Child Care & Early Education
RESEARCH CONNECTIONS

ICPSR 03804

Early Head Start Research and Evaluation (EHSRE) Study, 1996-2010: [United States]

United States Department of Health and Human Services. Administration for Children and Families

Codebook Appendices, Volume 1

ICPSR | INTER-UNIVERSITY CONSORTIUM FOR POLITICAL AND SOCIAL RESEARCH
P.O. Box 1248
Ann Arbor, Michigan 48106
www.icpsr.umich.edu

About *Research Connections*

These data are made available by the Child Care and Early Education *Research Connections* project. *Research Connections* promotes high quality research in child care and early education and the use of that research in policymaking.

Research Connections is operated by the National Center for Children in Poverty at the Mailman School of Public Health, Columbia University and the Inter-university Consortium for Political and Social Research at the Institute for Social Research, University of Michigan, through a cooperative agreement with the Office of Child Care, Office of Family Assistance and the Office of Planning, Research, and Evaluation, Administration for Children and Families in the U.S. Department of Health and Human Services.

About the Data

These data are based on research conducted as part of the Early Head Start Research and Evaluation Project. Mathematica Policy Research, Inc. was responsible for the creation of the public use files.

Terms of Use

Information about Copyrighted Content

I. INTRODUCTION TO EHSRE USER DOCUMENTATION

The national Early Head Start Research and Evaluation (EHSRE) study began in 1995, at about the same time that the first Early Head Start programs were funded.[1] It was designed to provide a rigorous, comprehensive evaluation of the impacts of Early Head Start programs on low-income families and their young children through the children's third birthdays.

Funded by the Administration on Children, Youth and Families, Early Head Start is a comprehensive, two-generation program that focuses on enhancing children's development while strengthening families. Designed for low-income pregnant women and families with infants and toddlers up to age 3, the programs provide a wide range of services through multiple program options. Services include child development delivered in home visits, child care, case management, parenting education, health care and referrals, and family support. Early Head Start programs meet families' and communities' needs through one or more official program options: (1) home-based, (2) center-based, (3) combination (in which families receive both home visits and center experiences), and (4) locally designed. Because a program may offer multiple options, we characterized programs for research purposes according to the options they

[1] The data included in the data files documented here were collected for the national Early Head Start Research and Evaluation Project funded by the Administration on Children, Youth and Families (ACYF), U.S. Department of Health and Human Services under contract 105-95-1936 to Mathematica Policy Research, Princeton, NJ, and Columbia University's Center for Children and Families, Teachers College, in conjunction with the Early Head Start Research Consortium. The Consortium consists of representatives from 17 programs participating in the evaluation, 15 local research teams, the evaluation contractors, and ACYF. Research institutions in the Consortium (and principal researchers) include Administration for Children and Families (Rachel Chazan-Cohen, Judith Jerald, Esther Kresh, Helen Raikes, and Louisa Tarullo); Catholic University of America (Michaela Farber, Lynn Milgram Mayer, Harriet Liebow, Christine Sabatino, Nancy Taylor, Elizabeth Timberlake, and Shavaun Wall); Columbia University (Lisa Berlin, Christy Brady-Smith, Jeanne Brooks-Gunn, and Alison Sidle Fuligni); Harvard University (Catherine Ayoub, Barbara Alexander Pan, and Catherine Snow); Iowa State University (Dee Draper, Gayle Luze, Susan McBride, Carla Peterson); Mathematica Policy Research (Kimberly Boller, Jill Constantine, Ellen Eliason Kisker, John M. Love, Diane Paulsell, Christine Ross, Peter Schochet, Cheri Vogel, and Welmoet van Kammen); Medical University of South Carolina (Richard Faldowski, Gui-Young Hong, and Susan Pickrel); Michigan State University (Hiram Fitzgerald, Tom Reischl, and Rachel Schiffman); New York University (Mark Spellmann and Catherine Tamis-LeMonda); University of Arkansas (Robert Bradley, Mark Swanson, and Leanne Whiteside-Mansell); University of California, Los Angeles (Carollee Howes and Claire Hamilton); University of Colorado Health Sciences Center (Robert Emde, Jon Korfmacher, JoAnn Robinson, Paul Spicer, and Norman Watt); University of Kansas (Jane Atwater, Judith Carta; and Jean Ann Summers); University of Missouri-Columbia (Mark Fine, Jean Ispa, and Kathy Thornburg); University of Pittsburgh (Beth Green, Carol McAllister, and Robert McCall); University of Washington School of Education (Eduardo Armijo and Joseph Stowitschek); University of Washington School of Nursing (Kathryn Barnard and Susan Spieker), and Utah State University (Lisa Boyce and Lori Roggman).

offer families. Thus, programs were grouped according to three approaches: (1) home-based, (2) center-based, and (3) mixed-approach.

Detailed descriptions of the EHSRE study design, data collection, and evaluation results are available elsewhere (Administration on Children, Youth and Families 2001; and Administration for Children and Families 2002a).

Only a limited amount of the EHSRE data is included in this public use file release within the Child Care and Early Education Research Connections (CCEERC). The file only includes variables used in analyses for the Early Head Start final impact report (ACF 2002a), although a few items that would enable identification of families or sites have been excluded, and some items in this file may have only been used in exploratory analyses but not reported on. An expanded set of EHSRE data, including source data from direct interviews and assessments, additional constructed variables, and qualitative data, will be made available to qualified researchers as restricted use files through the Murray Research Center. The restricted use files release will include technical support enabling better access to the complex network of files included within the Early Head Start data set. Limited technical support will be available to accompany the public use file in the CCEERC.

The following chapters provide background information to further the user's level of understanding of the EHSRE study and its data files. Also available in Early Head Start Research and Evaluation (EHSRE) Study, 1996-2010: [United States]: Codebook Appendices, Volume II is (1) a copy of the baseline (Head Start Family Information System/HSFIS) forms, (2) copies of the parent interviews, with variable names indicated next to the questions in the HSFIS forms and parent interviews, (3) copies of the child assessments and videotape protocols, (4) a list of the variables available in the public use data file and (5) a table ("crosswalk") detailing the

correspondence among questions and variables across the parent interviews (conducted when children were about 14, 24, and 36 months old). Additional documentation on other data sources will accompany the release of the restricted use data files.

Detailed information on the study design and evaluation results when children were 3 years old can be found in Volumes I and II of the final evaluation report, *Making a Difference in the Lives of Infants and Toddlers and Their Families: The Impacts of Early Head Start* (2002a).

II. STUDY DESIGN

Seventeen Early Head Start programs participated in the national Early Head Start Research and Evaluation study. Once programs were selected to participate in the study, they began enrolling families and working with Mathematica Policy Research, Inc. (MPR) staff and their local research partners to carry out the requirements of random assignment. In the remainder of this chapter, we provide an overview of site selection, sample enrollment, and random assignment. The following chapters provide information about the data collected for the analysis.

A. SITE SELECTION

When the first 68 Early Head Start programs (Wave I programs) were funded in late 1995, they agreed, as a condition of funding, to participate in local and national research if selected to do so. In March 1996, 41 university research teams submitted proposals to the Head Start Bureau—in partnership with Wave I Early Head Start program grantees—to conduct local research and participate in the national evaluation. The Administration on Children, Youth, and Families (ACYF) purposely selected 15 research sites, using a number of criteria: (1) programs had to be able to recruit twice as many families as they could serve; (2) programs had to have a viable research partner; and (3) in aggregate, programs had to provide a national geographic distribution that represented the major programmatic approaches and settings and reflected diverse family characteristics thought to be typical of Early Head Start families nationally. Applying these criteria resulted in fewer center-based programs than desired, so in 1996 ACYF selected one additional center-based program from Wave I, and in late 1997 selected another center-based program (without a local research partner) from Wave II programs (75 of which were funded in mid-1996), resulting in the full sample of 17 programs.

Because the 17 research programs were not randomly selected, findings based on the data cannot be formally generalized to all Early Head Start programs funded during 1995 and 1996. However, as shown in the study's final impact report (ACF 2002a), the features of the 17 programs, as well as the characteristics of their enrolled families and children, are similar to those of all Early Head programs funded in 1995 and 1996. Thus, to the extent that the quality and quantity of services offered in the 17 programs are similar to those offered nationwide, findings from these data are likely to pertain to Early Head Start programs more broadly.

B. SAMPLE ENROLLMENT

Although Wave I grantees entered Early Head Start with varying degrees and types of experiences (see ACF 2002a,b), all had been asked not to enroll any families until it was decided whether they would be selected for the research sample. Because all programs had agreed, in submitting their original proposals, to participate in the random assignment process if they were selected for the research sample, it was not necessary to persuade any of the programs to cooperate. Thus, as soon as the programs were selected, beginning in spring 1996, MPR staff began working with their staffs to implement the random assignment process in conjunction with each program's regular enrollment procedures.

Although poverty-level families with children up to age 3 are eligible for Early Head Start, only low-income families with children up to 12 months old at the time of enrollment were eligible for the evaluation.[2] Programs participating in the research were asked to focus recruiting

[2] In a small number of cases it was discovered when dates of birth were verified after enrollment that the applicant child was up to 15 months old. These cases have been retained in the sample, although the variable indicating child age in months at enrollment has been top-coded for eight cases at 12 months. Actual (rounded) ages of all focus children are also provided for each data collection point.

and enrollment during the sample enrollment period on these families. Some sites also enrolled pregnant women. Overall, one-fourth of the families enrolled while pregnant with the focus child.

Except for recruiting about twice as many families as they could serve, programs were expected to recruit as they would in the absence of the research, with special instructions to be sure to include all the types of families that their program was designed to serve (including those whose babies had disabilities). MPR and ACYF created detailed procedures to guide the random assignment process (see ACF 2002a, Chapter II, for a detailed description).

C. RANDOM ASSIGNMENT

As soon as programs determined through their application process that families met the Early Head Start eligibility guidelines, they sent the names to MPR, and MPR staff entered the names and identifying information into a computer program that randomly assigned the families either to the program or to the control group (with equal probabilities). Program staff then contacted the program group families, while representatives of the local research partners notified the control group families of their status.

Control group families were not allowed to receive Early Head Start services until their applicant child reached the age of 3. They could, however, receive any other services available in their community.

Some program staff were concerned that random assignment might, by chance, result in denial of services to families with particularly high service needs. ACYF was very clear, however, that the study findings should pertain to all families and children that Early Head Start was designed to serve, including infants and toddlers with disabilities. To address program concerns, ACYF and MPR established a process by which programs could apply to have a

family declared exempt from participating in the research. No exemptions were requested, however.

Sample enrollment and random assignment began in July 1996 and was completed in September 1998. In most sites, sample intake occurred over a two-year period, although some took less time. The extended enrollment period was due in part to the extra work involved in recruiting twice as many families as programs were funded to serve, and in part to the process of new programs working out their recruitment procedures. Two programs completed sample enrollment in late 1997, and one (the 17th site) did not begin sample intake until fall 1997. Thus, the study population for the evaluation includes Early Head Start-eligible families who applied to the program between late 1996 and late 1998.

During the sample intake period, 3,001 families were randomly assigned to the program (1,513) and control (1,488) groups. The samples in most sites include between 150 and 200 families, divided fairly evenly between the two research groups.

Early Head Start staff implemented random assignment procedures well. We estimate that about 0.7 percent of all control group members received any Early Head Start services (that is, were "crossovers"), and most sites had no crossovers.[3] Furthermore, our discussions with site staff indicate that information on nearly all eligible families who applied to the program during the sample intake period was sent to MPR for random assignment (that is, site staff did not provide Early Head Start services to families who were not submitted for random assignment). Hence, we believe that the research sample is representative of the intended study population of

[3]Site staff reported that 10 control group families in 5 programs received Early Head Start services. One program had 4 crossovers, one program had 3 crossovers, and 3 programs had 1 crossover each.

eligible families, and that any bias in the impact estimates due to contamination of the control group is small.

Random assignment yielded equivalent groups: the average baseline characteristics of program and control group members are very similar (see Appendix D in ACF 2002a, Volume II). This is as expected, because MPR used computer-generated random numbers to assign families. Therefore, the only difference between the two research groups at random assignment was that the program group was offered Early Head Start services and the control group was not. Subsequent analyses showed that the baseline characteristics of program and control group members who completed subsequent interviews and assessments were comparable (ACYF 2001; ACF 2002a). Thus, differences in the subsequent outcomes of the two groups can be attributed to the offer of Early Head Start services with a known degree of statistical precision.

III. DATA COLLECTION

Comprehensive data from multiple sources were collected for the Early Head Start Research and Evaluation study. MPR prepared all the follow-up data collection instruments and trained all field staff. In all sites but one (where MPR collected the data), data collection field staff were hired by the local research teams, who were responsible, under subcontract to MPR, for collecting the data and monitoring data quality. Respondents were offered modest remuneration and a small gift to complete each set of interviews and assessments. Appendix B in ACF 2002a, Volume II, describes the data collection procedures in greater detail.

This section provides an overview of all data sources, and some information on the restricted use data files containing these data. While the public use file only contains constructed—not source—variables, data sources for the constructs that are included in the public use data file are noted. The response rates to the interviews and assessments, and the timing of interviews are also noted.

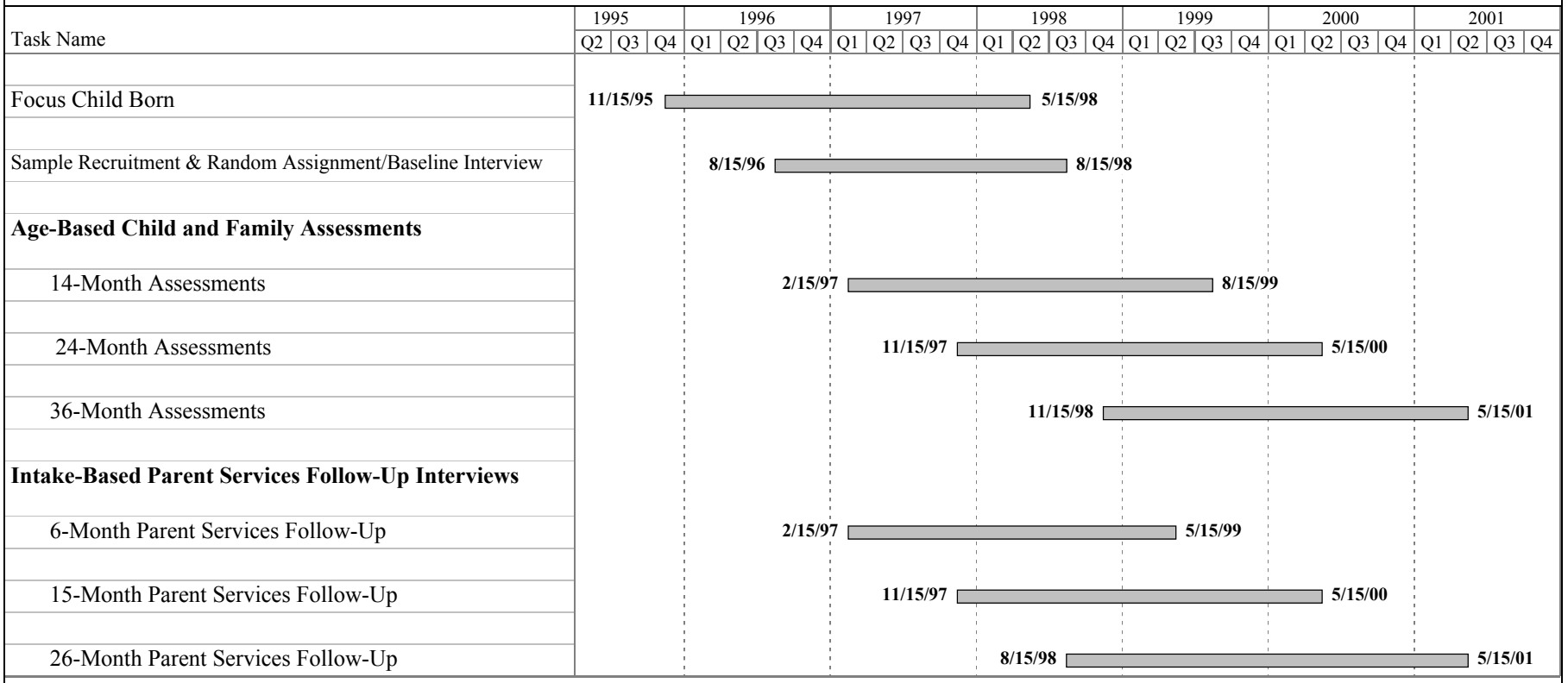
A. DATA SOURCES

Baseline data were collected when participants applied to participate in Early Head Start. Follow-up interviews and assessments were conducted based on two different collection schedules. The follow-up data were collected at time points based on (1) the number of months since random assignment, and (2) the age of the focus child (see Figure 1). Each family's use of services and progress toward self-sufficiency were seen as likely to be a function of the amount of time since the family applied for Early Head Start services. Therefore the interview schedule for these data was based on the number of months since random assignment. Other data—particularly those related to child and family development—were more likely to be a function of

FIGURE 1

SAMPLE RECRUITMENT AND DATA COLLECTION FOR THE EARLY HEAD START EVALUATION

Date Ranges Represented in the Public Use Files



Dates in public use files were rounded and some outliers were bottom- or top-coded.

the increasing age of the focus child over time. Thus the data collection schedule for these interviews and assessments was tied to children's birth dates.

The data sources include:

1. ***Baseline Data from the Head Start Family Information System (HSFIS) Program Application and Enrollment Forms and the MPR Tracking System.*** Most baseline data came from Head Start Family Information System (HSFIS) Program Application and Enrollment Forms that were completed by families at the time of program application. Some information on research status (program/control group assignment) and some characteristics of the applicant, mother, and focus child is from MPR's sample tracking system. The HSFIS data were entered by MPR staff into HSFIS software, which stored the data in DBase-format files. These were converted to SAS format for processing. The baseline data files include data for all families who were eligible when they applied to Early Head Start, whether or not they participated in the program or in follow-up data collections, except for a few cases excluded due to confidentiality concerns, as described below under data file preparation. Data from these forms are organized into three separate data files: the main baseline/HSFIS file, a file containing data for other eligible children in the household, and a file containing data for other family members. In addition to information from the HSFIS forms, the main baseline/HSFIS data file contains some information from the MPR tracking system and variables constructed from baseline information that were used to define subgroups for analyses. *Constructed variables from baseline/HSFIS data were used to define subgroups for the impact analyses and are included in the public use file.*

2. ***Parent Services Follow-Up Interviews (PSI) Targeted for 6, 15, and 26 Months After Random Assignment.*** These data contain information on (1) the use of services both in and out of Early Head Start (such as the receipt of home visits, and of services related to case management, parenting, health, employment, and child care); (2) progress toward economic self-sufficiency (such as employment, welfare receipt, and participation in education and training programs); (3) family health; and (4) children's health. PSIs were conducted using computer-assisted personal interviewing (CAPI) software, with responses entered directly into laptop computers by field staff. Most PSIs were conducted by telephone with the focus child's primary caregiver, although some interviews were conducted in person for those not reachable by phone. The PSI data are contained in three data files corresponding to the three targeted data collection points. *Constructed variables from the PSIs were analyzed for the impact report and are included in the public-use file.*

3. ***Parent Interviews (PI) Targeted for Completion When Children Were 14, 24, and 36 Months Old.*** These interviews obtained a large amount of information from the primary caregivers about their child's development and family functioning. These data usually were collected in person using hard-copy questionnaires, but some PIs or portions of them were conducted by telephone when necessary. Responses were data entered by MPR staff. Some portions of the PIs could be completed as self-administered questionnaires (SAQs) when appropriate; those data were merged with the interviewer-administered in the data files. The PI data are contained in three data

files corresponding to the three targeted data collection points. *Constructed variables from the Parent Interviews were analyzed for the impact report and are included in the public-use file.*

4. ***Child and Family Assessments Targeted for Administration When Children Were 14, 24, and 36 Months Old.*** Field interviewers conducted direct child assessments (including Bayley assessments, Peabody Picture Vocabulary Test-III (PPVT), and Test de Vocabulario en Imagenes Peabody (TVIP)), and recorded information from their observations of children's behavior and home environments for child and family assessments in Child Record Booklets (CRBs). They also videotaped semistructured parent-child interactions. Codes were developed from the videotapes and are included in the data files. Data from CRBs (including the PPVT and TVIP at 36-months), videotape codes, and Bayley assessments are contained in a single merged data file for of the three targeted data collection points that includes a record for each child who completed any of the instruments at that age. *Constructed variables from these sources were analyzed for the impact report and are included in the public-use file.*

5. ***Child Care Provider Interviews and Observations Targeted for Administration When Children Were 14, 24, and 36 Months Old.*** Interview and observation data were collected from child care providers for children who were in child care arrangements that met particular criteria when they were approximately 14, 24, and 36 months old. Different data collection instruments were used, and separate data files are provided, for children in child care centers and children cared for by family child care providers or relatives. However, matching variables were given the same names even when from different types of providers, so the data from both types of providers may be used together for some types of analyses. Also, Child-Caregiver Observation System (C-COS) data recorded on the Child-Focused Observation Form was collected for children in both types of care settings at 24 and 36 months. With two types of provider files at each of the three targeted data collection points plus two C-COS files there are a total of eight data files of child care provider source data.

6. ***Father Interviews Targeted for Collection When Children Were 24 and 36 Months Old.*** In addition to asking mothers about their child's father, biological fathers and father figures in 12 sites were interviewed directly about fathering issues at the time of the 24- and 36-month birthday-related interviews (but not when children were 14 months old).[4] The 36-month data collection also included a supplemental questionnaire to be administered only to men who had not completed the 24-month interview, to ask some background questions that were included in the 24-month but not the 36-month interview. These data are in three data files. [But if we want to also release the data for the handful of "second dads" that were interviewed, which are in

[4]The father study was supported with funding from the National Institute of Child Health and Human Development, the Ford Foundation, and the Office of the Assistant Secretary for Planning and Evaluation, as well as ACYF.

separate files, that would double.] *Constructed variables from the Father Interviews were analyzed for the impact report but are not included in the public-use file.*

7. ***Videotaped Father-Child Interactions When Children Were 24 and 36 Months Old.*** In seven of the sites where fathers and father figures were interviewed, interviewers also videotaped semistructured father-child interactions when children were 24 and 36 months old and recorded information from their observations of children and fathers in Father Child Record Booklets. Data from the Father Child Record Booklets and coded variables developed from the videotaped interactions are contained in a single merged data file for each of the two targeted data collection points. *Constructed variables from the Father videos were analyzed for the impact report but are not included in the public-use file.*

8. ***Exit Interviews When Children Reached 36 Months of Age.*** The exit interview obtained final updated summary information on the use of services and included questions for program group members about their experiences in Early Head Start. Whenever the 36-month PI was scheduled 30 days or more after the 26-month PSI, field staff administered the full exit interview to program families and all except the program experiences questions to control group families. If the 36-month PI was scheduled for less than one month after the 26-month PSI, and if the respondent was in the control group, no updated information on services was needed and the exit interview was not conducted. If the 36-month PI was scheduled for less than one month after the 26-month PSI, and if the respondent was in the program group, no updated information on services was needed and the interviewer skipped the service use questions and asked only the questions on program experiences. Responses were recorded on hard-copy questionnaires by field staff and were data entered by MPR. For file and variable naming purposes, the exit interview is considered the fourth in the series of PSIs.

A longitudinal study is underway that will follow and interview program and control group families through the time the focus children enter kindergarten to assess the longer-term outcomes experienced by study children and families. Information about public-use files from that phase of the study will be available at a later date.

B. RESPONSE RATES

Table III.1 displays overall response rates for key data sources by research status, as well as response rates for various combinations of data sources. Interview respondents are sample members who provided data that could be used to construct key outcome variables. Nonrespondents include those who could not be located, as well as those who could be located

but for whom complete or usable data were not obtained (see Appendix B in ACF 2002a, Volume II).

Response rates were higher for the PSIs and the PIs than for the Bayley and video assessments. Furthermore, as expected, response rates decreased somewhat over time. The rate was 82 percent to the 6-month PSI, 75 percent to the 15-month PSI, and 70 percent to the 26-month PSI. It was 78 percent to the 14-month PI, 72 percent to the 24-month PI, and 70 percent to the 36-month PI. At 14 months, it was 63 percent to the Bayley assessment and 66 percent to the video assessment, while at 36 months, it was 55 percent to each. 57 percent of sample members completed all three PIs, 39 percent completed all three video assessments, and 35 percent completed all three Bayley assessments. The percentage who completed both the 24- and 36-month interviews were 5 percentage points higher than those who completed all three interviews.

Importantly, response rates were similar for program and control group members for all data sources. However, they were consistently 2 to 6 percentage points higher for the program group.

In general, the same families responded to the different interviews. For example, among those who completed a 36-month PI, 87 completed a 24-month PI, and 81 percent completed both a 14- and 24-month PI. Similarly, among those who completed a 36-month video assessment, 99 percent also completed a 36-month PI, and 92 percent also completed a 36-month Bayley assessment.

Sensitivity analyses indicated that weighting for nonresponse would make little difference, so impact analyses, which controlled for differences between program and comparison groups, were not weighted for nonresponse. However, nonresponse weights for selected outcomes are available in the restricted use files.

TABLE III.1

RESPONSE RATES TO KEY DATA SOURCES
(Percentages)

Data Source	Program Group	Control Group	Combined Sample
Parent Service Interviews (PSIs)			
6-Month	83.9	79.3	81.6
15-Month	76.1	74.4	75.2
26-Month	71.1	67.9	69.5
15- and 26-Month	63.0	59.9	61.5
All three	58.6	54.4	56.5
Parent Interviews (PIs)			
14-Month	79.1	77.1	78.1
24-Month	73.9	70.4	72.2
36-Month	73.2	67.4	70.3
24- and 36-Month	64.4	58.2	61.4
All three	59.4	53.9	56.7
Bayley Assessments			
14-Month	64.2	61.2	62.7
24-Month	61.5	57.1	59.4
36-Month	58.1	52.4	55.3
24- and 36-Month	46.5	40.6	43.6
All three	37.0	32.6	34.8
Video Assessments			
14-Month	66.5	65.2	65.8
24-Month	62.2	57.5	59.9
36-Month	57.8	52.7	55.3
24- and 36-Month	48.1	42.7	45.4
All three	40.8	37.0	38.9
Combinations			
PSI 15 and PI 24	65.6	63.2	64.4
PSI 26 and PI 36	63.9	58.7	61.3
PI 24 and Bayley 24	60.5	56.5	58.6
PI 24 and Video 24	61.5	57.1	59.4
Bayley 24 and Video 24	55.9	51.9	53.9
PI 24, Bayley 24, and			

Data Source	Program Group	Control Group	Combined Sample
Video 24	55.4	51.5	53.5
PI 36 and Bayley 36	57.4	52.0	54.7
PI 36 and Video 36	57.4	52.4	54.9
Bayley 36 and Video 36	53.2	47.9	50.6
PI 36, Bayley 36, and Video 36	52.8	47.6	50.2
PI 24 and Bayley 36	52.2	46.0	49.2
PI 24 and Video 36	52.4	47.0	49.7
Video 24 and PI 36	55.8	48.8	52.3
Video 24 and Bayley 36	47.2	40.9	44.1
Sample Size	**1,513**	**1,488**	**3,001**

Source: Administration for Children and Families. *Making a Difference in the Lives of Infants and Toddlers and Their Families: The Impacts of Early Head Start. Volume I.* Washington, DC: U.S. Department of Health and Human Services, 2002, pp. 46-47.

Note: The response rates in this table pertain to the full sample of program and control families. For the public use files, 24 families, most of whom did not respond to most waves of data collection, were removed from the sample included in the data files to protect confidentiality.

C. TIMING OF INTERVIEWS

Most interviews were conducted near their target dates (see Appendix B in ACF 2002a,Volume II). For example, the average 15-month PSI was conducted 17 months after random assignment, and 80 percent were conducted between 12 and 18 months. Similarly, the average 26-month PSI was conducted 28 months after random assignment, and 76 percent were conducted within 30 months of random assignment. On average, the 6-, 15-, and 26-month PSI interviews were conducted 5 months before the 14-, 24-, and 36-month birthday-related instruments, respectively.

The average 24-month PI was conducted when the child was 25 months old, and 92 percent were conducted when the child was between 23 and 27 months old. The average 36-month PI was completed when the child was 37 months old and 91 percent were completed before the child was 40 months old. The corresponding figures for the Bayley and video assessments are similar to those of the PIs, however 96 percent of 36-month Bayleys and videos were completed before the child was 39 months old.

The distributions of interview and assessment completion ages were similar for program and control group families. Thus, it is not likely that impact estimates on outcomes (such as the child language measures) are affected by differences in the ages of program and control group children at the time the data were collected. As discussed in Appendix C in Volume II of ACF (2002a), a pertinent norming sample was not available to age-norm most measures.

IV. DATA FILE PREPARATION

The source data are generally stored in a separate data file for each type of interview or data source, including separate files for each of the 3 waves of most types of follow-up data. However, data from certain closely related sources have been put into merged files to reduce the total number of separate data files: at each of the 3 "birthday related" age points data from various "child and family assessments" (Bayley assessments, video interactions coding, PPVT/TVIP, and Child Record Booklet) are in one file and information from all of the various instruments used to collect information on childcare providers has been merged into a file for data from centers and one for data from family providers. Constructed variables are in separate data files, except for the baseline data file which contains both source and constructed variables.

In preparing each data file, records are included only for those families with data for that source (or from one of the several sources in the merged files), and thus the number of observations varies between files. The main baseline/HSFIS data file (and the public use file) contains one record for each family in the research sample, except for some cases that were omitted as described below due to confidentiality concerns. The same family ID number variable (IDnum) is included in all data files, and can be used to link all data for each case.

The following sections describe the preparation of the data files in more detail.

A. DATA CLEANING

The primary goal in checking and cleaning the data was to minimize data errors and create data files that are as accurate as reasonably possible. In general, we took a conservative approach to data cleaning. That is, data reflect the information recorded on the hard-copy instrument or in CAPI, except to correct any obvious recording errors (such as transposed digits, or incorrect dates that could be verified from other sources) and any errors that may have crept in

during the data entry process.[5] Hard-copy documents were double-checked as needed to resolve apparent discrepancies involving critical data items.

File preparation included filling in "logical skip" codes (.B in SAS, -2 in SPSS) for all items that were skipped based on instructions in the interview (according to responses to earlier questions or other information about the family).

In general, item responses were not deleted to make the skip logic consistent. For example, for some families, certain questions were completed even though they should have been skipped based on responses to earlier questions. The rationale for this approach is that if the respondent was able to give an answer to the later question, it could be that the error was actually in the response to the earlier question, and there was not adequate information for determining how to correct the error. However, apparent errors made by interviewers in "interviewer instruction" items were corrected when there was adequate basis (from earlier in the interview, or other sources) to identify and correct such errors.

In constructing "date" variables from month/day/year parts, missing data were imputed as follows: If the month was missing and the year was between 1960 and 1994 (before the beginning of the study), the month was imputed to be May. If the month was missing and the year was 1995 or later (during the study) the date variable was set to missing.

B. SCALE SCORES

In computing a score for a multi-item scale for inclusion in a data file, analysts consulted scale developers and experts regarding how to handle missing items for particular scales. In the absence of guidelines for a specific scale, however, analysts implemented the following

[5]Except for HSFIS, data that were entered from paper instruments were verified (double entered) to minimize data entry errors. The HSFIS software does not provide for double entry.

convention for handling missing items: If more than 25 percent of the items were missing for an individual case, or the scale was not based on a simple calculation, the scale was coded as missing. If 25 percent or fewer of the items were missing from a multi-item scale and the scale was based on a simple calculation such as a mean or sum, the score was calculated using imputed values for the missing items. For example, if the total score was the sum of the item scores or the average item score, we assigned the average item value across nonmissing items to the missing items. The actual item responses are included in the source data file and documentation unless copyright or agreements with test publishers prohibit it.

C. CONFIDENTIALITY ISSUES

To ensure respondent confidentiality and reduce the risk of individual disclosure, a few cases and some variables were omitted from the public and restricted use data files, and changes were made to other variables in these data files, as follows:

- All data were excluded for 24 families for whom MPR records indicate a miscarriage or death or adoption of the focus child within three years after random assignment. Depending on the date of death, few of these families were eligible for any data collection, and very few completed any interviews.
- Because sample sizes in individual study sites are relatively small, no information is provided on families' site location. The newly created case IDs (IDnum) are not assigned by site. The ID numbers of child care providers and data collectors, both of which could provide the means to link cases by site, are excluded from the files.
- Many date variables were originally entered as separate month, day, and year fields, then combined into date-type variables. Due to confidentiality concerns, the separate date parts have been omitted, and the dates, including the date of random assignment, dates of data collections, and the focus child's date of birth, have been rounded to the midpoint of the calendar quarter, and some have also been bottom and/or top-coded.[6] The files include the focus child's age in months at the time of the data collection.

[6] Variables with small frequencies in the lowest categories were bottom-coded by collapsing those categories into the highest of the low-incidence categories. Similarly, variables with small frequencies in the highest categories were top-coded by collapsing those categories into the lowest of the high-incidence categories. Variable labels and/or notations in the codebook note when this was done and how it was implemented for particular variables.

- Dates of birth for persons other than the focus child have been omitted from the files, and in their place, the person's age in years at the time of the data collection is provided. Some ages were also bottom- and/or top-coded; other household members were top-coded at age 66, and most other limits are specified in variable labels.

- Some information on household composition, including some relationship codes, have been regrouped or modified in other ways to protect confidentiality. Due to some of the types of recoding, some variables may not now "add up" properly as the un-modified ones would have. For example, a case may have had the number of "aunts and uncles" reduced to a top-coded value, but the total number in the household was within the typical range, so it was not adjusted, and as a result the total is no longer equal to the sum of the members of each type.

V. NAMING AND CODING CONVENTIONS

To facilitate the use of the data, variables and restricted use data files have been named according to common naming conventions. The names of both restricted use data files and variables within the data files include a document-specific prefix combined with further specifying information. The prefixes indicate the type of document and the "wave" from which the file or variable came, as described below. Most variable names for items taken from survey instruments and similar documents are based on the document prefix plus the item number in the hard-copy instrument. These naming conventions and standard codes used in the data files are described in more detail below.

A. DOCUMENT RELATED NAMING CONVENTIONS

1. Baseline/HSFIS Data

The main baseline/HSFIS data file is identified by the prefix "BL." Constructed baseline subgroup variables and items from MPR's tracking system have mnemonic names, while variables from the HSFIS include a prefix to indicate the forms from which they were taken:

- Prefix (first digit): H
- Form indicator (second digit): A=**A**pplication form, E=**E**nrollment form

2. Parent Services Interview and Exit Interview Data

The parent services interview and exit interview data are identified by 2- or 3-digit prefixes:

- Prefix (first digit): P
- Wave indicator (second digit): 0 = **0**6-month PSI, 1 = **1**5-month PSI, 2 = **2**6-month PSI, 3 = Exit Interview
- "Cumulative" indicator (third digit, constructs only): V (from eVer) indicates that the variable is cumulative up to the wave specified in the second digit.

- Examples: P0_HEAR is whether the focus child had received hearing testing, according to the 6-month PSI, P1V_HEAR is whether the focus child had ever received hearing testing by the 15-month PSI, using information from both the 6-month and 15-month PSIs.

3. Birthday-Related Data

The multiple elements of the birthday-related data collection are identified in file and variable names with prefixes consisting of the following digits:

- Prefix (first digit): B
- Wave indicator (second digit): 1 = **1**4-month data collection, 2 = **2**4-month data collection, 3 = **3**6-month data collection, V = "e**V**er" used for constructed variables combining information from all 3 waves
- Document indicator (third digit, and sometimes fourth digit):
  - P = **P**arent interview
  - R = child **R**ecord booklet
  - B = **B**ayley
  - V = **V**ideotaped interactions codes (fourth digit indicates type of task: 3=**3**-bag (the parent-child semistructured play task), H=**H**igh chair, T=**T**eaching, P=**P**uzzle)
  - D = father (**D**ad) interview
  - *F = **F**ather videotaped interactions codes (**third** digit indicates type of task: V=overall **V**ideo info, 3=**3**-bag, T=**T**eaching, P=**P**uzzle)
- Examples: B2V3PINT is a variable from the 24-month video 3-bag task coding, BVP_MAL1 is constructed using information from all 3 waves of parent interviews.

4. Constructed Variables

The majority of constructed variable names include a prefix indicating the data source in the first digits, as described above, then mnemonic characters for the remainder of the names. Some based on baseline/HSFIS and on PSI data have fully mnemonic names.

B. RECURRING VARIABLES

Certain key variables are repeated in each data file. These key variables include:

- **IDnum** uniquely identifies each case (family) and can be used for matching cases across data files.
- **Program** is a research status flag: 1=program group, 0=comparison group.
- **RandDt,** the random assignment date, is the date that the family entered the Early Head Start study and was randomly assigned to the program or comparison group (rounded to the midpoint of the calendar quarter and bottom- and top-coded).
- **CDOB** is the focus child's (FC) date of birth (rounded to the midpoint of the calendar quarter and bottom- and top-coded).
- **CSEX** is an alpha (character) variable denoting the focus child's gender: F=female, M=male, U=unknown.
- **[prefix]_data** is a 1/0 flag indicating that data are available for this document/source. (Occasionally a value of 0.5 signifies partial data.)
- **[prefix]_dat(e)** is the date of interview/data collection (rounded to the midpoint of the calendar quarter and bottom- and top-coded).
- **[prefix]_mth(s)** is the age in months of the focus child at the time of the interview/data collection.
- **[prefix]_vrsn** is the version number of the instrument. After data collection began, some documents were revised, including adding code categories, changes in which questions were included, and sometimes in "skip logic." In the 36-month Parent Interview, there were also some differences in items between English and Spanish versions of the questionnaires.

C. VARIABLE CODING CONVENTIONS

Conventions were followed for coding missing values, "other—specify" responses, and common response sequences, as described below:

1. Missing Values

The standard codes for missing numeric data[7] for all EHSRE data files, which differ depending on the file format, are shown in the chart below. In text and SPSS files various types

[7] The single exception in the public use file to the negative missing value codes is the variable CMTHS, which is the age in months of the Focus Child at the random assignment date, and includes valid negative numbers and NO missings.

of missing data are represented by negative number codes. In SAS files a corresponding set of SAS "special" missing value codes are used. The creation of an SPSS file from a text file should include SPSS syntax to assign the negative numbers as missing values for numeric variables. The creation of a SAS file from a text file should include SAS statements to convert the negative numbers to the special SAS missing codes as listed in the following chart. Note that the few missing values that occur in date variables in the EHSRE public use file simply contain the SPSS SYSMIS and SAS "." missing codes (since the behavior of the special missing codes is "peculiar" in SPSS).

Missing Category	SPSS Code Numeric	Date	SAS Code
Don't Know	-1	12/31/59[8]	.A
Logical Skip	-2	12/30/59	.B
Refused	-3	12/29/59	.C
Not Applicable	-4	12/28/59	.D
Item Missing	-5	12/27/59	.E
Section Missing (may be due to phone interview, child not present)	-6	12/26/59	.F
Special Missing - CAPI problem; child data unscorable / uncodable / procedural problem	-7	12/25/59	.G
Not in Version (was added after data collection began)	-8	12/24/59	.H

2. Coding "Other (Specify)" Responses

MPR used the following procedure when "Other (specify)" was selected as the response to a question:

[8] The 1959 dates for missing values in the SPSS files are related to the fact that in SAS (where the files were created), 1/1/60 is "day 0" and other dates are counted forward or back from that. So, for example, a "not applicable" -4 in a date variable was converted to 4 days before 1/1/60, i.e., 12/28/59.

- If the response fit into an existing category, the item was recoded to reflect that.
- If the response did not fit into an existing category, but there were at least 10 respondents who gave the same answer, a new category was created and the item was recoded. The new categories are listed in the documentation for those questions.
- If the response did not fit into an existing category and fewer than 10 respondents gave the same answer, the item code remained "other" (99).
- In public use files, some categories were collapsed for confidentiality reasons.

3. Other Coding Conventions

Other coding conventions include the following:

- Unless other codes are clearly printed in a document, all yes/no items are coded 0 = No, 1 = Yes.
- In most cases "Other" has been recoded to 99, especially in any documents showing the code Other = 0.
- In a few cases, as indicated in the documentation, multiple responses were given to questions requesting one response for which there is a single variable. If the multiple responses included a combination of "other" and a coded response, 100 was added to the coded response. Two separate coded responses were indicated in different digits of the response code.

VI. BASELINE/HSFIS DATA

This section contains detailed information on the baseline/HSFIS data. It is provided for researchers who want to fully understand the structure of these data for developing their own constructs. It may also be of use to researchers who wish to more fully understand the source data that went into the existing baseline constructed variables, but the constructs may also be used without study of this section of documentation.

A copy of the HSFIS forms is provided later in this appendix. The variable names have been handwritten onto the forms and all code values are indicated.

The main baseline/HSFIS data file contains the following sections of data (which are described in further detail below):

- Information on the case's program status and some characteristics of the applicant, mother, and focus child from the MPR tracking system, plus other constructed variables from baseline information that were used to define subgroups for analyses
- Information on the applicant (from HSFIS Application section 1, Applicant Demographics)
- Summary variables pertaining to all family members (from HSFIS Application sections 1, Applicant Demographics, 2, EHS Eligible Children, and 3, Other Family Members) (Detailed data on family members other than the applicant, father, and focus child, including eligible children and others, are in separate data files.)
- Information on the father (if present) (from HSFIS Application section 3, Other Family Members, and Enrollment Section 2, Other Family Members)
- Information on family circumstances (from HSFIS Application section 4, Family Composition and Resources, and Enrollment Section 3, Family Circumstances)
- Information on the mother's pregnancy (from HSFIS Enrollment Form Supplement A, Health and Pregnancy History), if the focus child was unborn at application
- Information on the focus child (from HSFIS Application section 2, EHS Eligible Children, and Enrollment Section 1, EHS Eligible Children), if the child was already born at application

Each variable has a name and a label. Tracking system variables have mnemonic names. The variable names for HSFIS items are based on the HSFIS form and item number. Variables that start with 'HA' refer to variables in the application form, and variables that start with 'HE' refer to variables in the enrollment form. For example, HA1_8 refers to item 1.8 on the application form, and HE3_3 refers to item 3.3 on the enrollment form. Variables that start with 'HE5' refer to enrollment information for pregnant women in Supplement A. (Some exceptions had to be made for items with many parts. For example, for Application form item 2.18, on types of risks, the "_18" that would normally be in columns 3 to 5 has been replaced by "ES," "BM," and "EN" for Established, Biological/Medical, and Environmental risks.) In addition to actual HSFIS items, variables that end with '_DATA' indicate whether or not data are available for a particular HSFIS section. The labels describe the content of the data items.

Most HSFIS items are ordered within each section on the data file as on the paper forms.

For some HSFIS items that are in the form of a list, we have added a variable (not on the HSFIS form) that is a count of the number of items in the list with positive responses. The names of such variables generally end with "_N."

The following sections describe the baseline variables and their location in the baseline data files in more detail.

A. SAMPLE TRACKING INFORMATION

The MPR tracking system variables describing the family's program status, random assignment date, and characteristics of the applicant,[9] mother, and focus child, have mnemonic

[9]In some cases, the HSFIS forms were completed with a father as the applicant, but the mother was present in the family and was the respondent for follow-up data collection. In such cases, the tracking and HSFIS data have been adjusted to contain information on the mother as the applicant.

variable names and are near the beginning of the main data file. The tracking variables about the applicant and focus child are preferred to the corresponding HSFIS variables because the tracking variables are more complete, and HSFIS has no information on focus children born after application/random assignment.

B. HSFIS APPLICANT INFORMATION

Information on the applicant comes from HSFIS Application section 1, Applicant Demographics. There were some data problems in the HSFIS computer system with the questions for teens (1.19 to 1.24), so these items have a higher than usual level of missing data.

C. HSFIS FAMILY MEMBERS INFORMATION

Summary variables pertaining to all family members (from HSFIS Application sections 1, Applicant Demographics, 2, EHS Eligible Children, and 3, Other Family Members) have been constructed to supplement the HSFIS household composition questions (HA4_1, etc., which contain some questionable data). These variables contain counts of family members listed in the HSFIS forms along several dimensions:

- According to the amount of time the person is reported to live with the applicant (all, some, or none of the time, or missing) (Variables: HA_liv1, HA_liv2, HA_liv3, HA_liv9)
- According to the relationship of the person to the applicant. For those living with the applicant (including some of the time or missing), categories are based on the relationship code of the person to the applicant; if this is missing, the applicant is categorized according to whether he is the biological father of the focus child, or whether "living with" is known or missing. Separate categories for those NOT living with the applicant indicate whether person is a "focal" adult (yes, no, missing). (Variables: HA_rel00 to HA_rel25, see labels for definitions.)
- According to the age of the person (excluding those who do ***not*** live with applicant). If age is missing, categories indicate whether the person is the adult applicant (focal adult: yes, no, missing). (Variables: HA_0_5, HA_6_17, HA_1829, HA_3049, HA_5065, HA_GE66, HA_focal, HA_NONfo, HA_DKfoc)

- To reduce the risk of individual disclosure, the total number of household members reported in HA_liv1 is top-coded at 8, the number of biological children of the applicant (HA_rel03) is top-coded at 5, and each of the other categorical variables is top-coded at 4 or lower. In addition, some of the original relationship categories have been grouped: 04 = "other custodial child" includes adopted, foster, grand-, and step-children; and 08 = "other parent" includes foster, adoptive, step, and god-parents and legal guardians.

Selected detail data on any other eligible children in the family (from HSFIS Application section 2, EHS Eligible Children) and on any other family members (from HSFIS Application section 3, Other Family Members, and Enrollment Section 2, Other Family Members) are available in separate data files.

D. HSFIS FATHER INFORMATION

Information on the "father" (if present) (from HSFIS Application section 3, Other Family Members, and Enrollment Section 2, Other Family Members) is included in the main HSFIS data files. The father variables contain information on either the father of the focus child or the spouse or significant other of the applicant.[10] The fathers' ages are bottom coded at 17 and top coded at 40.

E. HSFIS PREGNANCY AND FOCUS CHILD INFORMATION

The baseline data file contains information on either the mother's pregnancy, if the focus child was unborn at application (from HSFIS Enrollment Form Supplement A, Health and Pregnancy History), or on the focus child, if born (from HSFIS Application section 2, EHS Eligible Children, and Enrollment Section 1, EHS Eligible Children). Data items in the section that do not apply to a family are coded "not applicable" (.D in SAS, -4 in SPSS).

[10]In the only case where data were available for both, the information for the resident boyfriend was included for the "father" and the nonresident biological father data are in the restricted use "other family members" file.

In some instances, the focus child was born after program application but before random assignment. Child information exists in the application form for some of these children, but not for others. For families with child information recorded in the application, the data file contains this information but not any pregnancy information from Supplement A. However, for cases without child information, the data file contains only the pregnancy information.

VII. BIRTHDAY-RELATED PARENT INTERVIEWS

This section contains detailed information on the parent interviews (PIs). It is provided for researchers who want to fully understand the structure of these data for developing their own constructs. It may also be of use to researchers who wish to more fully understand the source data that went into the existing PI constructed variables, but the constructs may also be used without consulting this section of documentation.

A crosswalk was created to show the correspondence among common questions included in the PIs, which were conducted when children were 14, 24, and 36 months old. The crosswalk is provided later in this appendix. Copies of each of the PIs are available in Early Head Start Research and Evaluation (EHSRE) Study, 1996-2010: [United States]: Codebook Appendices, Volume II. The variable names have been handwritten onto the interviews and all code values are indicated. The following sections describe the information in the crosswalk and provide key information about the data files containing data from the birthday–related parent interviews.

A. CROSSWALK

On the parent interview crosswalk, the first column contains the question description. This is sometimes abbreviated, and the hard-copy instruments should be considered the more accurate reference. The second column contains the variable names. The third column indicates any scales that the variable is part of. Then there is a column for each of the instruments (14-month, 24-month, and 36-month) showing the question number for that question in that instrument. The question numbers in the 14-month questionnaire were the basis for the variable names. In the crosswalk, variable name prefixes are generally shown as “BnP,” where the “B” indicates birthday-related data, the “n” indicates that the second digit (wave indicator) may vary,

and "P" indicates data from a parent interview. The 14-month ("year one") variables have "1," the 24-month variables have "2," and the 36-month variables have "3" as the second digit.

B. PARENT INTERVIEW DATA FILES

Several aspects of the PI data files are notable. These include:

- Responses were generally recorded on hard-copy questionnaires by field staff and were data entered by MPR. Certain sections of the parent interviews could optionally be administered using self-administered questionnaires (SAQs) completed by the respondents themselves. The data files include variables indicating the actual source of such data for each respondent, but the responses from the SAQs are generally integrated into the data files with responses from the main questionnaires.

- Within each wave of parent interviews, after interviewing began, we revised the questionnaires to correct problems and reduce administration time. Thus, parent interviews were completed using multiple versions of the instrument that sometimes contained different code categories, different data items and/or different interviewer instructions (such as skip instructions). In addition, changes to MPR's data entry system during processing of the 14-month data caused some additional variations in the 14-month data files. A version number is indicated in the variable BnP_vrsn, which was determined based on patterns of "not in version" codes. The crosswalk and variable labels generally indicate data items missing from certain versions, and the missing items are coded -8(SPSS)/.H(SAS). In some cases (indicated in variable labels) the "new" version variables were reconstructed from corresponding data collected in a different fashion in the earlier versions of the instrument. In addition, specific version numbers have the following meanings:

  - In the 14-month parent interview files, the final version is designated version 4. Version 3 is also based on the final version of the interview, but an earlier version of our data entry program. Most sites have some observations from version 1 and five sites have a few observations from version 2 (the versions in use when interviewing began).
  - In the 36-month parent interview files, versions 1.5 and 2.5 indicate that Spanish versions of the questionnaire were used.

- In the 24-month and 36-month data files, "interview type" is indicated in the variable BnP_type, which was determined based on patterns of "not applicable" and "missing" codes. Telephone interviews, which skipped the 24-month Parent Woodcock-Johnson (Section 7B) and questions based on the interviewer's observations of the home, are coded BnP_type = 4. A few 24-month parent interviews that were completed by telephone, but included a Woodcock-Johnson or a SAQ that was completed at another time, are coded B2P_type = 3. Interviews completed when the child was not present, or when the Bayley and video were not completed are coded BnP_type = 2. Regular in-person interviews are coded BnP_type = 1. It must be

emphasized that *the type determination was based solely on the patterns of "not applicable" and "missing" codes*, since we have no other clear records of this. Items that were skipped for types 2, 3, and 4 are coded .F/-6.

- Despite "circle one" instructions, several 14-month parent interview questions had frequent instances of multiple responses. To accommodate this, some of the most common responses were separated into distinct variables. Each variable contains either the code value, if that response was given, or zero, if it was not given (or missing value codes, as appropriate). Questions for which additional response variables were added are: 8.5, 9.1, 9.2, and 9.3. The specific codes for which variables were added are labeled in the variable names and on the marked-up interview. (In the 24-month and 36-month parent interviews, the questions corresponding to 9.1, 9.2, and 9.3 were modified to allow for two responses each, and there was no question corresponding to 8.5.)

- In a few cases, multiple responses were given where there are not separate variables for the categories. If it was a combination of "other" and a coded response, 100 was added to the coded response (e.g., "6" plus "other" gives "106"). Two separate coded responses were put into different digits of the variable (e.g., "1" and "3" is coded "31").

- Other questions were already in multiple response format, with separate variables for different responses, but new codes were defined in later versions of the interview. Some of the new response codes were added to the file as distinct variables, generally after the variable for "other," while other codes are just reported within the "other" variable. Typically, each variable contains either the code value, if that response was given, or zero, if it was not given (or missing value codes, as appropriate), while the "other" variable may contain specific added codes, plus "99" for uncoded "other" responses, and "0" if a case gave no "other" response.

VIII. REFERENCES

Administration on Children, Youth and Families. *Building Their Futures: How Early Head Start Programs are Enhancing the Lives of Infants and Toddlers of Low-Income Families. Volumes I-II.* Washington, DC: U.S. Department of Health and Human Services, 2001. Available at http://www.acf.dhhs.gov/programs/core/ongoing_research/ehs/ehs_intro.html. and http://www/mathematica-mpr.com/3rdLevel/ehstoc.html.

Administration for Children and Families. *Making a Difference in the Lives of Infants and Toddlers and Their Families: The Impacts of Early Head Start. Volumes I-III.* Washington, DC: U.S. Department of Health and Human Services, 2002a. Available at http://www.acf.dhhs.gov/programs/core/ongoing_research/ehs/ehs_intro.html and http://www/mathematica-mpr.com/3rdLevel/ehstoc.html.

Administration for Children and Families. *Pathways to Quality and Full Implementation in Early Head Start Programs.* Washington, DC: U.S. Department of Health and Human Services, 2002b. Available at http://www.acf.dhhs.gov/programs/core/ongoing_research/ehs/ehs_intro.html and http://www/mathematica-mpr.com/3rdLevel/ehstoc.html.