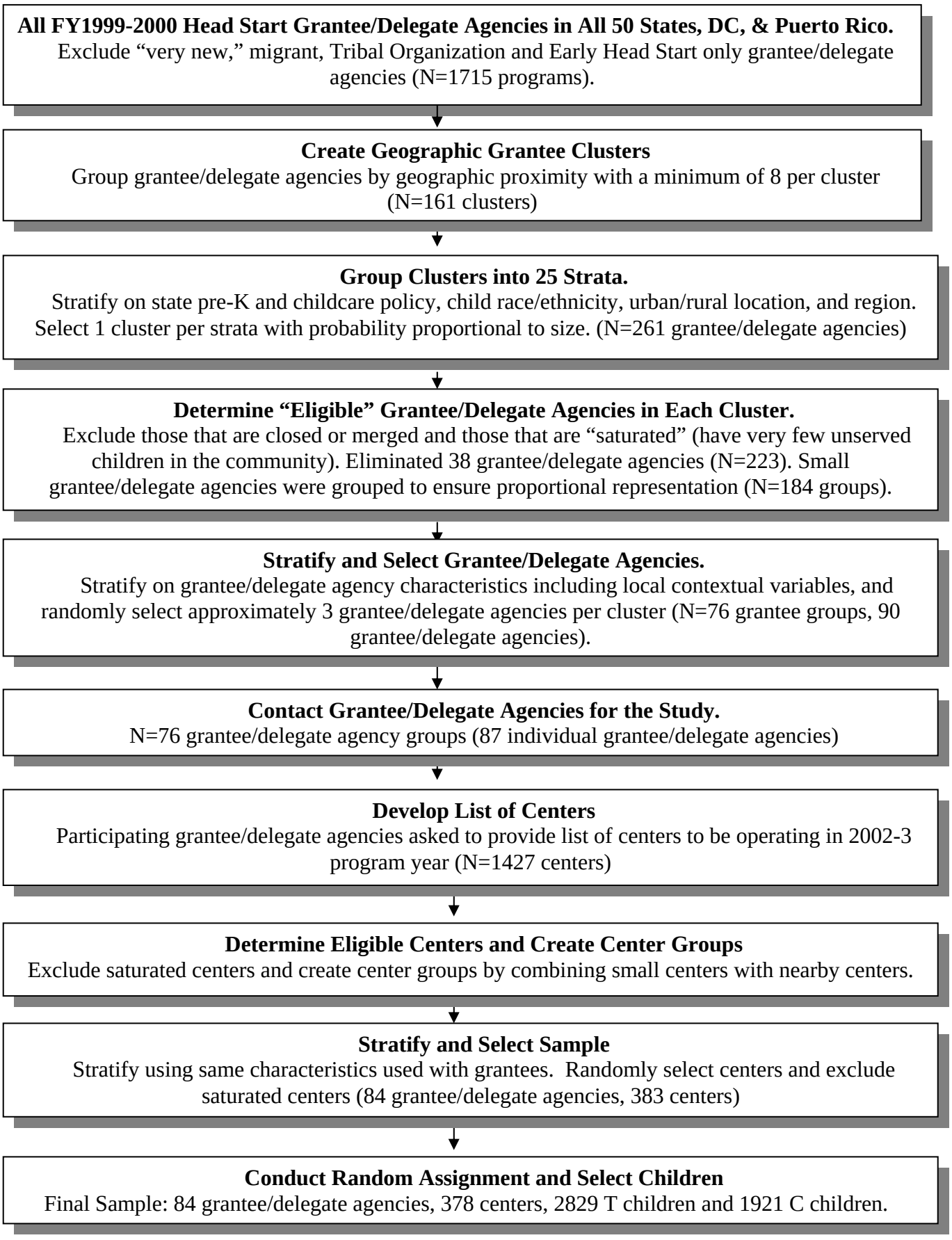
[15] The study sample includes both Head Start grantees and their delegate agencies. Grantees are organizations that have fiscal and administrative responsibility for programs in their jurisdiction. In some cases, they can subcontract with agencies to handle administrative oversight over some or all of these programs. Throughout this report we use the term grantee/delegate agencies to refer to both types of agencies.

Furthermore, the Advisory Committee recommended that the sample of Head Start grantee/delegate agencies should reflect variation in a variety of characteristics, including *"region of the country, race/ethnicity/language status, urban/rural, and depth of poverty in communities," and ". . .design of program as a one-year or two-year experience for children; program options (e.g., center-based, home-based, part-day, full-day); auspice (e.g., Community Action Agency, public school, non-profit organization); community-level resources; alternative child care options for low-income children; and, the nature of the child care market and the labor market in the community studied."*

The Grantee/Delegate Agency Sampling Process

To meet the legislative requirements, the study used a multi-stage sample selection process to select Head Start programs. The process, depicted in Exhibit 1, began by selecting a large sample of grantee/delegate agencies, screening these programs for inclusion in the impact study, and then selecting a sample of Head Start centers within the sampled grantee/delegate agencies. This process is described below:

Exhibit 1: Head Start Impact Study: Sample Selection Process



1. ***Initial grantee/delegate agency selection—***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) 1999-2000, after excluding (1) grantee/delegate agencies serving **only** special populations (migrant 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*."[16]
   This pool of Head Start programs was organized into 161 "geographic clusters" (to increase the ability to closely monitor random assignment and obtain high quality data) and then grouped into 25 strata to control for factors such as region of the country, urban/rural location, race/ethnicity, and variation in state pre-kindergarten and child care policies. One cluster was then randomly selected with probability proportional to total enrollment from each of the 25 strata, providing a total of 261 grantee or delegate agencies in the sampled clusters (to improve efficiency, random sub-sampling was done in three very large urban clusters).
2. ***Determining grantee/delegate agency eligibility***—To be eligible for inclusion in the randomized study, grantee/delegate agencies had to have enough "extra" or additional newly-entering applicants 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 Head Start programs where local staff were currently unable to serve all the eligible children seeking enrollment for fall 2003. This eligibility was determined from information gathered through telephone calls to all 261 grantee/delegate agencies, augmented with information provided by federal Regional Office staff and obtained from secondary sources such as local Child Care Resource and Referral Agencies. This screening process eliminated 28 grantees/delegate agencies (a reduction of 11 percent). Additionally, ten other grantee/delegate agencies had been closed or merged, further reducing the pool of eligible programs to 223 grantee/delegate agencies.
3. ***Selecting grantee/delegate agencies—***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." These groups (some of which consisted of a single grantee or delegate agency) were then stratified along several dimensions: urban location (central city, other urban, rural/small town); auspice (school based vs. all other agency types); percent Hispanic and percent Black enrollment; program options offered (part-day only, full-day only, both); and the percent of total enrollment represented by newly-entering 3-year-olds. Approximately three grantee/delegate agency 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 agency groups comprised of 90 individual grantee/delegate agencies.
4. ***Grantee/delegate agency recruitment—***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 sites 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—leaving 87 grantee/delegate agencies in 76 grantee/delegate agency groups.

[16] Defined as in operation for less than two years.

5. ***Identifying operating 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. This resulted in a list of 1,427 Head Start centers in the 87 grantee/delegate agencies that could potentially be included in the Head Start Impact Study.
6. ***Determining center eligibility and selecting a sample of study centers—***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 about 11 percent and the loss of three grantee/delegate agencies, but no grantee groups). Next, small centers were combined with nearby centers, and the resulting "center groups" were then stratified using the same characteristics employed in the selection of grantee/delegate agencies (excepting those that do not vary within grantee/delegate agencies such as region). A main sample consisting of an average of three center groups was 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 the randomized 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 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. (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. 38 of these centers were used. The final sample was 383 (448-103+38) centers).

As described below, this sample of Head Start grantee/delegate agencies and centers is designed to yield a sample of children that, when properly weighted, represents the national population of newly-entering children and their families (with the exclusions noted above).

How Representative Is the Program Sample?

Two questions arise regarding such a national sample:

1. How well does the study sample match the overall population from which it was drawn? Specifically, how well does the sample of grantee/delegate agencies match the complete universe of Head Start programs, after excluding newly-opened programs and those serving **only** migrant, Native American, or Early Head Start children?
2. Did decisions to exclude programs from the study sample, particularly saturated programs, contribute to measured differences between the study sample and the overall population from which it was drawn?

Answers to such questions speak to the external validity—or generalizeability—of the study sample.

Comparison of Grantee/Delegate Agency Sample with National Population

This first question is addressed by Exhibit 2 using data from the 1998-99 fiscal year PIR (the most recent data available at the time the study sample was selected). The first column lists important characteristics of grantee/delegate agencies that are available from the PIR, the second column provides the national distribution of all Head Start agencies on these characteristics in 1998-99, and the third and final column provides the characteristics of the initially selected study sample on these same characteristics, again as measured in 1998-99, *after the sample of programs was weighted to account for how programs were selected, including adjustments for the exclusion of grantees and centers that could not be included due to saturation.*

Statistically significant differences are indicated by the use of shading in Exhibit 2. As shown, the weighted study sample of Head Start programs is statistically similar to the PIR on all but three characteristics. The three observed differences indicate that the sample of grantee/delegate agencies may have overrepresented larger programs and overrepresented Hispanic/Spanish-speaking children as compared to the PIR. Two of the characteristics, Hispanic race/ethnicity and Spanish language, are highly correlated with each other. Thus, we effectively have two characteristics for which there are statistically significant differences. It should be noted, however, that these results are based on preliminary estimates of standard errors (which are used to determine statistical significance) and the more complex estimates of sampling variances that will be developed later are likely to eliminate or reduce the number of observed significant differences.

Exhibit 2: Comparison of Children Served by all Head Start Grantee/Delegate Agencies versus Impact Study Grantee/Delegate Agencies

Percent of Children by Category	Total Head Start Grantee/Delegate Agency Population, 1998-99 PIR	Head Start Impact Study Grantee/Delegate Agency Sample Population, 1998-99 PIR (N=90)
Auspice:		
Community Action Agency	39.2%	38.1%
School System	13.2%	7.2%
Other Non-profit	38.9%	45.0%
Government Agency	8.8%	9.7%
Funded Enrollment		
Less than 147	5.6%	3.2%
147-268	12.2%	10.3%
269-487	21.3%	16.3%
More than 497	60.9%	70.0%
Length of Day Served:		
Full Day	56.0%	61.1%
Part Day	44.0%	38.9%
Options Offered:		
Center based	92.0%	94.5%
Home based	3.9%	3.2%
Combination	1.8%	0.7%
Locally designed	2.2%	1.6%
Metro Status		
Urban	24.3%	22.8%
Rural	75.7%	77.2%
Head Start Region		
Northeast	23.9%	24.1%
North Central	18.8%	19.7%
South	35.1%	35.1%
West	22.3%	21.2%
State Pre-K Programs		
Are similar to HS	25.4%	25.2%
Partly similar to HS	18.5%	19.1%
Remaining States	56.0%	55.6%
Child Age:		
Under Age 3	0.4%	0.4%
Age 3	34.8%	33.5%
Age 4 or older	64.8%	66.1%
Child Race/Ethnicity:		
White	31.7%	30.0%
Hispanic	27.7%	33.4%
Black	36.7%	33.0%
Other	3.9%	3.6%
Child Language:		
English	77.4%	69.5%
Spanish	19.0%	25.2%
Other	3.6%	5.0%

Notes: (1) Shading indicates statistically significant differences. (2) Data for national population and sample are from 1998-1999 PIR. (3**)** Population totals exclude grantees/delegates that: serve only migrant, Native American, or Early Head Start children; those that do not serve 3- and 4-year-olds; those operating in territories other than Puerto Rico; and those newly created in FY1999-2000 or later. (4) ACF funded enrollment was used for auspice, metro status, region, and state preschool programs; total funded enrollment was used for length of day and option; actual enrollment used elsewhere (actual enrollment can exceed funded enrollment). (5) For size categories, ACF funded enrollment was used to create categories. (6) These percentages are based on total enrollment, not just newly entering 3- and 4-year-olds.

Effects of Exclusion Due to Saturation

Exclusion of Grantee/Delegate Agencies: One factor that may account for the differences noted in Exhibit 2 is the exclusion of saturated grantee/delegate agencies. This is because saturated programs have a lower percentage Hispanic enrollment and are smaller than non-saturated programs (see Appendix B).[17]

As noted above, 38 (15%) of the initially selected grantee/delegate agencies were dropped from consideration for inclusion in the impact study. Of these, 28 were determined to be saturated and unable to provide "extra" children for the control group, and ten grantees/delegate agencies had either closed or had merged with another grantee/delegate agency. After eliminating these 38 programs, the final frame consisted of 223 grantee/delegate agencies in 25 geographic clusters. The 28 saturated programs represent an estimated 11 percent of the eligible Head Start program universe. However, due to the smaller than average size of the excluded programs, they represent only 3.9 percent of the total number of newly-entering 3- and 4-year-old applicants. Similarly, ten closed or merged programs are estimated to represent 60 such programs on the PIR frame containing under 3 percent of total newly entering 3- and 4-year-old applicants.

There is potential for bias due to the exclusion of these saturated Head Start programs from the sampling frame if the saturated programs are systematically different from the non-saturated programs and if the characteristics on which they differ are correlated with the outcome measures that will be examined for the impact analysis (see the next chapter). As shown in Appendix B, preliminary analyses indicate that the saturated programs in the sample are systematically different from the non-saturated programs. They are smaller, more likely to be school-based, and have a smaller percentage Hispanic enrollment than the non-saturated programs.[18] As a consequence, these potential differences will have to be accounted for during the estimation of program impacts through the use of both weighting and statistical adjustment.

Exclusion of Head Start Centers: As also discussed above, within the sample of 76 grantee/delegate agency groups, all of the individual Head Start centers were screened for inclusion in the study sample. As also discussed in Appendix B, this procedure resulted in the deletion of 154 centers from a frame of 1,423 centers. The excluded 154 centers represent an estimated eight percent of the Head

[17] Program weights were adjusted for these exclusions, but the adjustment may not totally compensate for any potential bias.

[18] As with the previous discussion, these estimates may change as we conduct further tests because (1) the variances at the grantee/delegate agency level are not very stable because the number of saturated programs is small, and (2) the variance estimates do not yet include the between-PSU component of variance. These factors may result in an overstating of the statistical significance of the differences.

Start center universe, approximately 5 percent of the total number of 3- and 4-year-old children enrolled in Head Start, and 4.7 percent of newly-entering 3- and 4-year-old applicants. As with grantee/delegate agencies, the 154 saturated centers are significantly smaller, have lower Hispanic enrollment, and have a larger percentage of newly-enrolled 3-year-olds than the non-saturated centers. Again, as noted above, these differences will have to be taken into account during the statistical analysis of program impacts.

Overall, the estimated coverage rate of the population of newly-entering 3- and 4-year-olds for the Head Start Impact Study sample is 92 percent, which is obtained by multiplying the coverage rate for Head Start grantee/delegate agencies and by the coverage rate for centers [i.e., (100-3.9% for programs)*(100-4.7% for centers)]. Stated another way, the overall under-coverage rate is 8 percent.

Understanding Saturation: To understand the context in which saturation occurs, we examined trends in child poverty from 1997 to 2001. Data from consecutive March Supplements to the Current Population Survey, collected in March 1998 through March 2002 for the previous calendar year, show that between 1997 and 2000, the estimated number of 3- and 4-year-old children in poverty in the United States fell by 475,415. The number then rose again by 94,533, in the most recent year with data, 2001. Overall, the estimated number of 3- and 4-year-old children in poverty in the United States fell about 21 percent between 1997 and 2001 (from 1,790,183 to 1,409,301). Thus, by the start of the Head Start Impact Study in 2002, Head Start grantee/delegate agencies likely had a smaller eligible population to draw from than in the mid-1990s, and this may have contributed to the observed level of service saturation.

Further evidence of this decline in the eligible population comes from the PIR data, which provide information about a grantee/delegate agency's funded and actual enrollment[19]. Among the 261 grantee/delegate agencies in the initial pool selected for the Head Start Impact Study, some were under-enrolled. On average, about one in five was under-enrolled by at least five percent per year in the three years preceding the study, a slight improvement from the one in four under-enrolled in 1997-1998. Further, the PIR data suggest that for many grantee/delegate agencies, under-enrollment (defined as being at 94 percent of funded enrollment or less at a given time point) appears to be a consistent problem:

[19] Total funded enrollment is the number of children the grantee/delegate agency has been funded to serve, regardless of funding source. Actual enrollment is the total number of children enrolled at any time during the year.

- 14.9 percent of the 261 grantee/delegate agencies experienced under-enrollment at least once during the four year period but not consistently across any year.
- 26.4 percent experienced a problem with under-enrollment consistently across at least one year but not all four years.
- 11.1 percent experienced chronic under-enrollment across all four years.

The data also indicate that challenges with under-enrollment fluctuate from year to year, i.e., grantee/delegate agencies that have problems one year may not have problems the next, and vice versa. Thus, saturation is a complicated issue for Head Start programs, as they are constantly trying to achieve the right balance among population shifts, changes in program regulations, pre-kindergarten resource availability, fluctuating economic conditions, and other factors.

Child-Level Selection and Assignment Process

Child Sample Selection

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.[20]

The study assigned local site coordinators to work with grantee/delegate agencies in each of the 25 geographic clusters to ensure that parents received this information with their applications. These site coordinators were also responsible for obtaining data from **all** applications for the 2002-03 program year (to ensure equal treatment of all applicants), and listing these data on a roster that was subsequently key-entered by central office study staff. Returning children, and a small number of grantee-requested "high-risk" exclusions,[21] were eliminated from these lists, and checks were made for duplicate records. The high-risk exclusions were made on a case-by-case basis with each grantee/delegate agency and in close consultation with Administration for Children and Families staff. 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 test these children and include them in the study sample (e.g., blindness). Each grantee was limited to one exclusion per center,

[20] Non-admitted (control group) children selected for the study sample were prohibited from participating in Head Start during 2002-03. Those who were 3-year-olds could, however, re-apply for Head Start in 2003-04 and may be admitted if eligible.

[21] 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) the Field Test demonstrated that the potential

but in fact only 276 exclusions were taken out of a total of approximately 18,000 newly-entering 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. Site coordinators 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 non-Head Start 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.

The goal was to randomly select, on average, 27 children from the expanded list at each of the sampled center groups: 16 to be assigned to the Head Start group and 11 to be assigned to the non-Head Start group. For an average center group, the 11 non-Head Start control group children represented about nine percent of total enrollment. Where necessary, stratification was used, such as in situations where the degree of saturation varied by program option (part-day vs. full-day) or age cohort. In some cases, where fewer children than expected were actually available, a smaller sample of children was selected for the study.

The original legislative mandate required that the Head Start Impact Study "*to the extent practicable*" address possible variation in program impact related to "*the length of time a child attends a Head Start program (and) the age of the child on entering the Head Start program.*" This requirement reflects the hypothesis that different program impacts may be associated with one versus two years of Head Start experience. It also reflects a trend of increased enrollment of 3-year-olds in some grantee/delegate agencies presumably due to the growing availability of preschool options for 4-year-olds (often state-sponsored programs). Consequently, the study includes samples of both **newly-entering** 3-year-olds (studied through two years of Head Start participation, kindergarten, and first grade) and **newly-entering** 4-year-olds (studied through one year of Head Start participation, kindergarten, and 1st grade). The sample of 3-year-olds is slightly larger than the sample of 4-year-olds to protect against the

---

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.

possibility of higher study attrition resulting from an additional year of longitudinal data collection for the younger children.[22]

Random Assignment of Children

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, only five centers (1.3 percent) did not fully cooperate with the study requirements, resulting in 378 centers with successful random assignment that can be divided into the following three groups:

- **Obtained Full Sample:** Random assignment was completed at 173 Head Start centers that provided the full expected sample of children.
- **Obtained Smaller Sample:** Random assignment was completed at 150 Head Start centers 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 that provided a larger than expected sample (i.e., because application rates for newly-entering children were higher than originally estimated, sample sizes were increased to compensate for other centers that were unexpectedly low).

In total, **4,750 newly entering children** were randomly assigned and included in the Head Start Impact Study:

Age Cohort	Head Start (Treatment) Group	Non-Head Start (Control) Group	Total Sample
3-year olds	1,552	1,045	2,597
4-year-olds	1,277	876	2,153
Total	2,829	1,921	4,750

As indicated above, about 60 percent of the sample was assigned to the Head Start group and about 40 percent was assigned to the non-Head-Start group. This imbalance reduces the precision of the impact estimates by just two percent (compared with a balanced 50-50 design). However, it provided several important benefits: (1) it significantly increased the ability to recruit grantees and centers by decreasing the number of extra children needed for the control group; (2) decreased the loss of sites due to saturation; and (3) saved considerably on data collection costs because treatment group members (who

[22] This equal sampling of 3- and 4-year-old enrollees was done 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 4-year-olds include 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.

participate in Head Start) require less effort to track and interview over time than children in the non-Head-Start control group.

Comparison of Treatment and Control Children and Their Families at the Time of Random Assignment

Was random assignment implemented well enough to support the intended impact analysis? This question is addressed in Exhibit 3, which shows a comparison of children randomly assigned to the Head Start and non-Head Start groups on characteristics that were measured and available at the time of random assignment (these data were drawn from parental applications for Head Start). As shown, there are no statistically significant differences between the two randomly assigned groups indicating that they are well matched on the available characteristics. Consequently, we can conclude that the initial randomization was done with high integrity and that the samples can provide the necessary confidence in the validity of the eventual impact estimates.

Exhibit 3: Comparison of Head Start and Non-Head Start Study Groups: Baseline Child and Family Characteristics Measured at Random Assignment

Characteristic	Head Start (Treatment) Group	Non-Head Start (Control) Group	Difference: (Head Start) – (Non-Head Start)
Child Age Cohort:			
3-year old	52.8%	53.2%	-0.4%
4-year-old	47.2%	46.8%	0.4%
Child Gender:			
Boys	49.6%	48.7%	0.9%
Girls	50.4%	51.3%	-0.9%
Child Race/Ethnicity:			
Hispanic	42.2%	43.0%	-0.8%
Black	27.0%	25.5%	1.5%
White	28.1%	29.1%	-1.0%
Other	2.7%	2.5%	0.2 %
Child Language:			
English	65.7%	64.7%	1.1%
Spanish	31.4%	32.2%	-0.8%
Other	2.8%	3.1%	-0.3%
Parent Language:			
English	68.0%	68.1%	-0.1%
Spanish	30.5%	30.9%	-0.4%
Other	1.5%	1.1%	0.4%
Income Eligibility (for Head Start):			
Yes	93.0%	91.7%	1.3%
No	7.1%	8.3%	-1.3%

Notes: (1) Data source: Roster information used at time of random assignment; (2). These percentages reflect individual level data on newly entering 3- and 4-year-olds and do not include saturated grantees and centers; and (3) There were no statistically significant differences between groups.

Chapter 3: Data Collection Procedures, Measures, and Data Sources

This chapter begins with an overview of the study's overall data collection plan and then describes the Field Test that was conducted prior to implementing the full-scale study. The chapter next describes the measures and data sources being used to obtain the necessary information about children, parents, and programs.

Overview of Planned Data Collection

Data collection for the main study began in fall 2002 and will continue through spring 2006, following children from entry into Head Start (either 3 or 4 years old) through the end of their preschool years, end of kindergarten, and end of 1st grade (see Exhibit 4). The study is collecting comparable data for both Head Start and non-Head Start group children, including interviews with parents, direct assessments of children's development, 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 teacher/care provider ratings of individual study children.

The Field Test

The Field Test, conducted from fall 2001 through spring 2002, tested the feasibility of all study procedures, including the recruitment of programs to conduct random assignment, the random assignment procedures themselves, and the data collection instruments and procedures.

Exhibit 4: Schedule of Planned Data Collection

School Year **3-Year-Old Cohort (C-1)** **4-Year-Old Cohort (C-2)**		**2002-2003** **3-year-old Preschool** **4-year-old Preschool**		**2003-2004** **4-year-old Preschool** **Kindergarten**		**2004-2005** **Kindergarten** **Grade 1**		**2005-2006** **Grade 1**	
Data source	*Cohort*	*Fall*	*Spring*	*Fall*	*Spring*	*Fall*	*Spring*	*Fall*	*Spring*
Child Assessment	C-1	✓	✓		✓		✓		✓
	C-2	✓	✓		✓		✓		
Parents/Primary Caregiver Interview	C-1	✓	✓	✓	✓	✓	✓	✓	✓
	C-2	✓	✓	✓	✓	✓	✓		
Administrative Records	C-1						✓		✓
	C-2				✓		✓		
Staff Interviews	C-1		✓		✓		✓		✓
	C-2		✓		✓		✓		
Observed Quality of Care Settings	C-1		✓		✓		✓		✓
	C-2		✓		✓		✓		

The sample for the Field Test included eight grantee/delegate agency groups in geographic clusters that were not part of the full-scale study. The clusters were selected to vary in terms of the availability and comprehensiveness of program or child care options (in addition to Head Start) within the community. The grantee/delegate agencies and/or centers varied across auspices, number of unserved children, local service richness, grantee size, ethnicity of population served, and urbanicity.

Within the selected grantee/delegate agencies, the Field Test involved approximately 400 3- and 4-year-old new applicants to Head Start spread across a total of 21 Head Start centers. This early feasibility assessment served several purposes. First, the Field Test was an important way to understand the recruitment and application process in Head Start programs and essential to testing and refining random assignment procedures. Second, it allowed for the testing of data collection measures and procedures and provided an opportunity to learn more about the best strategies to deal with the complexities of both Head Start and other available programs for low-income children. Valuable lessons learned and used to inform the design of the full-scale study included the following:

- **It is critical to build and maintain study team, regional office, and grantee/delegate agency partnerships.** For the full-scale study, Recruitment Teams, led by senior study staff and including a local study site coordinator, communicated with both federal regional office and grantee/delegate agency staff throughout the recruitment and random assignment phases of the project. Recruitment teams made site visits to each grantee/delegate agency and maintained frequent communication with key staff.
- **It is important to gain a thorough understanding of the recruitment and application process of each grantee/delegate agency and to "tailor" random assignment and data collection procedures to fit local processes.** Recruitment teams for the full-scale study gathered information on local program operations and worked collaboratively with staff to design study procedures that maintained the necessary scientific rigor while, to the extent possible, minimizing disruption to usual recruitment and enrollment procedures.
- **It is necessary to understand that Head Start enrollment does not always happen as expected or when expected, and this can affect whether sites can implement random assignment and provide the desired sample size.** Consequently, random assignment for the full-scale study was conducted in "waves" from spring through fall 2002 (e.g., using monthly batches of new applicants). In addition, a small number of centers initially recruited for the study were eliminated due to saturation.
- **It is important to address program concerns about serving the neediest families, monitoring compliance, notifying parents of enrollment decisions, and staff responsibilities and burden.** Recruitment teams met with program administrators, teaching and social service staff, and parents to provide an orientation to the study and address concerns. Study Fact Sheets were also developed for each group to answer frequently asked questions.

- **High response rates could be achieved from both the Head Start and non-Head Start control groups.** An 80 percent response rate (parent and child combined) was attained from the fall 2001 field test data collection and maintained in the spring data collection.

Data Sources and Measures

To address the Congressional mandate, and respond to the key research questions, a wide variety of data sources and measures are being used to assess the effects of Head Start on fostering and enhancing child development, including direct child assessments, parent/primary caregiver interviews, interviews with providers of early care and program services to participating study children, observations of early care settings, and the collection of secondary data. The following types of data are being collected:

- **Child outcomes** through: 1) direct assessments of children's development, 2) parent/primary caregiver reports of children's development, 3) teacher/care provider report of children's development, and 4) administrative school records. Child outcomes will be measured in the key domains of approaches to learning, language usage and emergent literacy, math, cognition and general knowledge, physical well-being, motor development, social and emotional development, and eventually, school success.
- **Characteristics and quality of children's home environments** through: parent/primary caregiver interviews including such topics as parenting practices, household composition, parents' health and mental health status, safety of environment, child stress/risk, and receipt of health and other services.
- **Characteristics and quality of the primary preschool and child care arrangement** through: (1) interviews with center-based directors, (2) surveys of teachers or interviews with care providers, and (3) observations of these settings.

Each type of data is described below in more detail. It is first important to understand how children's care settings have been defined for the present study because this had influenced the choice of instruments and what data have been collected for individual children.

Defining and Assessing the Child's Care Setting

Head Start children may attend other programs during the hours they are not in Head Start and will be exposed to numerous other experiences at home and elsewhere that shape their development. Similarly, children assigned to the control group will also have a variety of experiences that affect their development, such as care provided in their own home, as well as time spent in other settings. Although all care experiences are important, practical considerations and differences among the settings limit what data can be collected and in what depth across these various settings. Consequently, criteria were developed to help identify, categorize, and prioritize the range of settings in which children spend time.

These criteria were then used to select the one setting that will later serve as the most appropriate comparison to Head Start for each child in the control group. The characteristics and quality of this setting were assessed through caregiver interviews and interviewer observations. The setting chosen was based on a hierarchical decision tree, prioritizing out of home experiences with center-based care over family day care. For those children who did not have any out-of-home program or care experiences, a small number of additional items were used to assess the home setting as the primary daytime care setting.

The information used to define each child's care setting is derived from the parent interview. Parents are asked to identify all of the places where sampled children spend time between the hours of 8am and 6pm weekdays during the school year. This information is used to categorize each child's experiences into the following Monday through Friday timeframes:

- The care setting (including parental care) where they spend the most time between 9am and 3pm—the portion of the day when most children are in Head Start.
- The non-parental child care arrangement where they spend the most time between 8am and 6pm. This arrangement is termed the **Primary Non-parental Care Setting** and serves as the data collection site for comparative data on the quality of care setting.
- Any additional non-parental child care arrangements where they spend time between 8am and 6pm.

Exhibit 5 summarizes the data sources for each of the different settings in which children may be found. These data sources are discussed in greater detail below.

Exhibit 5: Assessment of Setting Characteristic by Type of Care Setting

	Parent Report[23]	Center Director	Teacher/Care Provider Report	Observation	Care Provider Interview	Teacher Survey
Primary Non-Parental Care Setting[24] (Monday-Friday, 8am-6pm) Hierarchical Selection						
1. Head Start	X	X	X	-ECERS-R -Arnett		***X***
2. Center-Based (most time, at least 5 hours per week)	X	X	X	-Teacher directed activities -5 question observation		X
3. Non-Center-based (most time, at least 5 hours per week)						
-relative	X		X	-FDCRS -Arnett	***X***	
-non-relative	X		X	-5 question observation	X	
4. Own home						
-with relative	X		X	5 question observation	X	
-with non-relative	X		X		X	
Primary Care Setting is Child's Parent (Monday-Friday, 9am-3pm)						
1. Own Home with parent	X			5 question observation		
Additional Child Care Settings (Monday-Friday, 8am-6pm)						
1. Center-based	X					
2. Non-Center-based						
-relative	X					
-non-relative	X					
3. Own Home						

[23] Based on the winter cataloging of care settings.

[24] If child has more than one setting between 8am-6pm, only one will be chosen for measuring quality through observation and provider report. To choose the setting, the hierarchy (as listed in the table with #1 Head Start being highest in priority) will be followed, and if child has more than one setting of a certain type, the setting with the most hours will be chosen.

Direct Child Assessments

Child assessments provide direct measures of how well Head Start and non-Head-Start preschool programs, or other child care services, are achieving the goal of assisting children to be physically, socially, and educationally ready for success in kindergarten. These measures, which represent the key developmental outcomes for the Head Start Impact Study, are being assessed using a 35- to 45-minute battery of standardized tasks administered by specially selected and trained child assessors. The assessments are being administered at the location where the child spends the most time, Monday through Friday from 9am to 3pm.

The assessment battery (see Exhibit 6) is composed of a short series of tasks that are feasible and interesting for preschoolers to carry out and that have been shown to be predictive of later school success and achievement or academic difficulties (test citations are provided in Appendix E). There is an emphasis on tasks that relate to the acquisition of reading skills because reading is so central to success in school and to later functioning in society. For non-English speaking children, the Interviewers used a Language Decision Form at the time of the assessment to determine the appropriate language for conducting the assessment (See Appendix F).

For children requiring assessment in Spanish, a bilingual Interviewer/Assessor administered the assessment battery in Spanish in fall 2002, and also administered two subtests in English—the PPVT-IIIA and the Woodcock-Johnson III Letter-Word Identification. In spring 2003, the children assessed in Spanish in fall 2002 were assessed primarily in English, along with the continued administration of two Spanish language measures: the TVIP and the Batería Woodcock-Muñoz Identificación de letras y palabras. One exception is Puerto Rico where, because instruction is in Spanish, all children were assessed only with the complete Spanish battery in spring 2003.

For children who could not be assessed in English or Spanish, either a bilingual Interviewer/Assessor or an interpreter for the child's language were used. The Interviewer/Assessor or interpreter used the English version of the booklet, translated the

Exhibit 6: Child Assessment Battery by Study Year and Child Language: Preschool Years

Domain	Measure	Child's Language at Beginning of Head Start Year					
		English		Spanish		Other	
		Fall 2002	Spring 2003	Fall 2002	Spring 2003	Fall 2002	Spring 2003
Language Development And Literacy	PPVT III (adapted)	E	E	E	E	-	E
	TVIP (adapted)	-	-	S	S	-	-
	CTOPPP Print Awareness	E	E	S	E	-	E
	CTOPPP Elision	E	E	S	E	-	E
	Woodcock-Johnson III Letter Word Identification	E	E	E	E	-	E
	Woodcock-Johnson III Spelling	E	E	-	E	-	E
	Woodcock-Johnson III Oral Comprehension	E	E	-	E	-	E
	Woodcock-Munoz Letter Word Identification	-	-	S	S	-	-
	Woodcock-Munoz Dictation	-	-	S	-	-	-
	Story and Print Concepts	E	E	S	E	O	E
	Letter Naming Task	-	E	-	E	-	E
Mathematics	Counting Bears	E	E	S	E	O	E
	Woodcock-Johnson III Applied Problems	E	E	-	E	-	E
	Woodcock-Munoz Applied Problems	-	-	S	-	-	-
Perceptual Motor	McCarthy Draw-a-Design	E	E	S	E	O	E
	Woodcock-Johnson III Spelling (H1-H7)	E	E	-	E	-	E
	Woodcock-Munoz Dictation (G1-G6)	-	-	S	-	-	-
Other Cognitive Ability	Color Names	E	E	S	E	O	E
Sustained Attention	Leiter-R (adapted)	E	E	S	E	O	E
Assessor Ratings	Task persistence	X	X	X	X	X	X
	Attention span	X	X	X	X	X	X
	Body movement	X	X	X	X	X	X
	Attention to directions	X	X	X	X	X	X
	Comprehension of directions	X	X	X	X	X	X
	Verbalization	X	X	X	X	X	X
	Ease of Relationship	X	X	X	X	X	X
	Confidence	X	X	X	X	X	X

Notes: "-" =subtest not administered during the data collection period; "E"=subtest administered in English; "S"=subtest administered in Spanish; "O"=subtest administered in language of child other than Spanish or English; "X"= an observed assessor rating not a direct Child Assessment.

instructions into the child's language and administered four subtests: McCarthy Draw-A-Design, Color Names & Counting, Leiter-R-Adapted, and Story & Print Concepts. For spring 2003, these children were assessed in English.

Parent/Primary Caregiver Interviews

The parent/primary caregiver interview, conducted concurrently with the child assessments in both fall 2002 and spring 2003, collects information in the following areas: (1) parental beliefs and attitudes toward their child's learning and satisfaction with their child's child care experience; (2) family household and demographic information, including parent-child relationships and the characteristics of the child's home environment; (3) parent ratings of their child's behavior problems, social skills, and competencies; (4) parent's perception of their child's accomplishments; (5) parent's perception of their relationship with their child; and (6) child and family receipt of a variety of comprehensive services.

The parent interview also includes items that address key parenting behaviors that may be affected by participation in Head Start, such as attitudes about how parental authority is conveyed, encouragement of the child's exploration and independence, and parental activities, particularly reading. In spring 2003, parents were also asked about services they received and the help they received in coordinating the services. Other topics included the child's transition from preschool to kindergarten and any information or services the family received to assist with this transition. These data are being collected because it is hypothesized that such behaviors may be affected by Head Start, and changes in parenting behavior may have a subsequent impact on children's development.

Staff Surveys and Interviews

Data are also being collected through surveys and/or interviews with staff in the primary non-parental care setting. In center-based programs, interviews are being conducted with the center director and paper-pencil surveys are being completed by the teacher in the classroom where the child spends the most time. For children whose primary non-parental setting is a family daycare home, data comparable to that collected through the center-based director interview and teacher survey are collected through a single combined care provider interview with the child care provider.

In spring 2003, data were collected from each child's teacher/care provider in the child's primary non-parental care setting. This included information on: their demographic characteristics; education, training, and experience with young children; the nature of the setting in which they work; and the types of services they provide to the selected study children. In particular, items on literacy-promoting activities are included in three questionnaires: the Teacher Survey, Care Provider Interview, and Center Director Interview. During the kindergarten year, the Teacher Survey will collect information about the kindergarten program, provisions that are made for the child's transition to kindergarten, and whether the teacher obtained any information from the Head Start program or other care settings about the child's development or special needs.

Teachers and care providers are also asked to rate each of the study children who are in their classroom or care in the following areas: relationships with adults; classroom behavior and conduct, problem solving and initiative, social relationships, creativity, musical and movement ability, language ability, and mathematical ability. Although not as objective as direct assessments or observations by impartial observers, parent/primary caregiver and teacher/care provider ratings of children's accomplishments and behaviors are an important source of information because these individuals see children over extended periods of time and in a variety of situations. These ratings, when combined with the direct child assessments, provide for fairer and more robust appraisals of children's skills and competence and such assessments have been found to predict later child outcomes.

The Center Director Interview provides information on the operation and quality of the Head Start and non-Head Start center-based programs. Issues addressed in this interview include: child recruitment and enrollment; staffing; teacher professional development; parent involvement; curriculum; classroom organization and activities; child assessment; home visits; and kindergarten transition. The Care Provider Interview collects comparable information regarding child care for non-Head Start children in non-center-based settings or those who are at home with a relative or non-relative other than the parent/primary caregiver. The interview includes questions on the number of children in the child care setting, staffing, typical child activities, and beliefs on how children should be taught and managed.

Care Setting Observations to Assess Quality

In addition to interviews and surveys, direct observations are being used to assess the quality of each child's primary non-parental care setting. These tools provide direct measures of the extent to which Head Start centers and other child care arrangements employ skilled teachers/care providers and provide developmentally appropriate environments and curricula for the children. Trained observers spend enough time in each setting to ensure observation of a major portion of the daily schedule and a variety of center, classroom, and family daycare activities.

The following standardized observational instruments and coding schemes that have been widely tested in child development research are being used in the Head Start Impact Study:

- ***The Early Childhood Environment Rating Scale (revised) (ECERS-R),*** provides a global rating of classroom quality based on structural features of the classroom including personal care, furnishings, language and reasoning activities, gross and fine motor activities, social activities, and provisions for adults and teachers. For observations in non-center-based settings outside the child's own home the related ***Family Day Care Rating Scale (FDCRS)*** is being used.
- ***The Classroom Observation of Teacher-Directed Activities Checklist*** is designed to collect more definitive information about the kinds of teacher-guided instructional activities that take place in the classroom.
- ***The Arnett Scale of Teacher/Provider Behavior*** is a rating scale consisting of 26 items organized under five areas: sensitivity, punitiveness, detachment, permissiveness, and promotion of independence.

In the interest of having some observational measure of quality across all types of settings, a five-question observational instrument designed for use in the child's home as well as in center-based and family daycare homes is being used. These supplemental items are completed by interviewers for all primary care settings regardless of whether the setting is Head Start, another center-based program, the child's home, or someone else's home and are completed in addition to the standardized observational instrument appropriate to each type of setting, as noted above. Interviewers rate each child's primary care setting in four areas: overall safety, basic hygiene standards, availability of educational materials, and overall positive and negative interaction between provider and child.

At the kindergarten and first grade level, the study will rely primarily on existing indicators of school quality that have been found to be important determinants of student performance on tests of academic achievement and other indicators of school success. Examples

include: school grade composition and enrollment size; economic level of the student population (free and reduced-price lunch eligibility); racial and ethnic composition of the students; average pupil/teacher ratios; average per pupil expenditures; the availability of federal funds (e.g., Title I) and other special programs (e.g., after school programs, and programs for Limited English Proficiency students); teacher and other staff characteristics; measures of school safety; measures of average student achievement by subject area; and, attendance and dropout statistics.

Specifically, the schools attended by the study children will be linked to annual data collected from every public school in the U.S. by the Department of Education's National Center for Education Statistics (NCES). Two primary data sources will be used: The Common Core of Data (CCD) for Public and Private Elementary Schools, and the Schools and Staffing Survey – Data for Public and Private Elementary Schools. The NCES data will be augmented by linking to state- and district-level data that are publicly available as school "report cards" under state accountability systems.

Chapter 4: Preliminary Information on Child and Family Characteristics; Fall 2002 Data Collection

This final chapter begins with a presentation of preliminary information about the samples of Head Start and non-Head Start children and their families using information from the fall 2002 baseline parent surveys. The overall baseline response rates are presented, along with the response rates for the Head Start and non-Head Start control groups, and a discussion of the implications of these response rates for later data analysis. The chapter ends with a description of the range of preschool and child care arrangements used by children who were randomly assigned to the non-Head Start control group and for whom a parent interview was completed. These non-Head Start children will serve as the basis for calculating the impact of Head Start on child and family outcomes.

Comparison of Head Start and Non-Head Start Groups

Response Rates

Data collection for fall 2002 and spring 2003 has been completed. For the fall 2002 data collection, data entry, initial checking, and statistical weighting have all been completed. Work has also begun to assess the psychometric properties of the direct child assessments. Overall response rates for both child assessments and parent/primary caregiver interviews have been very good, with a combined rate of 80 percent for fall 2002 and a projected rate of 83 percent for spring 2003.[25] However, as depicted in Exhibit 7, the fall 2002 response rate for the group of treatment families was higher than that of control families.

[25] The spring 2003 response rates have not yet been finalized.

Exhibit 7: Comparison of Response Rates for Head Start and Non-Head Start Groups, Fall 2002

Instrument	Fall 2002	
	Head Start (Treatment)	**Non-Head Start (Control)**
Child Assessment	85%	72%
Parent Interview	85%	74%

Preliminary response rates for spring 2003 indicate a higher overall response rate than in the fall (i.e., 83% vs. 80%) and a slightly smaller differentiation between treatment and control group response rates (i.e., 88% Treatment vs. 78% Control for child assessments and 86% Treatment vs. 79% Control for the parent interviews).

Respondent Characteristics

It is important to explore whether the differential response rates for the treatment and control groups led to differences in the average characteristics of the two groups. Exhibit 3 showed that at the time of random assignment there were no differences in these two groups on the characteristics measured (child age, gender, race and parent and child language and income eligibility). As shown in Exhibit 8, when looking at fall 2002 respondents only, there is one significant difference on the same variables measured at random assignment. Specifically, families with children whose primary language is English were more likely to respond if they were in the treatment group, rather than the control group.

Further data on similarities and differences between the two groups of respondents is also available from the fall 2002 parent/primary caregiver interview and these provide a greater array of characteristics. Exhibits 9 and 10 provide information from the parent/primary caregiver interview on the characteristics of the respondents (both children and their families) assigned to the Head Start and non-Head Start study groups. Data in these tables are weighted (using the final weights) to represent the national population of new Head Start applicants at all Head Start grantees/delegate agencies. (Appendix D provides the same information using unweighted data, data weighted with and without adjustments for non-response, and an analysis of the differences. This technical information is intended to help readers understand the effect of weighting, which will have important implications for later impact analyses.) As shown in Exhibits 9 and 10, the two groups are well matched on all baseline individual-level characteristics tested. Again, as was the case with the data from the random assignment rosters, the only difference found was related to language.

Exhibit 8: Comparison of Head Start and Non-Head Start Study Groups Using Roster Data: Fall 2002 Respondents only

Characteristic	Head Start (Treatment) Group	Non-Head Start(Control) Group	Difference: (Head Start) – (Non-Head Start)
Child Age Cohort:			
3-year old	53.3%	53.1%	0.2%
4-year-old	46.7%	46.8%	-0.2%
Child Gender:			
Boys	49.0%	47.9%	1.1%
Girls	51.0%	52.1%	-1.1%
Child Race/Ethnicity:			
Hispanic	42.1%	44.1%	-2.0%
Black	27.3%	25.0%	2.5%
White	28.2%	29.4%	-0.2%
Other	2.3%	2.6%	-0.3%
Child Language:			
English	66.3%	62.3%	4.0%**
Spanish	31.4%	34.2%	-2.8%
Other	2.3%	3.5%	-1.2%
Parent Language:			
English	68.6%	66.7%	1.9%
Spanish	30.5%	32.5%	-2.0%
Other	1.0%	1.0%	0.1%
Head Start Income Eligibility:			
Yes	93.3%	91.7%	1.6%
No	6.7%	8.3%	-1.6%

**= $p \leq 0.05$

Note: Data source: Roster information used at time of random assignment

Exhibit 9: Comparison of Child Demographic Characteristics: Head Start vs. Non-Head Start Groups, Fall 2002 Parent InterviewData

Child Characteristics	Head Start (Treatment) Group	Non-Head Start (Control) Group	Difference (Treatment-Control)
Age on 9/1/02:			
3 years old	60.2%	60.8%	-0.6%
4 years old	39.8%	39.2%	-0.6%
Gender:			
Boys	49.5%	49.4%	0.1%
Girls	50.5%	50.6%	-0.1%
Race/Ethnicity:			
White, Non-Hispanic	28.8%	28.3%	0.5%
Black, Non-Hispanic	24.9%	23.7%	1.2%
Hispanic	40.3%	40.5%	-0.2%
Other	6.1%	7.6%	-1.5%
Home Language:			
English	67.7%	64.8%	2.9%**
Spanish	28.9%	30.5%	-1.6%
Other	3.3%	4.6%	-1.3%
Disability:			
Yes	13.8%	12.4%	1.3%
No	86.2%	87.6%	-1.3%

**= $p \leq 0.05$

Note: **Data source: Fall 2002 Parent Interview.**

Exhibit 10: Comparison of Maternal Demographic Characteristics: Head Start vs. Non-Head Start Groups, Fall 2002 Parent Interview Data

Maternal Characteristics	Head Start (Treatment) Group	Non-Head Start (Control) Group	Difference (Treatment-Control)
Age:			
Under 18 years	0.1%	0.0%	0.1%
18-21 years	9.2%	9.8%	-0.6%
22-25 years	28.8%	28.0%	0.8%
26-30 years	28.7%	32.6%	-3.9%
Over 30 years	33.1%	29.6%	3.5%
Marital Status:			
Married	45.3%	48.3%	-3.0%
Separated	7.7%	7.9%	-0.2%
Divorced	8.5%	7.5%	1.0%
Widowed	0.7%	0.3%	0.4%
Never Married	37.8%	36.0%	1.8%
Race/Ethnicity:			
White, Non-Hispanic	31.0%	32.5%	-1.4%
Black, Non-Hispanic	24.6%	24.0%	0.6%
Hispanic	38.8%	38.6%	0.1%
Other	5.5%	4.9%	0.6%
Education Level:			
No High School/GED	34.8%	35.6%	-0.9%
GED	5.5%	6.5%	-1.0%
High School Diploma	28.8%	27.4%	1.4%
Some Postsecondary	22.6%	22.8%	-0.2%
Associate Degree	3.7%	4.0%	-0.3%
Bachelor's Degree	3.4%	2.7%	0.8%
Graduate Degree	1.1%	1.0%	0.1%
Employment Status:			
Full-time	32.7%	32.7%	0.1%
Part-time	16.1%	18.1%	-2.1%
Other	51.2%	49.2%	2.0%
Monthly Household Income:			
$250 or Less	3.4%	2.1%	1.3%
$251-$500	8.3%	9.0%	-0.7%
$501-$1,000	25.6%	25.0%	0.7%
$1,001-$1,500	24.2%	25.2%	-0.9%
$1,501-$2,000	17.3%	18.6%	-1.3%
$2,001-$2,500	9.9%	7.5%	2.4%
Over $2,500	11.2%	12.6%	-1.4%
Receipt of Public Assistance:			
Yes	53.8%	52.1%	1.8%
No	46.2%	47.9%	-1.8%

Notes: (1) **Data source: Fall 2002 Parent Interview;** (2) There were no statistically significant differences.

Early Care Experiences of Children Assigned to the Control Group

This section describes the preschool and child care arrangements used by children who were randomly assigned to the non-Head Start control group. It is important to emphasize that this preliminary descriptive information does not reflect the characteristics, or the preschool and child care arrangements, used by **all** low-income families. Because there is strong reason to believe that families applying to Head Start may be different from the overall population of low-income families, at least in terms of motivation to enroll their children in a preschool program prior to kindergarten, the data cannot speak to the arrangements used by low-income families overall. In addition, the data do not provide information about the overall availability of preschool and child care arrangements nationally or for all parts of the communities involved in the study.

Instead, the descriptive information specifically portrays the alternatives used by families that were eligible for, and applied to enroll their children in, Head Start but did not gain access to Head Start services. The alternatives ultimately chosen by these families depend on both parents' preferences and the availability and accessibility of various options. However, this preliminary analysis does not attempt to disentangle the extent to which the observed arrangements are a function of supply versus demand, nor does it address the quality of preschool or child care arrangements used by children in the Head Start and non-Head Start groups. These considerations will be addressed in future analyses.

Data and Methods

The information reported in this final section use data collected from parents in fall 2002 regarding where their child regularly spends time Monday through Friday and who is responsible for their child during that time. This information, as described earlier, is used to explore children's weekday arrangements in two ways: (1) the ***main daytime arrangement,*** defined as the setting where a child spends the majority of his or her time between 9am and 3pm, Monday through Friday and (2) ***all daytime weekday arrangements,*** defined as one or more non-parental arrangements used regularly between the hours of 8am and 6pm, Monday through Friday, for at least 5 hours per week (this includes the main preschool or child care arrangement).

The data are weighted to represent the population of newly entering children in communities that are not saturated with Head Start services and adjusted for survey non-response (see Chapter 2). Data are presented separately for children in the 3- and 4-year-old sample

cohorts[26] because (1) families of younger preschool-age children may use different preschool and child care arrangements than families of older preschool-age children, and (2) impacts will be assessed separately for 1-year versus 2-year participants in Head Start.

Children's Care Settings: Non-Head Start Group

As shown in the first section of Exhibit 11, parent care serves as the main arrangement in fall of 2002 for about half of the children assigned to the non-Head Start group (45 to 48 percent depending on the cohort). Center-based care is the principal alternative among children who do not have parent care as their main arrangement.[27] Overall, center-based care serves as the main arrangement for 39 to 40 percent of children assigned to the non-Head Start group in fall 2002. All other types of non-parental care, which include care in someone else's home and in the child's own home by someone other than the parent, make up the main arrangement for a total of 13 to 14 percent of children in the non-Head Start group.

[26] Children were assigned to cohorts at the time they applied to Head Start based on how local program staff classified them for the purpose of making enrollment decisions. Therefore, these cohorts may not correspond exactly to the age of children at the time they would have begun Head Start.

[27] A small number of parents of children in the non-Head Start group reported that their children were enrolled in Head Start. Preliminary data about these arrangements indicate that some of these children were enrolled by Head Start grantees selected for the study, that some of these children gained access to Head Start through grantees not selected for the study, and that some of these children were not enrolled in federally-funded Head Start program but were in a different program that parents referred to as "Head Start." Our estimate is that about 11 percent of children assigned to the control group found their way into Head Start. For this analysis, all of these children were categorized as being in a center-based child care or preschool arrangement.

Exhibit 11: Percent of Non-Head Start Study Group Children in Various Preschool and Child Care Arrangements, Fall 2002[28]

Type of Arrangement	*Main Daytime Arrangement (9am-3pm)*		*All Daytime Weekday Arrangement(s)* (8am-6pm)*	
	3-year- old cohort	*4-year-old cohort*	*3-year-old cohort*	*4-year-old cohort*
Parent Care	**45**	**48**	N/A	N/A
Non-Parental Care:	**55**	**52**	**63**	**62**
Center	40	39	42	45
Non-Relative's Home	6	5	9	7
Relative's Home	4	4	9	8
Child's Home w/Relative	4	4	7	7
Child's Home w/Non-relative	<1	<1	<1	1

* Children could be in more than one arrangement

These preliminary descriptive analyses of the fall 2002 baseline data also indicate that the main arrangements for children not admitted to Head Start do not differ substantially by age. This is a somewhat surprising finding given that research about the use of center-based programs by 3- and 4-year-olds in population-based samples tends to show that 4-year-olds are more likely than younger children to be enrolled in center-based programs.[29] Furthermore, public funding for pre-kindergarten and preschool programs is often targeted at 4-year-olds which would suggest that low-income parents of 4-year-olds might find it easier to access center-based services than parents of 3-year-olds. As a result, it was hypothesized that Head Start applicants in the 4-year-old cohort would be at least slightly more likely than the 3-year-old cohort to use a center-based alternative when they could not gain access to Head Start. However, parents of children in the 3-year-old cohort who want, but cannot access, Head Start tend to use other types of center-based care at the same rate as parents of children in the 4-year-old cohort, at least initially.

Some of the differences between the pattern of preschool and child care use for the 3- and 4-year-olds in the present study, as compared to the findings from other studies, could be due to

[28] "Main Daytime Weekday Arrangement" is the one arrangement where a child spends the most time between 9am and 3pm, Monday-Friday; "All Daytime Weekday Arrangement" includes any non-parental settings regularly used between the hours of 8am and 6pm Monday-Friday for at least 5 hours per week (multiple arrangements are allowed). For children for whom parent care is not the main arrangement, data were not collected to determine how many children regularly spend at least 5 hours per week in parent care.

[29] See: U.S. Department of Education, National Center for Education Statistics, (2002). ***The Condition of Education 2002***. Washington, D.C.: U.S. Government Printing Office (NCES Publication No. 2002-025). The Urban Institute, (2003), ***Percentage of Three- and Four-Year-Olds in Poverty in Different Types of Child Care Arrangements***. Unpublished calculations based on data from the 1999 National Survey of America's Families.

the close proximity between notification of not being able to enroll in Head Start and the fall 2002 data collection. That is, parents may not have had time to make alternative arrangements. The spring 2003 and subsequent data collection waves may show greater use of non-parental care among the non-Head Start group families.

In addition to their main arrangement, some children are enrolled in other preschool or child care arrangements during daytime hours. Moving to a consideration of *all* non-parental preschool or child care arrangements used between the hours of 8am and 6pm, Monday through Friday (for at least five hours per week), the second panel of Exhibit 11 shows that almost two-thirds of children in the non-Head Start group (62 to 63 percent) regularly use at least one non-parental child care or preschool arrangement. Again, the use of *any* center-based care is the most prevalent, with 42 to 45 percent of children using any center-based care during these hours. Use of other arrangements (i.e., care in a relative's home, a non-relative's home, or by a relative or non-relative in the child's home) is lower, with seven to nine percent of children in most of these settings for at least five hours per week (only 1 percent or fewer children spend at least five hours per week cared for by a non-relative in their own home). Consequently, even when considering the use of secondary arrangements, over one-third (37 to 38 percent) of children in the control group are exclusively in parent care during daytime hours.

Geographic Variation in Children's Settings

To determine whether the national pattern in Exhibit 11 typifies the patterns of arrangements in particular communities, the data were analyzed separately for each of the 25 geographically distinct clusters from which grantees/delegate agencies were selected for the Head Start Impact Study (see previous discussion of sample selection). These analyses show an extraordinary amount of variation in the patterns of preschool and child care arrangements from one geographic area to another. For example, the proportion of children in the non-Head Start group with parent care as their main arrangement ranges from a low of 4 percent to a high of 72 percent. The proportion of children in center-based care ranges from a low of 10 percent to a high of 93 percent.

Similar variation was observed when looking at differences in the use of *any* regular daytime arrangement across communities. For example, between 4 and 67 percent of children in the control group are exclusively cared for by their parents during daytime hours, and between 18 and 93 percent of children regularly spend at least 5 hours per week in a center-based program.

These findings, that national averages for the Head Start Impact Study's non-Head Start group generally do not typify the child care and preschool patterns in particular communities, are consistent with other research that shows a great deal of variation in the use of different types of arrangements across different communities.[30] As a result, the impacts of Head Start that will be reported in later publications will include an analysis of the extent to which impacts vary across this wide range of alternatives being used in different communities.

Comparison to Arrangements Used by a National Sample of Low-income Families

One of the key motivations for the Head Start Impact Study's reliance on equivalent Head Start and non-Head Start control groups is to ensure that measured effects of Head Start are not confounded by other family characteristics that could affect children's outcomes. For example, parents who apply to Head Start may, in general, be more motivated than other low-income parents to help their children prepare for kindergarten and more likely to seek formal out-of-home education and socialization opportunities for their children. This hypothesis, that the population of families applying to Head Start chooses different preschool and child care arrangements than the overall population of low-income families, is supported when the arrangements used by the non-Head Start group are compared to the arrangements used by all children living in poverty.

To carry out this comparison, data from the 1999 National Survey of America's Families (NSAF) were used to estimate the use of any non-parental care, and of different types of non-parental care, for at least 5 hours per week by 3- and 4-year-olds living in poverty nationally. NSAF data differ from Head Start Impact Study data in a few key ways that likely account for some, but probably not all, of the observed differences in the arrangements used by families in the Head Start Impact Study. First, NSAF data reflect arrangements used at any time during the entire week while data for the Head Start Impact Study only reflect arrangements used, at least in part, during weekday hours. Second, the methods used to sort children by age differ across the two studies, i.e., NSAF age groupings are based on the age of the child at the time of the

[30] Sonenstein, Freya, Gary Gates, Stefanie Schmidt, and Natalya Bolshun.. (2003), ***Primary Child Care Arrangements of Employed Parents: Findings from the 1999 National Survey of America's Families.*** Washington, D.C.: The Urban Institute. (Assessing the New Federalism, Occasional Paper No. 55.) O'Neil, Grace and Martin O'Connell (2001). ***State Estimates of Child Care Establishments: 1977 – 1997.*** Washington, D.C.: U.S. Census Bureau Population Division. (Working Paper Series No. 55).

parent/primary caregiver interview while Head Start Impact Study age groupings are based on the number of years children were expected to participate in Head Start before entering kindergarten.

Third, data for the Head Start Impact Study were collected in 2002, while NSAF data were collected in 1999. Although the effect of the time difference on the results is difficult to predict, there is evidence that rates of use of different child care arrangements have shifted over time[31] which may account for some differences in the two populations. Finally, NSAF data are representative of families living in poverty nationwide, while Head Start Impact Study data are nationally representative of the population of children applying to enroll in Head Start for the first time in communities where there are more Head Start applicants than available spaces.

Despite these differences, a comparison of the arrangements used by families in the Head Start Impact Study's control group against the arrangements used by all low-income families not reporting Head Start use shows two important contrasts, especially for the 3-year-olds (the patterns appear largely the same among 4-year-olds in the two populations). Compared to the overall population of children living in poverty, a higher percentage of 3-year-olds in the control group are in center-based arrangements. In addition, children in the Head Start Impact Study's 3-year-old cohort are less likely than low-income children overall to be in any relative care.

Although preliminary, these findings support the idea that families that apply to Head Start, especially when their children are 3 years old, are specifically seeking, and are more successful at finding, a center-based preschool or child care arrangement than low-income families as a whole. In other words, we expected parents who apply to Head Start to be different from other low-income parents, and these preliminary analyses support this hypothesis. This will, as discussed below, have implications for the later estimates of the impact of Head Start, and serves as a further reminder that data from the Impact Study cannot be generalized to describe all low-income children and their families.

---

[31] See: U.S. Department of Education, National Center for Education Statistics, (2002), ***The Condition of Education 2002***. Washington, D.C.: U.S. Government Printing Office (NCES Publication No. 2002-025). Sonenstein, Freya, Gary Gates, Stefanie Schmidt, and Natalya Bolshun., (2003), ***Primary Child Care Arrangements of Employed Parents: Findings from the 1999 National Survey of America's Families***. Washington, D.C.: The Urban Institute**.** (***Assessing the New Federalism,*** Occasional Paper No. 55.)

Implications

These preliminary descriptive analyses show that most families tend to turn to one of two care alternatives in the initial few months immediately after notification that their child will not be able to enroll in Head Start. Nationally, parent care is the alternative for about half of the children whose families wanted Head Start but could not gain access to the program, with center-based programs providing the usual alternative to Head Start for children who are not in parent care. The proportion of children cared for in non-parental home-based settings by a relative or non-relative is generally low. Within and across communities, the results also show substantial variation in the proportion of control group families that use each type of arrangement.

These results make it clear that the impacts of Head Start will not be evaluated against a pure "no-services" alternative. Rather, they will be assessed against a mixture of alternatives ranging from parent care to center-based programs, of which some may look very much like Head Start and some may look very different from Head Start. Although all types of alternatives, including parent care, may offer an environment that effectively supports children's development, parent care and center-based programs may be thought of as falling on opposite ends of a continuum in terms of the likelihood that the environment delivers a set of services and experiences that is similar to Head Start.

Based on the fall 2002 baseline data collection, children in the control group tend to be concentrated at the two ends of this continuum. Furthermore, the proportion of children using each type of arrangement varies dramatically from one community to another, with children in some communities evenly split between parent care and center-based arrangements and children in other communities more heavily concentrated in either parent care or center-based programs. This reality will complicate the interpretation of Head Start impact estimates that are based on comparing the average outcomes for children in Head Start to the average outcomes of control group children who spend their time in a wide variety of different arrangements.

Among control group children who are mainly in parental care, the analyses also show that approximately one in five children are enrolled in another arrangement for at least 5 hours per week. Because these other arrangements may offer children some of the same out-of-home enrichment opportunities as Head Start, it will be important to consider whether exposure to *any* group setting is related to outcomes of children in the control group. Consequently, impact analyses that attempt to understand how different alternatives to Head Start are related to

variations in the impact of Head Start from one community to another will need to consider *main* arrangements as well as use of *any* type of non-parental preschool or child care. The impact analyses will also need to account for potential changes in the types of care settings individual children use, especially as other alternative care settings may become more available over time.

These preliminary descriptive data on the preschool and child care arrangements for children in the non-Head Start group also indicate that families that apply for Head Start have different patterns of preschool and child care use than the overall population of families in poverty, at least in the few months immediately after not being able to enroll their child in Head Start. In particular, it appears that families that apply to Head Start when their children are 3 years old are more likely to access center-based preschool or child care than low-income families as a whole, even when they cannot gain access to Head Start. This finding underscores the importance of conducting research on the impacts of Head Start using a rigorous random-assignment methodology.

Head Start services are delivered on the premise that a comprehensive, high quality early childhood program can make a difference in child and family outcomes. As a result, basic information about the preschool and child care arrangements of children assigned to the non-Head Start group provides important background for the impact findings to be reported later. It is likely that greater access to good quality, comprehensive early childhood programs among children in the control group will be related to smaller differences in outcomes between these children and children assigned to the Head Start treatment group. Although the preliminary results reported here do not speak to the type or quality of services received in each type of arrangement, future reports will build on these findings to understand the extent to which children in the control group are in arrangements that are similar to Head Start, in terms of both the type and the quality of the services received.

Chapter 5: Study Status

The Head Start Impact Study began in October 2000 with initial activities focused on developing the study design, recruiting Head Start grantee/delegate agencies, and field testing study measures and procedures. The selection and recruitment of participating Head Start centers, families, and children continued through spring 2002, culminating with the collection of child, family, and program data in fall 2002 and spring 2003. Activities are on schedule for the planned final report of findings in December 2006. Accomplishments to date include:

- During the 2000-01 Head Start year, conducted an advance field test involving eight Head Start grantees, 21 Head Start centers, and approximately 400 children to test the feasibility of all study procedures.
- During the 2001-02 Head Start year, completed the selection and recruitment of a national probability sample of 84 grantee/delegate agencies and 378 Head Start centers to participate in the national study.
- During summer 2002, completed random assignment, in all recruited centers, resulting in 4,750 children being randomly assigned to either a Head Start treatment group or a non-Head Start control group.
- Completed fall 2002 data collection, with high overall response rates for both child assessments and parent interviews with a combined rate of 80 percent for fall 2002.
- Completion of data entry and checking for errors, as well as statistical weighting of the fall 2002 data. The initial assessment of the psychometric properties of the child assessments is currently underway.
- Completion of spring 2003 data collection with approximately 83 percent response rates for both child assessments and parent interviews.
- Have started data entry and checking for errors, as well as statistical weighting. Initial assessment of psychometric properties for the spring 2003 data collection will continue on an ongoing basis following the completion of this process for the fall 2002 data.
- A meeting was held of the Advisory Committee on Head Start Research on June 16 and 17, 2003, to update the committee on the progress of the study activities and to seek additional input and recommendations on the remaining work of the study. A summary of the Advisory Committee meeting can be found at http://www.acf.hhs.gov/programs/hsb/research/hsreac/index.htm

Overall, Advisory Committee members commented positively on the research design of the Impact Study and the success of implementation to date, especially given the complexity of the Congressional mandate and the ethical concerns of randomly assigning low-income children to a control group. Several members noted that when completed, the Impact Study will provide a rich array of data that will not only meet the Congressional requirements, but will also allow for an extensive range of secondary data analyses.

A final report of the study findings is scheduled for December 2006.

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: Analysis of Programs Excluded During Sampling

Comparison of Saturated and Non-saturated Head Start Programs

The frame used to select Head Start grantee/delegate agencies initially contained all the programs on the PIR in each of the 25 clusters (referred to as primary sampling units, or PSUs) that were determined to be eligible for the impact study based on PIR data. The total number of programs in the selected 25 PSUs was 355. Prior to sampling Head Start programs in each PSU, we further screened the programs by telephone to determine their eligibility for the study. To reduce the amount of telephone screening in three very large clusters, programs were sub-sampled resulting in a reduction of 86 programs; eight FACES programs were also deleted to avoid sampling them for the Head Start Impact Study. This resulted in a list of 261 grantee/delegate agencies that were screened for inclusion in the study. During this screening process, 28 programs were determined to be saturated and 10 programs were found to have closed or had merged with another program. After eliminating these 38 programs, the final frame consisted of 223 programs in the 25 PSUs.

The 28 saturated programs represent an estimated 12 percent [5.3%, 19.0%][32] of the eligible Head Start program universe, and approximately 4.3 % [2.2%, 6.4%] of the total Head Start age 3-4 year old enrollment and 3.9% [2.0%, 5.7%] of the newly-entering 3-4 year old Head Start enrollment. The 10 closed programs represent 60 closed programs on the PIR frame containing 3.0% of the total enrollment and 2.7% of first year enrollment. The total enrollment data used in these calculations were taken from the 1999-2000 PIR, and the newly-entering enrollment data came from the telephone screening.

There is potential for under-coverage bias due to the exclusion of saturated Head Start programs from the sampling frame. Newly-entering Head Start children in these saturated programs had no chance of selection for the study and therefore are not represented in the study sample. The potential for bias arises if the saturated programs are systematically different from the non-saturated programs retained on the frame, and if the characteristics on which they differ

[32] The 95% confidence intervals are given in brackets, as calculated from standard errors produced by Wesvar using jackknife replication.

are correlated with the outcome measures for the children they enroll. If the children in these excluded programs represent only a small percentage of the Head Start population, then the potential for bias is much less.

Tables B.1 and B.2 compare saturated and non-saturated programs by the few qualitative characteristics and enrollment variables available from the PIR. The programs were weighted to account for the PSU stage of sampling and for the sub-sampling of programs in three very large clusters prior to the telephone screening. This is necessary to draw conclusions about the entire Head Start population and not merely programs in the 25 sampled PSUs. Tests of statistical significance were done using WesVar with jackknife replicate weights to account for the Head Start Impact Study sample design.

Tables B.1 and B.2 show that the saturated programs are smaller, more likely to be school-based, and have smaller percent Hispanic enrollment than the non-saturated programs. Although they appear to be more often located in the Midwest, differences in the distribution of saturated vs. non-saturated programs by Head Start regions are not statistically significant. A cautionary note is that variances at the program level are not very stable because the number of saturated programs is small. In addition, they do not include the between-PSU component of variance due to sampling PSUs[33], thus they are underestimates and the p-values may be overstating the statistical significance of the differences.

[33] These adjustments will be incorporated for the spring 2003 replicate weights.

Table B.1. Comparison of Saturated and Non-saturated Head Start Programs by Characteristics of Enrolled Children

Characteristic	Saturated Programs	Non-saturated Programs	P-value for t-test of difference in means or proportions *
Percent Hispanic Enrollment	9%	26%	**.001**
Percent Black Enrollment	20%	33%	.134
Percent Age 3 Enrollment	52%	49%	.535
Percent First Year Age 3 Enrollment	79%	87%	.222
Average Total Enrollment	188	571	**<.001**
Average Newly-Entering Children	113	388	**<.001**

*** Note: bold print indicates statistically significant difference.**

Table B.2. Comparison Of Saturated And Non-saturated Head Start Programs By Location

Characteristic	Saturated Programs	Non-saturated Programs	P-value for Chi-square test of association *
School-based			**.02**
Yes	66%	21%	
No	34%	79%	
Metro Status			.91
Urban (MSA)	66%	68%	
Rural (Non-MSA)	34%	32%	
Level of Pre-K Services in State			.60
Similar to Head Start	35%	25%	
Some Head Start-like	27%	20%	
Remaining States	38%	55%	
Head Start Region			.15
Northeast	24%	25%	
Midwest	48%	24%	
South	28%	39%	
Plains	0%	4%	
West	0%	8%	

*** Note: bold print indicates statistically significant difference.**

Comparison of Saturated and Non-saturated Head Start Centers

Enrollment data were also obtained for all of the operating Head Start centers in each sampled grantee/delegate agency using a standardized data collection form (called the Center Information Form or CIF). This information was used to determine the level of saturation for each center. Centers that would clearly be unable to provide a sample of non-Head Start control group

children were deleted from the frame prior to sampling; three sampled grantees were dropped at this stage because all of their centers were determined to be saturated.

This resulted in the deletion of 154 centers from a frame of 1,423, in addition to the seven centers in the three saturated programs. These 154 centers represent an estimated eight percent [4.7%, 11.4%] of the Head Start center universe, approximately 5.1 percent [3.0%, 7.1%] of the total Head Start age 3-4 enrollment, and 4.7 percent [2.7%, 6.8%] of newly-entering children. Overall, the estimated coverage rate of the newly-entering Head Start population for the Head Start Impact Study sample is 92 percent [89%, 95%], which is obtained by multiplying the coverage rate for programs and centers [(100-3.9% for programs)*(100- 4.7% for centers)], and the overall undercoverage rate is eight percent [5%, 11%].

Tables B.3 and B.4 compare saturated and non-saturated centers by various characteristics and by characteristics of enrolled children. As with grantee/delegate agencies, hypothesis testing was done in WesVar using jackknife replicate weights to account for the study sample design. The replicate weights do not yet incorporate the between-PSU variance component, therefore the p-values in these tables may be too liberal. In Table B.3 the chi-square test was not able to detect a significant difference with respect to type of program option offered, whether the program is based in a school, metro status, region, or level of Pre-K services available in the state. With respect to enrollment, Table B.4 shows that the saturated centers are smaller, less Hispanic, and have a larger percentage of newly-entering 3-year-olds than the non-saturated centers.

The saturation rate was calculated two ways: as the percent of centers in each program that are saturated, and as the percent of newly-enrolled children in saturated centers in each program. The average percent of saturated centers is 16.6 percent and ranges from 0 to 84 percent. The average percent of newly-entering children in saturated centers is only 13.2 percent and ranges from 0 to 69 percent, another indication that the saturated centers tend to be smaller.

Table B.3. Comparison Of Saturated And Non-saturated Head Start Centers By Center Characteristics

Characteristic	Saturated Programs	Non-Saturated Programs	P-value for Chi-square Test of Association
Program Option			.44
Full-Day Only	35%	28%	
Part-Day Only	52%	50%	
Other	13%	22%	
School-Based			.25
Yes	17%	11%	
No	83%	89%	
Metro Status			.64
MSA	74%	70%	
Non-MSA	26%	30%	
Head Start Region			.38
Northeast	32%	27%	
Midwest	34%	20%	
South	17%	31%	
Plains	12%	11%	
West	4%	11%	
Level of Pre-K Services in State			.21
Similar to Head Start	40%	22%	
Some Head Start-like	15%	18%	
Remaining States	45%	60%	

Table B.4. Comparison Of Saturated And Non-saturated Head Start Centers By Characteristics of Enrolled Children

Characteristic	Saturated Programs	Non-Saturated Programs	P-value for t-test of difference in means or proportions*
% Hispanic Enrollment	17%	30%	**.005**
% Black Enrollment	38%	26%	.204
% Newly-Entering Children	65%	66%	.985
% Newly-Entering 3-Year-Olds	54%	47%	.037
% on Waiting List	0%	15%	**<.001**
Average Number of Funded Slots	37	48	.036
Average Total Enrollment	26	47	**<.001**
Average Newly-Entering Children	16	31	**<.001**
Average Number on Waiting List	0	9	**<.001**

*** Note: bold print indicates statistically significant difference.**

Appendix C: Comparison of Fall 2002 Respondents and Non-Respondents

Non-response analysis for the child assessment and parent/primary caregiver interview survey instruments was conducted to determine the variables most correlated with response propensity for each instrument, i.e., to determine the factors that differentiate respondents from nonrespondents. Three definitions of response for each child were applied: 1) Child Assessment completed, 2) Parent Interview completed, or 3) both completed. The seven variables examined were the child's age, gender, race/ethnicity, region, metro status, program option applied for, and treatment vs. control group assignment. The source of these variables was the random assignment roster that was completed by Westat field staff for every applicant seeking to enroll in the sampled Head Start centers; the roster, therefore, contained data for both respondents and nonrespondents.

Weighted and unweighted response rates for each of the three definitions are given in Tables C.1a and C.1b, where the weight is the child base weight that accounts for all stages of sampling and permits the treatment and control samples to represent the entire newly-entering Head Start population. The tables show that the response rates for children assigned to the non-Head Start control group are below those for children assigned to the Head Start treatment group, by about 10-12 percentage points on average. Most of the difference is due to the greater difficulty in locating control group children, since they do not attend Head Start. It is also interesting to note that although the overall average response rate is approximately 80 percent, there is wide variation among grantee/delegate agencies. The response rates at the great majority of grantee/delegate agencies range between 60 and 100 percent, while a small number of grantee/delegate agencies (6-10) have rates ranging from 60 percent to 25 percent for the fall 2002 data collection.[34]

To analyze differences between respondents and nonrespondents, bivariate chi-square tests of association between response status and each of the variables age, gender, region, metro status, program option applied for, and race/ethnicity were performed separately for the treatment and control groups. A logistic regression model was also run with the response status as the dependent variable (1=respondent, 2=nonrespondent) and the seven variables mentioned above as

[34] The actual number of grantee/delegate agencies with these low rates differs slightly depending upon the respondent and type of instrument administered.

predictors. Both chi-square and logistic regression analyses were done using the child base weights and WesVar software to correctly account for the weights and the stratification and clustering in the Head Start Impact Study sample design. To prevent extremely large child weights from having undue influence on the results, 29 child base weights exceeding 1000 (0.61% of the sample) were truncated at 1000 and the analyses were rerun using this trimmed weight. The p-values for the chi-square tests are given in Table C.2 below and for the logistic regression analysis in Tables C.3 and C.4.

The Chi-square tests showed that aside from treatment vs. control status, only program option applied for within the control group was consistently significant across the three different response definitions. Response rates for control children applying for the Full-Day option are lower than for those applying for the Part-Day, which may be an indication that applicants to the Full-Day option tend to be parents who work full-time and are less available for the Parent Interview or Child Assessment. It may also indicate a higher refusal rate from those who wanted a full time option and were assigned to the control group. It is also worth pointing out that program report on this particular variable was inconsistently worded across sampled centers. As these are preliminary results, further analysis is ongoing. Head Start region was significant with the original weight but not the trimmed weight, which implies that the significance might have been caused by large outlier weights, and was not a true difference in response rates among the regions. Response rates also differed by gender for Child Assessment completes with the original weight but not the trimmed weight, again suggesting that the difference was caused by a few influential weights having extremely high values. As can be seen in Tables C.1a and C.1b, response rates for girls are only slightly higher than for boys.

In the logistic regression analysis, as expected, the control/treatment indicator variable was highly significant in predicting the propensity to respond for all three response definitions. Program option and gender were moderately significant predictors with the original child base weight, but not for the trimmed weights.

Table C.1a. Fall 2002 Parent Response Rates Based On Selected Characteristics For Children In The Head Start Treatment Group*

Child characteristics	Response rates based on:					
	Child Assessment completes		Parent Interview completes		Both instruments complete	
	Unweighted	Weighted	Unweighted	Weighted	Unweighted	Weighted
TREATMENT GROUP	86.04	86.82	86.31	87.52	84.66	85.73
Child's Gender						
Boys	84.89	85.20	85.71	86.66	83.71	84.28
Girls	87.27	88.64	87.07	88.60	85.75	87.39
Agency Region						
Northeast	77.51	80.56	77.83	82.01	74.93	78.63
Midwest	88.25	86.86	88.02	86.48	86.87	85.48
South	89.35	88.61	89.63	89.63	88.39	87.81
West	88.63	89.77	89.27	89.96	87.98	89.16
Agency Metro Status						
MSA	86.04	87.12	86.18	87.78	84.55	86.04
NonMSA	86.07	85.88	86.87	86.72	85.11	84.77
Program Option Applied For						
Full-Day	84.08	84.86	83.61	85.36	82.10	83.29
Part-Day	88.37	88.76	88.82	89.27	87.34	87.88
Both full day and part-day	71.43	75.51	79.59	82.96	71.43	75.51
Other**	85.34	87.41	87.93	88.33	85.34	87.41
Child's Race-ethnicity						
Hispanic	84.97	85.63	86.43	87.27	83.87	84.84
Black	86.23	87.35	85.43	87.35	84.36	85.98
White	87.96	88.90	87.70	88.92	86.60	87.53
Other	78.95	78.50	79.31	78.50	78.95	78.50
Child's Age						
Age 3	86.89	87.05	87.09	88.08	85.30	85.69
Age 4	85.01	86.56	85.33	86.89	83.87	85.78
Parent Language						
English	86.84	87.86	86.54	88.05	85.28	86.60
Spanish	85.09	85.93	86.76	87.80	84.06	85.17
Other**	63.89	55.46	55.46	64.86	63.89	55.46

* Data Source: Roster information used at time of random assignment
** These categories have very small cell sizes.

Table C.1b. Fall 2002 Parent Response Rates Based On Selected Characteristics For Children In The Non-Head Start Control Group *

Child characteristics	Response rates based on:					
	Child Assessment completes		Parent Interview completes		Both instruments complete	
	Unweighted	Weighted	Unweighted	Weighted	Unweighted	Weighted
CONTROL GROUP	73.74	76.27	75.62	77.43	73.45	75.71
Child's Gender						
Boys	72.27	74.32	74.59	76.25	72.16	73.91
Girls	75.22	78.02	76.67	78.47	74.73	77.30
Agency Region						
Northeast	68.10	72.41	69.76	72.48	67.39	70.03
Midwest	75.09	75.27	77.14	77.00	75.36	75.50
South	72.98	74.95	75.32	76.73	72.73	74.90
West	82.46	82.32	83.38	83.18	85.15	82.13
Agency Metro Status						
MSA	72.82	75.69	74.93	76.92	72.45	74.94
NonMSA	77.71	78.08	78.57	79.01	77.71	78.08
Program Option Applied For						
Full-Day	67.46	69.46	69.46	70.94	66.83	68.22
Part-Day	78.70	81.35	80.70	82.52	78.76	81.36
Both full day and part-day	67.74	65.07	64.52	61.98	64.52	61.98
Other**	80.49	82.77	81.71	83.19	80.49	82.77
Child's Race-ethnicity						
Hispanic	74.48	77.09	76.72	78.83	74.07	76.72
Black	71.58	75.60	72.86	74.88	70.86	73.82
White	74.58	74.82	76.55	76.81	74.86	75.00
Other	78.57	86.41	80.95	86.96	78.57	86.41
Child 'sAge						
Age 3	73.58	75.88	76.08	77.26	73.13	74.89
Age 4	73.95	76.72	75.03	77.63	73.83	76.64
Parent Language						
English	73.07	74.74	75.08	76.35	73.02	74.79
Spanish	75.78	79.73	77.56	80.56	75.00	78.43
Other**	64.71	72.98	58.82	55.43	58.82	55.45

* Data Source: Roster information used at time of random assignment

** These categories have very small cell sizes.

Table C.2: P-Values for Chi-Square Tests of Association between Response Status and Child Variables, Calculated Separately for Treatment and Control Groups*

Child characteristics by Control/treatment group	Response status based on:					
	Child Assessment completes		Parent Interview completes		Both completes	
	Original weight	Trimmed weight **	Original weight	Trimmed weight **	Original weight	Trimmed weight **
Child Gender						
treatment	.034	.450	.231	.423	.121	.492
control	.176	.220	.469	.465	.337	.237
Agency Region						
treatment	.016	.209	.058	.179	.030	.125
control	.120	.429	.113	.432	.155	.362
Agency Metro Status						
treatment	.729	.733	.740	.694	.802	.740
control	.663	.667	.681	.676	.621	.556
Program Option Applied For						
treatment	.515	.573	.588	.592	.609	.532
control	.003	**.016**	.007	**.042**	.014	**.023**
Child Race-ethnicity						
treatment	.635	.677	.631	.764	.826	.768
control	.526	.686	.403	.620	.612	.675
Child Age						
treatment	.859	.837	.516	.700	.954	.972
control	.729	.793	.822	.921	.560	.632

* Data source: Roster information used at time of random assignment.

** Note: bold print indicates statistically significant difference.

Table C.3: P-Values for T-Tests of Significance for Logistic Regression Parameters for Combined Treatment and Control Groups*

Predictor Variables	Response status based on:					
	Child Assessment completes		Parent Interview completes		Both completes	
	Original weight	Trimmed weight **	Original weight	Trimmed weight **	Original weight	Trimmed weight **
Intercept	0.003	0.003	0.004	0.018	0.010	0.002
Child's Age is 3	0.247	0.248	0.306	0.487	0.501	0.404
Child's Gender is female	0.052	0.314	0.062	0.141	0.168	0.310
Agency Region 1 -Northeast vs. West	0.048	0.069	0.267	0.390	0.056	**0.048**
Agency Region 2 -Midwest vs. West	0.182	0.185	0.641	0.710	0.348	0.256
Agency Region 3 -South vs. West	0.285	0.289	0.965	0.967	0.552	0.465
Agency Metro Status -Rural	0.812	0.871	0.442	0.527	0.869	0.891
Program Option Applied For 1 -Full-day vs. Other	0.347	0.380	0.855	0.895	0.402	0.337
Program Option Applied For 2 -Part-day vs. Other	0.882	0.890	0.446	0.528	0.954	0.952
Program Option Applied For 3 -Full- or Part-day vs. Other	0.421	0.415	0.817	0.909	0.509	0.413
Child Race-Ethnicity – Hispanic vs. White	0.721	0.694	0.957	0.942	0.768	0.692
Child Race-Ethnicity – Black vs. White	0.331	0.540	0.994	0.980	0.629	0.659
Child Race-Ethnicity – Other vs. White	0.925	0.974	0.132	0.268	0.903	0.933
Child is in treatment group	0.000	**0.006**	0.000	**0.000**	0.000	**0.003**

* Data Source: Roster information used at time of random assignment.

** Note: bold print indicates statistically significant difference.

Table C.4. P-Values for F-Tests of Overall Significance for Logistic Regression Variables for Combined Treatment and Control Groups*

Predictor Variables	Response status based on:					
	Child Assessment completes		Parent Interview completes		Both completes	
	Original weight	Trimmed weight **	Original weight	Trimmed weight **	Original weight	Trimmed weight **
Overall Fit	0.000	0.000	0.000	0.000	0.000	0.000
Child's Age (2)	0.247	0.248	0.132	0.268	0.501	0.404
Child's Gender (2)	0.052	0.314	0.306	0.487	0.168	0.310
Agency Region (4)	0.256	0.316	0.226	0.451	0.261	0.250
Agency Metro Status (2)	0.812	0.871	0.965	0.967	0.869	0.891
Program Option Applied For (4)	0.029	0.114	0.046	0.173	0.093	0.114
Child's Race-Ethnicity (4)	0.701	0.797	0.994	0.998	0.936	0.910
Treatment/Control (2)	0.000	**0.006**	0.000	**0.000**	0.000	**0.003**

* Data source: Roster information used at time of random assignment.

** Note: bold print indicates statistically significant difference.

Appendix D: Comparison of Fall 2002 for Head Start Treatment and Non-Head Start Control Group Children: Weighted and Un-weighted Data

To allow readers to understand the effect of weighting on the subsequent analysis of the impact of Head Start, Table D1 compares the Head Start treatment and non-Head Start control groups using three different weights: (1) using unweighted data; (2) weighted to reflect each child's overall probability of selection (Child Base Weights); and (3) using weights that are adjusted, or raked, [35] for nonresponse to the Parent Interview to restore the weighted distribution to that of the original sample (Final Child Weights).

In Table D1, the distribution of Parent Interview respondents ("Final Child Weights") closely matches that of the original sample ("Child Base Weights") with the exception of the percent Hispanic/Spanish speaking and percent White/English speaking children. This is because the raking adjustment was used to improve the reliability of the survey estimates and to correct for bias due to undercoverage. The weights of Hispanic children was reduced and those of White children slightly increased, so that the weighted sample distribution matches the PIR with respect to total enrollment by race/ethnicity. However, **none** of the differences in Exhibit D1 are statistically significant ($p > .13$ for all comparisons).

[35] These weights have been raked to PIR total enrollment by race/ethnicity, and trimmed to reduce the impact on variances of a few extremely large outlier weights. Raking or iterative proportional fitting is when weights are consecutively ratio-adjusted to marginal control totals of the population until the resulting weights converge to the totals for each dimension. In other words, raking ensures that the estimated totals match as closely as possible to known population totals.

Table D1: Comparison of Head Start and Non-Head Start Study Groups: Child and Family Characteristics Measured at Fall 2002 Parent Interview

Characteristic	Treatment Group (%)			Control Group (%)			Treatment and Control Groups Combined (%)		
	Unweighted	Child Base Weights	Final Child Weights[1]	Unweighted	Child Base Weights	Final Child Weights[1]	Unweighted	Child Base Weights	Final Child Weights1
Child's Age[2]:									
3-Year Old	60.9	59.6	60.1	61.0	59.5	60.7	60.9	59.5	60.4
4-year-old	38.8	40.3	39.8	38.9	40.4	39.2	38.9	40.3	39.5
Missing	0.3	0.1	0.1	0.1	0.1	0.1	0.2	0.1	0.1
Child's Gender:									
Boys	50.1	49.8	49.5	50.0	48.5	49.4	50.1	49.2	49.5
Girls	49.9	50.2	50.5	50.0	51.5	50.6	49.9	50.8	50.5
Child's Race/Ethnicity:									
Hispanic	41.1	44.3	40.2	41.2	46.0	40.5	41.1	45.1	40.3
Black	28.6	25.0	24.8	27.3	23.3	23.7	28.1	24.2	24.2
White	24.4	25.3	28.7	24.9	23.5	28.3	24.6	24.5	28.5
Other	5.7	5.2	6.1	6.6	7.2	7.6	6.0	6.1	6.8
Missing	0.3	0.2	0.2	0.0	0.0	0.0	0.2	0.1	0.1
Home Language:									
English	66.2	64.2	67.0	65.5	59.9	64.0	65.9	62.2	65.5
Spanish	30.4	31.9	28.6	30.1	34.8	30.1	30.3	33.3	29.4
Other	2.5	3.0	3.3	3.2	4.1	4.6	2.8	3.5	3.9
Presence of Child Disability:									
Yes	14.2	13.5	13.7	11.8	12.5	12.4	13.3	13.0	13.1
No	85.8	86.5	86.3	88.2	87.5	87.6	86.7	87.0	86.9
Mother's Age									
Less than 18-years old	0.2	0.1	0.1	0.1	0.0	0.0	0.1	0.1	0.1
18 – 21 years old	10.0	8.9	9.1	10.9	9.1	9.5	10.4	9.0	9.3
22 – 25 years old	29.5	28.4	28.2	26.9	27.3	27.1	28.5	27.9	27.7
26 – 30 years old	27.9	28.0	28.2	30.5	31.6	31.6	28.9	29.7	29.9
31 years or older	31.1	32.9	32.5	29.0	28.7	28.7	30.3	31.0	30.6
Missing	1.3	1.7	1.8	2.6	3.2	3.1	1.8	2.4	2.4
Mother's Race									
Hispanic	39.8	42.7	38.5	39.4	44.7	38.5	39.6	43.6	38.5
Black	28.5	24.7	24.5	26.8	23.3	23.9	27.9	24.0	24.2
White	27.1	27.5	30.9	29.2	27.2	32.4	27.9	27.3	31.6
Other	4.3	4.6	5.5	4.1	4.5	4.9	4.3	4.6	5.2
Missing	0.3	0.5	0.6	0.4	0.3	0.3	0.4	0.4	0.5

Table D1: Comparison of Head Start and Non-Head Start Study Groups: Child and Family Characteristics Measured at Fall 2002 Parent Interview (continued)

Characteristic	Treatment Group (%)			Control Group (%)			Treatment and Control Groups Combined		
	Unweighted	Child Base Weights	Final Child Weights	Unweighted	Child Base Weights	Final Child Weights	Unweighted	Child Base Weights	Final Child Weights
Mother's Education									
No HS Diploma or GED	36.4	35.3	34.5	37.8	36.4	35.3	36.9	35.8	34.9
GED	6.0	5.0	5.5	6.1	6.1	6.4	6.0	5.6	6.0
HS Diploma	27.4	28.4	28.6	26.1	28.1	27.1	26.9	28.2	27.8
Some Postsecondary, No Degree	21.6	22.0	22.4	20.9	21.3	22.5	21.3	21.7	22.5
Associate's Degree	3.7	3.6	3.7	3.9	3.6	4.0	3.8	3.6	3.8
Bachelor's Degree	3.4	3.8	3.4	2.8	2.7	2.6	3.2	3.3	3.0
Graduate school/degree	0.8	1.0	1.1	0.9	0.8	1.0	0.9	0.9	1.0
Missing	0.8	0.9	0.9	1.4	1.1	1.2	1.0	0.9	1.0
Mother's Employment Status									
Full-time	31.7	31.4	30.8	30.9	31.2	31.1	31.4	31.3	31.0
Part-time	14.8	14.9	15.1	15.7	18.0	17.2	15.1	16.4	16.1
Other	47.8	48.1	48.1	47.7	46.1	46.7	47.8	47.2	47.4
Missing	5.7	5.6	5.9	5.8	4.7	4.9	5.7	5.2	5.4
Mother's Marital Status									
Married	42.9	44.6	44.9	45.3	48.0	48.0	43.8	46.2	46.4
Separated	8.1	7.7	7.7	8.6	7.9	7.9	8.3	7.8	7.8
Divorced	8.1	7.9	8.4	7.4	7.5	7.4	7.8	7.8	7.9
Widowed	0.6	0.7	0.7	0.3	0.3	0.3	0.5	0.5	0.5
Never Married	39.5	38.0	37.4	37.7	35.5	35.7	38.8	36.8	36.6
Missing	0.8	1.1	0.9	0.8	0.7	0.8	0.8	0.9	0.8
Monthly Household Income									
$250 or less	2.5	2.4	2.3	2.0	1.5	1.4	2.3	2.0	1.9
251 – 500	6.7	5.9	5.7	7.1	6.0	6.1	6.8	5.9	5.9
501 – 1000	17.8	18.0	17.6	16.2	17.9	16.9	17.2	17.9	17.2
1001 – 1500	16.5	16.5	16.6	18.6	16.3	17.0	17.3	16.4	16.8
1501 – 2000	11.8	12.1	11.8	11.6	12.6	12.6	11.7	12.3	12.2
2001 – 2500	6.1	6.5	6.8	4.6	5.4	5.1	5.5	6.0	6.0
More than $2500	6.9	7.6	7.7	7.6	8.3	8.6	7.2	7.9	8.1
Missing	31.7	31.1	31.5	32.4	32.0	32.2	32.0	31.5	31.8
Receives Public Assistance?									
Yes	55.2	53.7	53.5	56.4	51.8	51.9	55.6	52.8	52.7
No	44.4	45.8	45.9	43.1	47.8	47.8	43.9	46.8	46.8
Missing	0.4	0.5	0.6	0.5	0.3	0.3	0.5	0.4	0.5

[1] Distribution based on children with completed parent interviews.

[2] 3-year-olds includes a few 2-year-olds and 4-year-olds includes a few 5-year-olds

Appendix E: Citations for Child Assessments, Scales, and Observation Instruments

CHILD ASSESSMENT BATTERY

Dunn, L.M., Dunn, L.L., and Dunn, D.M. (1997). *Peabody Picture and Vocabulary Test, Third Edition (PPVT).* Circle Pines, MN: American Guidance Service.

Dunn, L.M., Padilla, E.R., Lugo, D.E., and Dunn, L.M. (1986). *Test de Vocabulario en Imagenes Peabody.* Circle Pines, MN: American Guidance Service.

Leiter-R AM Battery (1997). Wood Dale, IL: Stoelting Co. (Subtest: Attention Sustained).

Lonigan, C.J., Wagner, R.K., Torgesen, J. K., and Rashotte, C. (2002). *Preschool Comprehensive Test of Phonological & Print Processing.* (Subtests: Print Awareness and Elision).

McCarthy, D. (1970, 1972). *McCarthy Scales of Children's Abilities.* San Antonio, TX: The Psychological Corporation. (Subtest: Draw-a-Design Task).

Woodcock, R. W., McGrew, K.S., and Mather, N. (2001). *Woodcock-Johnson III Tests of Achievement.* Itasca, IL: Riverside Publishing. (Subtests: Letter-Word Recognition, Spelling, Oral Comprehension, and Applied Problems).

Woodcock, R.W. and Munoz-Sandoval, A.F. (1996). *Bateria Woodcock-Munoz Pruebas de aprovechamiento-Revisada.* Itasca, IL: Riverside Publishing. (Subtests: Identificacion de letras y palabras, Dictado, and Problemas Aplicados).

Story and Print Concepts, Color Names and Counting, and Letter Naming Task - Developed for the FACES project.

TEACHER/CARE PROVIDER CHILD REPORT

High Scope Educational Research Foundation (1992). *Child Observation Record (COR).* Ypsilanti, MI: High Scope Educational Research Foundation.

Lutz, M.N., Fantuzzo, J.F., and McDermott, P. (in press). *Adjustment Scales for Preschool Intervention.*

Pianta, R.C. (1992). *Student-Teacher Relationship Scale.* Charlottesville, VA: University of Virginia.

QUALITY OF CARE OBSERVATIONS

Harms, T., Clifford, R. M. and Cryer, D. (1998). *Early Childhood Environment Rating Scale-Revised Edition (ECERS-R).* New York, NY: Teachers College Press.

Harms, T., and Clifford, R. M. (1989). *Family Day Care Rating Scale (FDCRS).* New York, NY: Teachers College Press.

Arnett, J. (1989). Caregivers in day-care centers: Does training matter? *Journal of Applied Developmental Psychology*, 10, 541-552.

PARENT INTERVIEW SCALES

Achenbach, T. (1996). *Child Behavior Checklist* (abbreviated version). University of Vermont: Center for Children, Youth, and Families.

Developing Skills Checklist—Home Inventory. (1990). Monterey, CA: CTB/McGraw-Hill.

Perlin, L.I., and Schooler, C. (1978). The structure of coping. *Journal of Health and Social Behavior*, 22, 337-356. (Pearlin Mastery Scale-Locus of Control).

Pianta, R.C. (1992). *Parent-Child Relationship Scale.* Charlottesville, VA: University of Virginia.

Radloff, L.S. (1977). The CES-D: A self-report depressions scale for research in the general population. *Applied Psychological Measurement*, 1, 385-401. (abbreviated version).

Appendix F: Language Decision Form

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