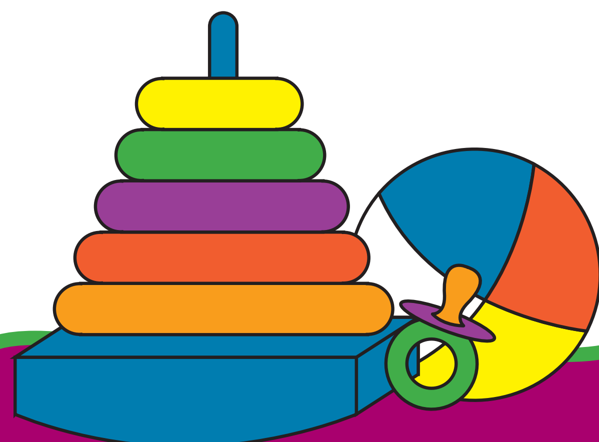
NCES
National Center for Education Statistics

Early Childhood Longitudinal Study, Birth Cohort (ECLS-B)

Early Childhood Longitudinal Study
Birth Cohort

User's Manual for the ECLS-B Nine-Month Public-Use Data File and Electronic Code Book

U.S. Department of Education
Institute of Education Sciences
NCES 2005-013



This page intentionally left blank.

NCES
National Center for Education Statistics

Early Childhood Longitudinal Study
Birth Cohort

U.S. Department of Education
Institute of Education Sciences
NCES 2005-013

Early Childhood Longitudinal Study, Birth Cohort (ECLS-B)

User's Manual for the ECLS-B Nine-Month Public-Use Data File and Electronic Code Book

January 2005

Christine Nord
Brad Edwards
Richard Hilpert
Laura Branden
Carol Andreassen
Anne Elmore
Desrene Sesay
Philip Fletcher
Jim Green
Ray Saunders
Rick Dulaney
Westat

Lizabeth Reaney
Kristin Denton Flanagan
Education Statistics Services Institute

Jerry West
National Center for Education Statistics

Notice of Copyrights and Permissions

Materials Copyright Notices:

SPONSORING AGENCIES

EARLY CHILDHOOD LONGITUDINAL STUDY, BIRTH COHORT (ECLS-B)

- National Center for Education Statistics, Institute of Education Sciences, U.S. Department of Education (ED)
- National Center for Health Statistics, U.S. Department of Health and Human Services (HHS)
- National Institutes of Health (NIH), U.S. Department of Health and Human Services

  National Institute of Child Health and Human Development
  National Institute on Aging
  National Institute of Mental Health
  National Institute of Nursing Research
  National Institute on Deafness and Other Communication Disorders
  Office of Behavioral and Social Sciences Research
  National Center on Minority Health and Health Disparities

- Economic Research Service, U.S. Department of Agriculture (USDA)
- Administration on Children, Youth and Families, HHS
- Maternal and Child Health Bureau, Health Resources and Services Administration, HHS
- Office of Special Education Programs, ED
- Office of the Assistant Secretary for Planning and Evaluation, HHS
- Office of Indian Education, ED
- Centers for Disease Control and Prevention, HHS
- Office of Minority Health, HHS

This page intentionally left blank.

ACKNOWLEDGMENTS

Over the past five years, many individuals and organizations have contributed to the design and conduct of the Early Childhood Longitudinal Study, Birth Cohort (ECLS-B). While it is not possible to name all the individuals who have made significant contributions to this study, we would like to recognize some of those who played a critical role during the development and implementation phases of the ECLS-B.

First, we would like to thank the 10,688 children and their parents who participated during the first wave of the study. The parents of these children invited us into their homes and allowed us to work with their children. We would also like to thank the more than 200 field staff and 16 field managers and supervisors who conducted the home visits both during the national data collection and as a part of several earlier field tests of the study design and instrumentation. And, we would like to express our appreciation to the State Vital Registration and Statistics Executives who provided the sample of 2001 birth certificates on which the study is based.

We gratefully acknowledge Marian MacDorman of the National Center for Health Statistics (NCHS) for all her efforts to make the ECLS-B a success. Her tireless work on the project's behalf to help gain the cooperation of the states and her review of the many sets of review documents submitted to the NCHS and State Institutional Review Boards were invaluable.

We wish to thank Natasha Cabrera of the University of Maryland at College Park (and formerly with the National Institute of Child Health and Human Development) for all her work in support of the health, special population, and father components of the study. We are especially grateful for her suggestions and encouragement in developing the father questionnaires of the ECLS-B, along with the assistance from the other members of the Developing a Daddy Survey (DADS) group.

Several others have consistently given their time in support of the study including: Vic Oliveira (Economic Research Service, U.S. Department of Agriculture); Michael Lopez, Louisa Tarullo, and Rachel Cohen (Administration on Children, Youth and Families, U.S. Department of Health and Human Services); Michael Kogan (Maternal and Child Health Bureau, U.S. Department of Health and Human Services); Linda Mellgren and Martha Moorehouse (Office of Assistant Secretary for Planning and Evaluation, U.S. Department of Health and Human Services); Karen Bourdon (National Institute of

Mental Health, National Institutes of Health); and Howard Hoffman (National Institute on Deafness and Other Communication Disorders, National Institutes of Health).

Westat, under the direction of the National Center for Education Statistics (NCES), designed and conducted the 9-month wave of the study. NCES received day-to-day support from the Education Statistics Services Institute (ESSI), American Institute for Research. NCES would like to recognize the contributions and efforts of several members of these two organizations' staffs. Carol Andreassen and Christine Nord of Westat led the instrument development activities. While the design of the ECLS-B child assessment battery and in particular the design of the Bayley Short Form–Research Edition were a team effort, no one invested more time, thought, and energy into these products than did Dr. Andreassen. Christine Nord contributed in a similar way to the parent interview and father questionnaires. NCES appreciates the work and dedication of each. They left a lasting imprint on the study that analysts will benefit from for many years to come. NCES would also like to thank Rick Dulaney and Ray Saunders of Westat for the leadership they provided to the programming, systems development, and data processing tasks associated with the study; Richard Hilpert of Westat for overseeing the field operations of the study; and Laura Branden for her tireless work with the NCHS and state Institutional Review Boards.

Kristin Denton Flanagan of ESSI and Liza Reaney (formerly of ESSI and now at Columbia University) supported NCES during the design and conduct of the ECLS-B and deserve special recognition. Each played an important role in the development of the ECLS-B and their efforts are reflected in the study instrumentation, data processing activities, and data products. NCES wishes to thank both of them for their years of service to the study.

A special thank you to Kendra Chandler Webb, age 9 (1994), for designing the ECLS logo.

TABLE OF CONTENTS

TABLE OF CONTENTS (CONTINUED)

TABLE OF CONTENTS (CONTINUED)

TABLE OF CONTENTS (CONTINUED)

TABLE OF CONTENTS (CONTINUED)

TABLE OF CONTENTS (CONTINUED)

TABLE OF CONTENTS (CONTINUED)

TABLE OF CONTENTS (CONTINUED)

TABLE OF CONTENTS (CONTINUED)

List of Tables

TABLE OF CONTENTS (CONTINUED)

List of Tables (continued)

TABLE OF CONTENTS (CONTINUED)

List of Tables (continued)

List of Figures

List of Exhibits

TABLE OF CONTENTS (CONTINUED)

List of Exhibits (continued)

TABLE OF CONTENTS (CONTINUED)

List of Exhibits (continued)

TABLE OF CONTENTS (CONTINUED)

List of Exhibits (continued)

TABLE OF CONTENTS (CONTINUED)

List of Exhibits (continued)

TABLE OF CONTENTS (CONTINUED)

List of Exhibits (continued)

DATA ANOMALIES AND ERRATA

In the course of preparing the data for release, some data anomalies and errata were identified. The anomalies are the result of allowable responses that seem odd when viewed in conjunction with other responses. The errata are due to errors in the programming of either the computer-assisted parent interview or the computer-assisted data entry (CADE) programs or in the design of the interview. These errors resulted in problems such as cases skipping questions that should have been asked or in not having an appropriate response option available. Both data anomalies and errata are described in this section.

Data Anomalies

- There are 4 cases where P1NUMCHO > P1NUMCH. P1NUMCH asks "Altogether, how many biological or natural children do you have?" P1NUMCHO asks, "How many of your biological or natural children live outside of your household?" The number of children reported in P1NUMCHO should be a subset of those reported in P1NUMCH. Either P1NUMCHO or P1NUMCH is in error in these cases.

- There are 31 cases where (P1NUMCH = P1NUMCHO) and X1RSPREL = 1 (respondent is the subject child's biological mother). This result indicates that all of the respondent's biological children are living outside of her household. This result is not possible because the subject child must be living with the respondent. Therefore either P1NUMCH or P1NUMCHO is wrong. Some respondents may have forgotten to include the subject child in the count of P1NUMCH. Other respondents probably misunderstood what was being asked in P1NUMCHO. Of the 31 cases, 10 were non-English-speaking households (X1LANGST = 1).

- There is a similar discrepancy with the parallel questions in the resident father questionnaire (F1NUMCH and F1NUMCHO). There are 5 cases where F1NUMCHO > F1NUMCH and the respondent to the resident father questionnaire is the biological father (F1RELCH = 1). Either F1NUMCHO or F1NUMCH is in error for these cases.

- There are 38 cases where (F1NUMCH = F1NUMCHO) and F1RELCH = 1 (biological father). This result indicates that all of the father's children live outside the household. Again, this result is not possible. Either F1NUMCH or F1NUMCHO is wrong.

- There are 11 cases where X1ASAGE (child's age at assessment) is younger than P1AGMILK (age child first began drinking milk as reported in the parent interview). Child's age at assessment is calculated using the child's birth date and the date of assessment entered into the Child Activity Booklet. The edit for P1AGMILK (age first gave child milk) checked the reported age against the child's current age (calculated using the child's birth date and the date of the parent interview). The 11 discrepancies occur because the assessment date differs from the parent interview date.

Although for the majority of cases the assessments were conducted on the same day as the parent interview, about 1 percent of the households requested that the interviewers return at a later time to conduct or complete the assessments. Usually the interviewers scheduled the second visit in the same week, but occasionally a longer time period elapsed.

- There are 2 cases that report an adult height of less than 4 feet (P1TALLUN, P1TALLFT, and P1TALLIN) and one case that reports a height of 7 feet, 11 inches. There is one case that reports an adult height of 2 feet in the father questionnaire (F1HGHFT).

- There are 48 cases that have a value of −1 (not applicable) for P1TALLIN, but that report valid heights at P1TALLFT. We suspect that these cases have reports such as "6 foot even" and that the interviewer entered "not applicable" at P1TALLIN.

- The responses to P1HHINCY (household income) and P1HMVALU (value of home) that are very low sometimes appear to be straight dollar amounts and sometimes appear to be dollar amounts reported to the nearest 1,000 or 10,000. Users who are particularly interested in the low income and home value ranges may want to review the data with other associated income (such as earnings) to edit these values. The data for P1HHINCY and P1HMVALU are as they were entered during the parent interview.

- There are 9 cases that report their pay before taxes (P1PAYBTX) as $0.1 or $0.01. Similarly, there are 10 cases that report the spouse's pay before taxes (P1SPINCM) as $0.01.

- The 3 cases that have an incomplete parent interview (X1STATP1=2) are twin cases where the interview was completed for the first twin, but was broken off before the interview was completed for the second twin.

- There are 10 cases where there is a biological father in the household who is identified as the spouse/partner of the biological mother respondent, but the response to F1RELCH ("What is your relationship to the child?") identifies someone other than the biological father as the respondent to the resident father questionnaire. Users will need to consider whether to include or exclude these 10 cases when analyzing the father data.

- There are 7 cases where there is a biological mother and another type of mother (e.g. adoptive mother) in the household. For the composite X1HPARNT (X1 CH Parents who reside in household), these cases take a value of 2 (Biological mother and other father). These cases are I_ID=015647, 017608, 130546, 141117, 847346, 397151, and 742941).

- There are 10 cases where the respondent was not the biological mother or the biological father, but the biological father lived in the child's household. For the composite on the status of the nonresident father questionnaire (X1STTNRQ), these cases take a value of 5 (no nonresident father). These cases could also have been assigned a value of 6 (not asked, respondent not biological mother) on X1STTNRQ. However, it was felt that category 5 (no nonresident father) provided users with more useful information.

- The ECLS-B protocol called for head circumference measurements to be collected for children who were very low birth weight. During data collection, the sampling birth weight measure (W1CBWT) was used to determine which children needed to have head circumference measurements taken. Prior to and during data collection, W1CBWT had missing data for some cases. For sampling and the collection of head circumference measures, these cases were assumed to be very low birth weight. In creating the composite for head circumference (X1CHCRFM), however, an updated birth weight measure from the National Center for Health Statistics (NCHS) natality file was used to determine whether measurements were needed (C1CHHCND). Cases that were determined not to be very low birth weight were given a value of −1 (not applicable) on X1CHCRFM even though measurements had been taken. The measurements collected remain on the file (C1CHHC1, C1CHHC2). A comparison of the birth certificate data from the natality file and W1CBWT yielded other discrepancies. The values for W1CBWT were not updated, however, because it is the variable that was used for sampling.

- There are 35 cases that have a household mother's weight recorded for the composite on mother's weight (X1HMGHT), but the first weight recorded during the interview (C1ADTWT1) was not the mother's weight. For 33 of these cases, the second weight recorded during the interview (C1ADTWT2) was the mother's and the composite takes the value of that second weight. The remaining two cases were foster mothers who were identified as such in an "other, specify" comment during the interview. It was decided to include them in the composite of mother's weight.

Errata

- The valid response ranges for P1NUMCH (number of biological children altogether) and P1NUMCHO (number of biological children living outside the household) are inconsistent. The valid response range for P1NUMCH is 1 to 12; while the valid range for P1NUMCHO is 1 to 17. Thus a response of 12 to P1NUMCH could mean 12 or more. Five (5) cases responded 12 to P1NUMCHO.

- P1RFEE, P1NFEE, and P1CFEE ask "Is there any charge or fee for the care {CHILD} receives from {relative/nonrelative/center} paid either by you or someone else?" The followup questions are "How much does your household pay to {relative/nonrelative/center} to care for {CHILD}?" The design does not allow "zero" as a correct response to the followup questions on how much the household pays for the child care. However, it became clear in reviewing the comments that some households reported that there was a charge or fee for the care, but that someone outside the household paid for that care. Thus “zero” should have been allowable. To overcome this error in the design, a new code 7 (HOUSEHOLD DOES NOT PAY – SOMEONE ELSE DOES) was added to P1RPDU1, P1NPDU1, and P1CPDU1. The cases that indicated that someone else paid had their values in P1RPDU1, P1NPDU1, or P1CPDU1 set to 7 and their values in P1RPAID, P1NPAID, P1CPAID set to 0.01 for the amount paid.

- A total of 1,188 respondents who answered “Yes” to P1LVBMAA (Did you live with your biological mother at any time until age 16?) and who should have been asked P1AGEBM (How old is your biological mother now?) were not asked P1AGEBM. Similarly, 2,246 cases who responded “Yes” to P1LVBFAA (Did you live with your biological father at any time until age 16?) were not asked P1AGEBF (How old is your biological father now?).

- Skip box SI151BX is incorrect. Three cases that responded “Don’t know” or “Refused” to P1SPWORK or P1SPVCTN should have been asked P1SPWKSN, but they were not. The parallel box (RI288BX) in the Respondent Information section is correct.

- The 66 cases who responded “Other, specify” to P1RNLNG (language used for parent interview) were accidentally skipped for the followup question RIPRNTRN (Was an interpreter used?). These cases have been set to −9 on R1PRNTRN and on X1INTRST (interpreter used status indicator).

- There is one case where a Child Activity Booklet was receipted that had valid data for the Bayley mental and motor scales, but because of a programming error, the data were not transferred from the CADE program to the data file. Because of this error, the case did not receive Bayley mental or motor scores. The case has a valid Parent-Child weight, but the Bayley scale scores are set to −9 (Not ascertained).

- There are two cases (I_ID=338198, and 843891) that should have been asked the questions about nonresident fathers. However, due to an error that occurred when the interviewer backed up to correct the identity of the household respondent, the computer-assisted parent interview program did not recognize that the father was not in the household. X1STTNRQ (status of nonresident father questionnaire) for these three cases has been set to “3” (No, nonresident father present but no mother permission).

- There is one case (I_ID=954750) that has a resident biological father in the household, but that was not identified as such. The problem occurred when the interviewer backed up to change the identity of the household respondent from biological father to biological mother. The computer-assisted parent interview program would not allow the interviewer to subsequently identify the father as the biological father. Thus she entered his relationship as “foster father” instead. This change, however, meant that the case did not follow the appropriate paths in the parent interview. Questions designed for resident biological fathers were skipped. The composite variables X1STTNRQ and X1FTHTYP have been corrected to indicate that there is no nonresident father and that the household father is the biological father. P1DAD_03 has also been corrected to indicate that the father is the biological father. The remaining data, however, are as recorded during the parent interview.

- There was an error in the CAPI program for variable R1PRNOT1. This variable indicates who was present during the home visit. It was intended to be a "code all that apply" item, but the program allowed the interviewer to enter only one response.

- Prior to finalizing the data, 53 cases were identified that consisted of relative/nonrelative female respondents to the parent CAPI interview. Originally, these cases had a value of X1MOMTYP = no resident mother. The composite X1MOMTYP was revised and a new category (Rel/nonrel resp–mother figure) was added to include these cases. Of these 53 cases, 24 were identified as having partners who completed the resident father questionnaires. Originally these cases had a value of X1FTHTYP = no resident father. The composite X1FTHTYP was revised and a new category (Rel/nonrel resps partner–father figure) was added to include these cases. In addition, 2 cases were identified that consisted of relative/nonrelative male respondents. Originally, these cases had a value of X1FTHTYP = no resident father. The composite X1FTHTYP was revised and a new category (Rel/nonrel resp–father figure) was added to include these cases. The parent demographic composite variables, however, were not revised for these cases. In addition, any imputations based on the results of X1FTYTYP, X1MOMTYP or the demographic composite variables were not revised to reflect the new values for X1FTHTYP and X1MOMTYP. To obtain the syntax necessary for updating the demographic composites to reflect the changes to X1FTHTYP and X1MOMTYP, go to http://nces.ed.gov/ecls.

- As described in chapter 2, interviewers were supposed to complete an interviewer remarks questionnaire (IRQ) as soon as possible after completing an interview and leaving the child's home. A substantial number of interviewers, however, did not complete this questionnaire (945 cases have missing IRQ data). Users should be aware that composites based on the IRQ data have missing data for these 945 cases. The following two composites are based on the IRQ: X1LNGPRT (language of parent interview) and X1INTRST (interpreter used status indicator).

This page intentionally left blank.

GETTING STARTED

This document highlights key information you will need to work with the Early Childhood Longitudinal Study, Birth Cohort (ECLS-B) data and directs you to the appropriate sections of the User's Manual so that you can get started quickly. To read more about any particular topic, go to the indicated section of the User's Manual. In this document, differences between the public-use and restricted-use files are highlighted, cautions and caveats about using the data are provided, and basic information about using the Electronic Code Book (ECB) is summarized.

Differences Between the 9-Month Public-Use and 9-Month Restricted-Use Data Files

The National Center for Education Statistics (NCES) typically releases both a restricted-use and a public-use data file for each data set it produces. Prior to being released as a public-use file, all data undergo a careful disclosure analysis to ensure the confidentiality of individual respondents. Variables that may pose a disclosure risk are modified to preserve confidentiality. Modifications include, but are not limited to, top or bottom coding outlier categories, collapsing categories for selected variables, and suppressing selected variables. See section 1.4.2 for a description of the modifications that are typically made to data in preparation for their release as a public-use file.

In order to preserve some information on topics for which detailed information was suppressed, 10 composites were added to the ECLS-B 9-month public-use file that are not in the restricted-use file. For example, one of the new composites, X1SYNDRM, is a summary indicator of whether a child has a congenital condition such as Down Syndrome, Turner's Syndrome, or spina bifida. The variables that indicate the specific conditions have been suppressed in the public-use file. See the end of table 7-6 for a description of the 10 composites that were added to the 9-month public-use file. See table 7-7 for a listing of the variables that have been modified or suppressed in the 9-month public-use file.

Cautions and Caveats

Other users of the ECLS-B data may have the same questions as you. NCES has developed a set of responses to users' most common questions. Please see the NCES web site for commonly asked questions and responses: http://nces.ed.gov/ecls.

In addition to the frequently asked questions and responses, there are other aspects of working with the data that are important to know, including the following:

- **Exclusions from the population of 2001 births.** Not all children born in 2001 were eligible to participate in the study. Children born to mothers under the age of 15 or those who were adopted at birth or shortly thereafter were excluded from the sample due to restrictions imposed by several state registrars and institutional review boards. However, such children are included in the target population, and sampling weights have been adjusted to account for the exclusion of these children from the sample since control totals for nonresponse adjustments will include all live births in 2001. On the other hand, infants who did not survive until 9 months will be considered as out of scope. See *Early Childhood Longitudinal Study, Birth Cohort (ECLS-B), Sampling Report for the Nine-Month Data Collection* (U.S. Department of Education, National Center for Education Statistics, forthcoming) for a discussion of the target population and sampling weights.

- **Not all sample children were 9 months of age at the time of the home visit.** The 9-month data file includes children who were born in 2001. Although every effort was made to assess the sample children and interview their families on or near the children's 9-month birthdays, some home visits occurred earlier and some later than the targeted age. Approximately 50 percent (50.5 percent) of the children were assessed within one month of their 9-month birthday (i.e., from 8.0 to 10.0 months). A small percentage of children (0.4 percent) were assessed before they were 8 months old. Due to practical operational limitations, a large percentage were assessed at 10 months or older (49 percent). Most of the children assessed at 10 months or older were assessed when they were 10 or 11 months of age (66 percent). The mean age at assessment was 9.7 months.

- **There are two different measures of the child's age in the file**. X1ASAGE is the child's age at the direct assessment. For most analyses, this is the age variable that you should use. In addition, there is X1BSIDAG, which is the child's age adjusted for prematurity. See section 7.5.1.1 for a discussion of how these two age variables were created.

- **Pay attention to missing data**. Users should always be sure to recode any missing data properly before conducting analyses. There are 5 different possible missing data codes in the file. See section 7.2 for a discussion of the different missing values codes and the circumstances in which they are used.

  Users may also want to omit some cases from analyses. For example, the Bayley Short Form–Research Edition (BSF-R) was designed to assess the development of children 8 to 11 months of age. Eighty-three percent of the children were assessed within this age range. The standard errors of the scale scores for children outside this age range may be high. Analysts may opt to exclude these cases from some analyses. See section 3.1.1 for a discussion of the BSF-R scale scores. See section 3.1.1.1 for further information on how and when to consider the impact of the child's age on assessment scores.

- **There may be no perfect weight**. The 9-month data file contains four sets of child-based weights. Because the ECLS-B is based on a sample of children, there are no weights available to produce household- or parent-level estimates. Although there are four sets of weights in the file, there are scenarios for which there may not be a perfect weight. For a discussion of the weights and guidance in selecting an appropriate one, refer to section 4.2.1.

- **Defining special populations.** The ECLS-B includes a number of analytic groups of interest that can be identified and studied separately. The ECLS-B was designed to support statistical analyses of selected race/ethnic groups, low birth weight babies, and twins. Specifically, the design allowed for statistical analyses of the following groups: American Indian, Chinese, Other Asian/Pacific Islander, Hispanic, Black (non-Hispanic) and White (non-Hispanic) births; very low (less than 1,500 grams), moderately low (1,500 to 2,500 grams), and normal (greater than 2,500 grams) birth weight births; and twin and non-twin (single birth and other multiple) births. Some of these groups were oversampled to achieve a sufficient number of cases (i.e., American Indian, Chinese, other Asian, low birth weight, and twin births). For other groups, a sufficient number were present in the sample so that oversampling was not required (i.e., White, Black, Hispanic, and normal birth weight births). See chapter 4, section 4.1.1 for information on how the groups were defined for the purposes of oversampling.

  In addition to the above groups, it is also possible to identify other populations for separate analyses, such as children living in households with incomes below the poverty threshold (X1POVRTY). Users who desire to study a specific subpopulation should search the ECB using the "NARROW" feature of the ECB to list variables that might help them identify their population of interest. See section 8.3.1 for a description of this feature.

- **There are several child's race/ethnicity variables contained in the file.** The file contains variables based on the child's race/ethnicity as reported by the parent respondent. It also contains race/ethnicity as reported on the birth certificate. Finally, there is race/ethnicity as it was defined for sampling purposes (see section 4.1.1 for a description of how race/ethnicity was defined for sampling purposes). The race/ethnicity from these three sources is not always the same. Race as defined for sampling purposes and on the birth certificate is based on the child's parents' race/ethnicity. The parent interview is the first time that race/ethnicity of the child is specifically identified. In addition, the parent interview uses the multi-race classification, which was not in operation for birth certificate data at the time of sampling. Thus, 2001 birth certificates did not allow for multi-race responses, and the sample was therefore selected based on a single race classification. In contrast, the ECLS-B classification employs the current NCES standards that allow for multi-race responses. Analysts will need to decide whether they wish to use race as defined for sampling purposes or race as reported for the child in the 9-month parent component in their analyses. For most analyses, users will probably want to use the race/ethnicity reported for the child in the 9-month parent component as it will be more consistent with how race/ethnicity is measured in other recent surveys.

Reference Guide

- **Electronic Code Book (ECB).** The ECB is designed to run under Windows 95®, Windows 98®, Windows 2000®, Windows XP®, or Windows NT® on a Pentium-class or higher personal computer (PC). The PC should have a minimum of 20 megabytes (MB) of available disk space. The ECB offers the most convenient way to access the data because it enables users to search the names and labels of variables, to examine question wording and response categories for individual items, and to generate SAS, SPPS for Windows, or Stata programs for extracting selected variables (see section 8.1.2 for a description of the several features of the ECB). Section 8.2 contains detailed instructions on how to install and open the ECB. The data user can create SAS, SPSS for Windows, and Stata programs that will generate an extract data file from the text (ASCII) data file on the CD-ROM. This text data file is referred to as the "child catalog" and is named Child9mP.dat in the CD-ROM root directory. The ECB CD-ROM also contains Portable Document Format (PDF) files of the associated questionnaires and of the User's Manual.

- **Data file.** The 9-month child catalog contains one record for each of 10,688 responding cases. This record contains data from all sources (e.g., parent CAPI instrument, parent self-administered questionnaire, father questionnaires (if applicable), and child assessments. It also contains selected birth certificate items. See chapter 8, section 8.7 for a description of the child catalog. See appendix B on the CD-ROM for the record layout for the child catalog. The file, named Child9mP.dat, is stored in the root directory of the CD-ROM as an ASCII file. It is strongly recommended, however, that users access the data using the ECB software available on the CD-ROM rather than access the ASCII file directly. When accessing the data with the ECB, you will need to have the CD-ROM in the drive.

- **Identification variables**. The 9-month data file contains a case identification variable (I_ID) that contains six digits and uniquely identifies each record. If a twin is sampled, a twin identification variable (I_TWINID) is used to identify the case identification number of the twin record. See section 7.1 for further information on these identification variables.

- **Availability of data on twins.** As mentioned earlier, the ECLS-B has an oversample of twins. In households with twins, the direct child assessments (i.e., BSF-R, Nursing Child Assessment Teaching Scale (NCATS) parent-child interaction videotape, and physical measurements) were administered separately for each child in a sampled twin pair. In the case of the parent CAPI instrument, certain sections applicable for both of the children were asked only once (e.g., family structure), while sections that could be different for each child were asked separately for each twin (e.g., child development, child health). When the data are presented in the file, each child has a separate ID and is presented as a separate case. The parent CAPI instrument data that were collected only once (e.g., family structure) are copied to both twin records, so that an analyst does not have to do any additional merging. As noted above, the variable I_TWINID in the file can be used to identify twin cases.

- **Instruments.** For the ECLS-B 9-month data collection, a child's primary caregiver (usually the child's mother) was interviewed using computer-assisted interviewing techniques. Self-administered questionnaires were used to collect sensitive information from the parent respondent. Hard-copy (paper) questionnaires were used to collect information from fathers, both those living with the child (resident) and those living elsewhere (nonresident). Chapter 2 provides an overview of the instruments. To help you decide what variables to use in your analyses, you should always review the actual instruments. Seeing the specific wording of the questions and the context in which they were asked is useful in understanding the results of your analyses and can help minimize errors. Appendix A on the ECLS-B ECB CD-ROM contains, with some exceptions, the 9-month instruments. The exceptions are measures that contain copyright-protected materials and instruments covered by agreements with the test publishers that restrict distribution. For this reason, the individual items from the BSF-R and from the NCATS are not included in appendix A.

- **Composite variables.** Numerous composites have been constructed for the ECLS-B data to make it easier for users to use the data set. Most composite variables were created using two or more variables that are in the data file or using information from other sources. Others are recodes of single variables. Composites based on the parent data include parent education (X1PARED), poverty status (X1POVRTY), and socioeconomic status (X1SESL). See section 7.5 and table 7-6 for details on all the composites contained in the 9-month public-use data file. It is strongly recommended that analysts consider using the composite variables because these variables represent the best data the study has, and some include sources not available in the data file.

- **Assessment scales.** A key feature of the ECLS-B data is the assessments conducted on each child. These assessments included direct and indirect measures of children's cognitive, socioemotional, physical, and psychomotor development. Chapter 2 provides a general description of the survey instruments, including the direct and indirect assessments. The direct child assessments included the BSF-R (an adaptation of the Bayley Scales of Infant Development, Second Edition [BSID-II]) and a teaching task from the NCATS. See section 3.1 for details on the direct assessments and the scores that are available for analysis. Section 3.1.1.1 discusses choosing the appropriate BSF-R score for analysis; section 3.1.2 describes the NCATS scales available for analysis.

- **Data on resident and nonresident fathers.** Two separate questionnaires were used to collect data from resident fathers and nonresident biological fathers. Most resident fathers were asked to complete the resident father questionnaire. The only exceptions were if the father was the respondent to the parent CAPI interview or if he was unavailable for an extended period (e.g., hospitalized or traveling). Only a subset of nonresident biological fathers were eligible for the nonresident father questionnaire. To be eligible, two criteria had to be met. First, the father had to have had some contact with either the child or the child's mother in the past 3 months and, second, the resident mother had to have granted permission for the nonresident father to be contacted. For information on the fathers from whom the ECLS-B collected data, see section 5.3.6. For information on which weights to use when analyzing father questionnaire data, see section 4.2.1.

- **Sample design and weights.** The ECLS-B employs a complex sample design. See chapter 4 for a description of the sample design. In order to obtain accurate estimates, you will need to select the appropriate weights. Section 4.2.1 describes the weights and provides advice for which weight to use for a given type of analysis. See exhibit 4-2 for a summary of the weights available for analysis.

1. INTRODUCTION

This manual provides guidance and documentation for users of the 9-month data of the Early Childhood Longitudinal Study, Birth Cohort (ECLS-B). This chapter presents an overview of the ECLS-B study. Subsequent chapters provide details on the instruments and measures used, the sample design, weighting procedures, response rates, data collection and processing procedures, and the structure and contents of the data file.

The ECLS-B is a multisource, multimethod study that focuses on the early home and educational experiences of children during their first 6 years. The study is sponsored by the U.S. Department of Education, National Center for Education Statistics (NCES) in the Institute of Education Sciences, in collaboration with several federal education and health policy agencies. The sponsoring agencies from the U.S. Department of Health and Human Services include the National Center for Health Statistics (NCHS), the National Institute of Child Health and Human Development (NICHD) and other components of the National Institutes of Health (NIH); the Administration on Children, Youth and Families (ACYF); the Office of the Assistant Secretary for Planning and Evaluation (ASPE); the Centers for Disease Control and Prevention (CDC); the Maternal and Child Health Bureau (MCHB); and the Office of Minority Health (OMH). Other sponsoring institutes from NIH are the National Institute of Mental Health (NIMH), the National Institute of Nursing Research (NINR), the National Institute on Aging (NIA), the National Institute on Deafness and Other Communication Disorders (NIDCD), the National Center on Minority Health and Health Disparities (NCMDH), and the Office of Behavioral and Social Sciences Research (OBSSR). The Economic Research Service (ERS) of the U.S. Department of Agriculture as well as the Office of Indian Education (OIE) and the Office of Special Education Programs (OSEP) of the U.S. Department of Education are also participating as sponsors for the study. Westat, a social science research company, is conducting the first two waves of the study for NCES.

The ECLS-B follows a nationally representative cohort of children born in 2001 from birth through first grade. It is part of the Early Childhood Longitudinal Study (ECLS), a longitudinal studies program comprising two cohorts—the birth cohort (ECLS-B) and a kindergarten cohort (ECLS-K). Together, these cohorts provide the breadth and depth of data required to more fully describe children's health, early learning, development, and education experiences. The birth cohort study focuses on those characteristics of children and their families, including children's early health care and in-home and out-of-home experiences, that influence children's development and first experiences with the demands of formal school (i.e., kindergarten and first grade). The kindergarten cohort follows a national sample of children who were in kindergarten during the 1998–99 school year from kindergarten through fifth grade. The ECLS-K measures key aspects of children's development and their environments (home and school) as they enter school and examines the influences of these on their academic achievement and experiences through elementary school. The two cohorts (ECLS-K and ECLS-B) are designed to overlap in time points (i.e., both have kindergarten and first grade data collections) and measures (e.g., similar direct child assessment and parent interview constructs). An overlapping cohort design aids in following the trajectory of children's development. For example, the ECLS-K has shown that differences between children's cognitive, social, and physical development exist at school entry (U.S. Department of Education, National Center for Education Statistics, 2000). The ECLS-B is designed to explain where, when, and how these differences emerge.

The parents of 10,688 children born in 2001 participated in the first wave of the ECLS-B study when the children were approximately 9 months old. Child assessments were conducted on a total of 10,224 of these children. Since the sampled children were born between January and December 2001, base year data were collected on a rolling basis between the fall of 2001 and the fall of 2002. The term "9 months" is used throughout this document to refer to the data collection that took place between the fall of 2001 and the fall of 2002, at which time most of the sampled children were 9 months of age. Because of the time that was required to locate families and to schedule and conduct the home visit, children ranged in age from approximately 6 months to 22 months at the time of the assessment (see chapter 3, section 3.1.1 for information about age and the use of the Bayley Short Form–Research Edition [BSF-R]). The first followup took place when the children were age 2 and additional waves of data collection are planned at 4 years of age and then in kindergarten and first grade.

Exhibit 1-1 presents the ECLS-B data collection schedule.

Exhibit 1-1. ECLS-B waves of data collection: 2001–07

Data collection	Date of collection
9-months	Fall 2001–fall 2002
2-year-old	2003
4-year-old	2005
Fall-kindergarten[1]	Fall 2006
Fall-first grade[2]	Fall 2007

[1]This collection is planned for the fall of the children's kindergarten year. Because the age requirements for school entry differ across the U.S., children in the sample will start kindergarten in different years. The majority of the children will enter kindergarten in fall 2006. All children will be included in the data collection, even if they are not in kindergarten in fall 2006.

[2]This collection is planned for the fall of the children's first grade year. The children will enter first grade in different years, with the majority entering in fall 2007. All children will be included in the 2007 data collection, regardless of their grade in school.

SOURCE: U.S. Department of Education, National Center for Education Statistics, Early Childhood Longitudinal Study, Birth Cohort, 9-month and 2-year-old data collections: 2001–03.

The central goal of the ECLS–B is to provide a comprehensive and reliable set of data that may be used to describe and to better understand children's early development; their health care, nutrition, and physical well-being; their preparation for school; key transitions during the early childhood years; their experiences in early care and education programs, kindergarten, and first grade; and how their early experiences relate to their later development, learning, and experiences in school.

There are several aspects of the ECLS-B design that help to achieve this goal as well as distinguish the ECLS-B from previous studies on early childhood development and growth. These include (1) oversampling of specific groups of children (e.g., American Indian and Asian and Pacific Islander infants, low birth weight infants, and twins), (2) collecting information directly from children's fathers as well as from their mothers, (3) videotaping a parent-child interaction, and (4) observing child care settings serving ECLS-B sampled children.[1] These features enrich the study design and provide in-depth descriptions of children's home and care experiences that can be linked to later development and experiences.

The ECLS-B may be used for both descriptive and analytic purposes. It will provide descriptive data on children's health status at birth and later; experiences in home, nonparental care, and

[1] Child care observations are conducted at waves 2 and 3 of data collection when children are 2 and 4 years of age, respectively. Information collected directly from the providers was not included in the 9-month data collection.

school environments; and children's development and growth through first grade. The ECLS-B will also provide a rich data set that will enable researchers to analyze how a wide range of family, nonparental care, school, community, and individual factors affect children's early experiences, development, and their success in school; to explore school readiness and the relationship between the early care and education experiences and later school performance; and to record children's cognitive, socioemotional, and physical growth from infancy through the first 2 years of formal schooling.

1.1 Background

Vital to any effort to improve the health, early care, and education for all of the nation's children is a research and data collection program that increases understanding of the dynamics that lead to differential school success. The ECLS-B helps to fill a gap in the existing data on early childhood development. It is the first study to follow a nationally representative sample of U.S. children from birth through the early formative years and on to school that collects data about the early experiences young children have in and out of their homes and the services they receive. The study measures children's early development prospectively, using a broad view of child development: the ECLS-B data contain repeated measures of children's cognitive development, physical health and growth, social development, and emotional well-being, along with information on family background and the educational quality of their home environments.

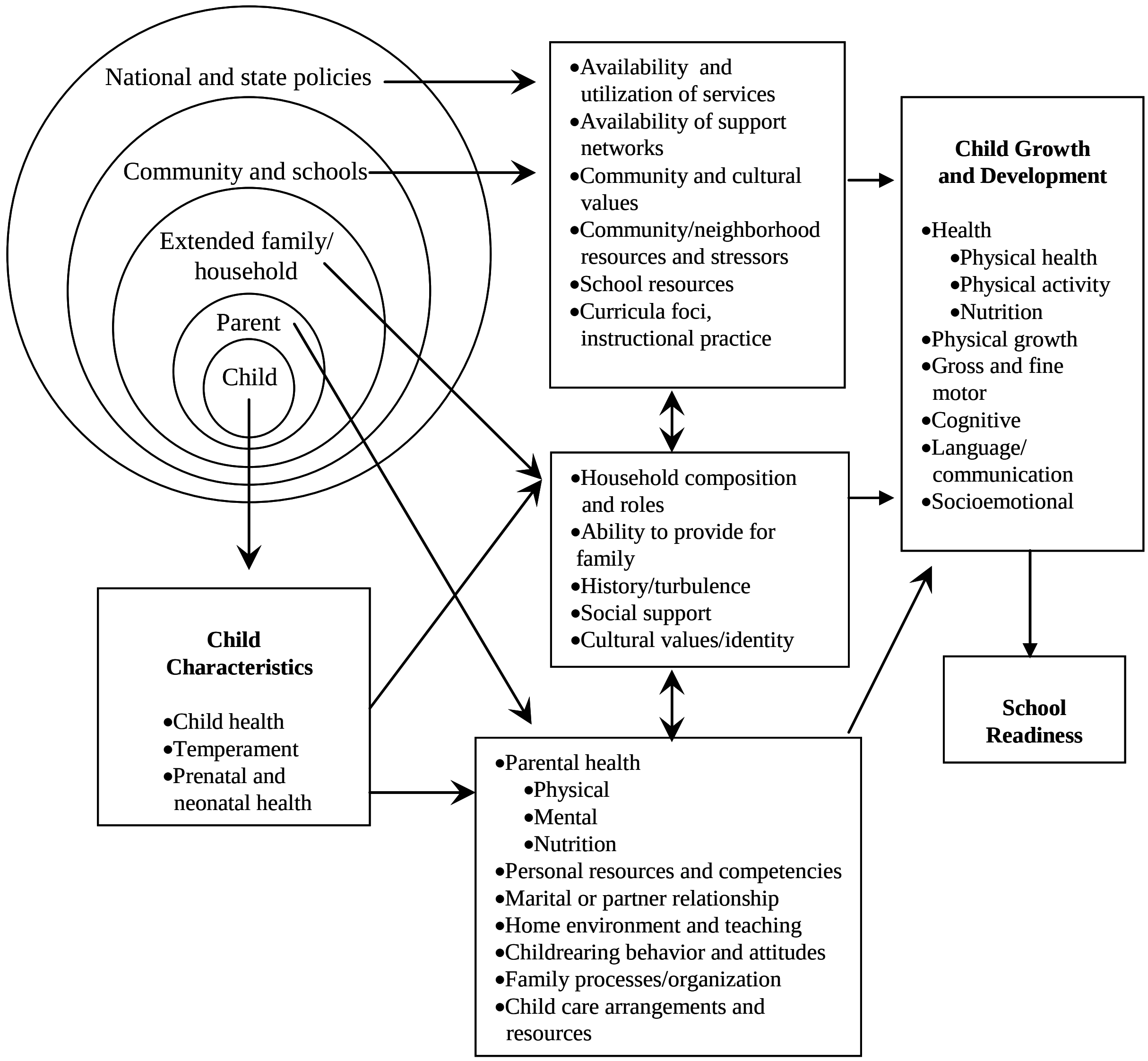
The ECLS-B allows researchers to examine how these developmental areas are affected by such factors as early home educational experiences, nutrition, chronic health conditions, provision of health care, father involvement, and participation in early care and education. Data on these experiences are also collected as they occur, except for retrospective information on prenatal care and newborn (neonatal) factors. This "in-time" collection produces a more accurate measurement of these antecedent factors and therefore a sound basis for drawing inferences about the relationship of early experiences to later growth and development in both cognitive and noncognitive domains. Thus, the ECLS-B can be used to effectively inform policies regarding children, their families, and their health, early care, and education.

Data collected during the first year of life (around 9 months) serve as a baseline for the study. A goal of the 9-month data collection has been to describe the diversity in the experiences, competencies, and backgrounds of U.S. infants. Because the ECLS-B sample includes substantial numbers of White, Black, Hispanic, Asian, and American Indian children, the ECLS-B data present many possibilities for studying cultural and ethnic differences in the educational preferences of families, the developmental patterns of children, and the educational resources and opportunities that different groups are afforded in the United States. The ECLS-B sample also includes substantial numbers of children with birth characteristics that may be indicative of later developmental difficulties (e.g., very low birth weight, moderately low birth weight, multiple birth). Thus, the ECLS-B data will enable researchers to analyze the differences in the growth and development of these children as well.

1.2 Conceptual Model

The design of the ECLS-B is based on the assumption that children’s preparation for school begins at (or before) birth and continues as they enter school. It is guided by a conceptual model of children’s development, care, and schooling that emphasizes the interaction between the child, family, health care access, early care and education programs, and community (figure 1-1). The ECLS-B recognizes the importance and inter-relatedness of factors that affect children’s health status, socioemotional development, and intellectual development and incorporates these with factors from the child’s family, community, early care and education, and (eventually) school-classroom environments. The study has paid particular attention to the role that parents and families play in supporting children’s growth, development, and learning through the first 6 years. As noted earlier, future waves of the ECLS-B will also gather information on the quality and experiences provided by early care and education arrangements.

Figure 1-1. ECLS-B conceptual model



SOURCE: U.S. Department of Education, National Center for Education Statistics, Early Childhood Longitudinal Study, Birth Cohort, 9-month data collection: 2001–02.

1.3 Study Components

The emphasis given to a broad view of child development and the different environments in which development and learning occur is critically important for the design of the ECLS-B. While parents are the primary reporters throughout the life of the study, at varying points data are also collected from birth certificates, care and early education providers, schools and teachers, and directly from the children themselves.

- Data from **parents/guardians** are collected through computer-assisted personal interviewing (CAPI) and through a self-administered questionnaire. The interview is conducted in the child's home at each data collection point. The parent/guardian interviewed is the individual who is the primary caregiver and the most knowledgeable about the care and education of the child. In nearly all of the cases (99.7 percent), this is the child's mother or female guardian. The parent/guardian interviews gather information about children's early health and development and about their experiences with family members and others. Parents/guardians also provide key information about themselves as caregivers, the home environment, and the neighborhood in which they live.

- **Children's** participation in the study begins at home at the 9-month data collection, in which children participate in activities designed to measure important developmental skills and attributes in the cognitive, socioemotional, and physical domains. The Bayley Short Form–Research Edition (BSF-R), which was designed specifically for the ECLS-B, is utilized in the 9-month data collection and consists of selected items from the Bayley Scales of Infant Development, Second Edition (BSID-II). It allows children's gross and fine motor skills as well as their receptive and expressive language skills and emotion regulation to be assessed. The Nursing Child Assessment Teaching Scale (NCATS) from the Nursing Child Assessment Satellite Training (NCAST) is used to assess the parent-child interaction for early precursors of cognitive and social skills. At the 9-month data collection, children's length, weight, and middle upper arm circumference (MUAC) are measured. Additionally, a measure of their head circumference is taken for children with very low birth weight. At the second wave of data collection (when the children turn 2 years old), the ECLS-B uses an updated BSF-R, a videotaped parent-child interaction (Two Bags Task), and an attachment rating. Children's height, weight, and MUAC are measured again, as is head circumference for very low birth weight children.

- **Fathers** complete self-administered questionnaires on the role they play in the development of their children. The father questionnaire captures information about activities fathers engage in with their children and their attitudes toward fatherhood. Fathers also provide key information about themselves as caregivers. This information is collected when the children are 9 months old and 2 years old, and will be collected again during future waves of the study. Both fathers who live with their children (resident fathers) and biological fathers who live elsewhere (nonresident fathers) complete self-administered questionnaires.

- **Birth certificates** contain a variety of useful data about the children, their mothers, and their families. They contain information on the date of birth, birth weight, gender, plurality (i.e., if multiple birth and birth order), and prematurity. Health characteristics of children, such as congenital anomalies, abnormal conditions of the baby and the baby's APGAR score, are included. Additionally information on parents' education, race, and ethnicity (including Hispanic origin) and mother's marital status is available. Birth certificates also include information on the mother's pregnancy history, prenatal care, medical and other risk factors during this pregnancy, and complications during labor and birth. Information on the obstetric procedures and delivery method is also available.

Beginning with wave 2 of the study, additional components that involve collecting information from children's child care providers, teachers, and schools will be fielded.

- **Care providers and preschool teachers**, like parents, represent a significant source of information on themselves (their backgrounds, teaching practices, and experience), the children in their care, and children's learning environments. Much of the data needed to describe the structure of children's care arrangements and education programs, to develop indicators of the quality of these arrangements and programs, and to profile the background and experience of the persons caring for these children can only be reported accurately by the care providers, teachers, and organizations themselves. Contacting children's care and education providers allows for the opportunity to collect information about children's development from sources other than their parents. Providers are interviewed when the children are 2 years old and again at 4 years of age. In addition to the interview, a subsample of care and education settings will be observed for key aspects in child care quality (e.g., learning materials, room arrangement, safety and health). These observations will also occur when the children are 2 and 4 years of age.

- Once the children enter formal schooling, **school administrators and teachers** become important sources of information on the physical and organizational characteristics of their schools, the schools' learning environments, educational philosophies, and programs.[2] As the ECLS-B cohort enters kindergarten and the first grade, their teachers become valuable sources of information on one of children's most immediate learning environments, the classroom. Teachers also represent important sources of information about children's development, both cognitive and social, and their academic achievement.

1.4 ECLS-B Data Files

The ECLS-B data files are released as part of a CD-ROM application. In addition to containing the actual data from the study, the CD-ROM also includes an Electronic Code Book (ECB)

[2] Prior to the kindergarten wave of the national data collection, NCES, its cosponsors, and contractors will examine the feasibility of collecting data from each child's school and teacher.

that allows easy access to information about the variables and their unweighted univariate frequencies. The CD-ROM also contains documentation on how to better understand the ECLS-B study, the data, the ECB software, and the instruments. A brief overview of these features is provided in section 1.4.1. The ECLS-B data are released in restricted-use and public-use versions. A brief overview of the differences between the restricted-use and public-use files is provided in section 1.4.2.

1.4.1 Availability of ECLS-B Nine-Month Data Files, Documentation, and Questionnaires

Several different materials are available on the ECLS-B 9-month CD-ROM:

- **ECLS-B 9-month child-level file.** The ECLS-B 9-month child-level file contains all the data collected from or about the children, including data from the child assessments, from children's mothers and fathers, and from their birth certificates. These data are all stored on a single data file (Child9mP.dat) on the CD-ROM. The Electronic Code Book (ECB), a software application on the CD-ROM, allows users to easily examine the variables in the data set by searching for variables and by providing their unweighted univariate frequencies. The data user can create syntax files suitable for use in SAS, SPSS for Windows, and Stata programs that will in turn generate an extract data file from Child9mP.dat containing only those variables an analyst wishes to use. See chapter 8 of this manual for specific directions on using the CD-ROM and the ECB. For users who cannot use SAS, SPSS, or Stata, a record layout for the data file (File Layout.pdf) is included in appendix B.

- **ECLS-B 9-month user's manual.** The ECLS-B 9-month CD-ROM also contains an electronic copy of the user's manual (i.e., this document) in Portable Document Format (PDF). To access this document, select the PDF file called *9-Month User's Manual Chapters 1-5.pdf* and *9-Month User's Manual Chapters 6-8.pdf* from the computer's drive (e.g., "D:") where the ECLS-B CD-ROM is located. To read these files, one must have Acrobat Reader® (version 4.05). Additional information on viewing, downloading, and printing PDF files can be found on the World Wide Web at http://nces.ed.gov/help/techissues.asp#pdf. The user's manual can also be accessed through the on-line help function of the ECB and is available on the World Wide Web at http://nces.ed.gov/ecls.

- **ECLS-B 9-month questionnaires.** The ECLS-B CD-ROM also contains copies of the ECLS-B 9-month questionnaires. These include the parent CAPI instrument, the parent self-administered questionnaire, the resident father self-administered questionnaire, and the nonresident father self-administered questionnaire. See chapter 2 for further details on the content of these instruments. These instruments are available in PDF. They are located in a folder (appendix A) on the CD-ROM. To access this document, select the file for the particular questionnaire of interest (e.g., *CAPI_parent.pdf* for the parent CAPI instrument) from the computer's drive (e.g., "D:") where the ECLS-B CD-ROM is located. To read this file, one must have

Acrobat Reader® (version 4.05). Additional information on viewing, downloading, and printing PDF files can be found on the World Wide Web at http://nces.ed.gov/help/techissues.asp#pdf. The instruments can also be downloaded from the ECLS website found on the World Wide Web at http://nces.ed.gov/ecls.

1.4.2 Differences Between ECLS-B Restricted-Use and Public-Use Files

Typically, NCES releases both a restricted-use and a public-use data file for each data set it produces. The restricted-use files generally contain a few more variables than the public-use files because they do not undergo the same level of disclosure risk analysis as the public-use files. However, as a safeguard to respondents, users wishing to obtain access to the restricted-use files must enter into a formal agreement with NCES that includes signing an affidavit stating that they will abide by NCES' standards of confidentiality or be subject to penalties. Because the public-use files are more broadly available, the public-use files undergo rigorous data disclosure analyses to protect the confidentiality of individual respondents. In preparing the public-use files, NCES takes steps to minimize the likelihood that an individual parent or child participating in the study can be identified. This is in compliance with the Privacy Act of 1974, as amended, the E-Government Act of 2002, and the Education Sciences Reform Act of 2002. These laws mandate the protection of confidentiality of respondents and limit the use of the data to research and statistical applications only, the one exception being when the information is acquired under court order by the U.S. Attorney General for investigation and prosecution of acts of terrorism (USA Patriot Act of 2001).

The process of preparing the data for release begins with a formal disclosure risk analysis. Variables identified as posing the greatest disclosure risk are altered, and in some instances, entirely suppressed, and in this way the public-use data files are created. Every effort is made to alter the files as little as possible, consistent with the requirement for confidentiality protection. After altering the variables, the disclosure risk analysis is repeated to verify that the disclosure risk has been reduced to acceptable levels.

The following data modifications account for the differences between the public-use and restricted-use data files for the ECLS-B 9-month data collection:

- Outlier values were top- or bottom-coded;

- Individual cases for which a particular variable posed an especially high risk of disclosure had the value of that variable altered (usually by no more than 5 to 10 percent) to reduce the risk;

- Some continuous variables were modified into categorical variables, and categories in certain categorical variables were collapsed; and

- Certain variables with too few cases and a sparse distribution were suppressed altogether, rather than modified; and

- Ten new composites were added to the 9-month public-use file. These composites provide summary information about topics for which more detailed information has been suppressed in response to the disclosure risk analysis. See the end of table 7-6 for a description of these new composites.

In addition to the above modifications applied to the public-use files only, both the public-use and the restricted-use files undergo some data swapping as another way to protect confidentiality. Swapping is a process by which values for potentially identifying variables are exchanged within pairs of records that have been matched according to key characteristics, such as household income, race/ethnicity of the child, child's sex, child’s health, and so forth. The purpose of swapping is to help guard against the risk that respondents may be recognized by users of the database who may be knowledgeable about one or more respondents. The modifications, including data swapping, that were implemented to avoid identification of parents and children do not affect the overall data quality.

Most researchers should be able to find all that they need in the public-use file. While very few of the variables were suppressed, there are a few users who might require the restricted file. Researchers interested in geographic data at a level more detailed than region will need the restricted-use file. There are also certain variables from the birth certificates that are only available on the restricted-use file. Those researchers examining certain rare subpopulations, such as children with specific non-English home languages, will find that the restricted-use file contains a few more variables. However, in many instances the sample sizes are too small to support these analyses. NCES recommends that researchers who are uncertain of which data release to use first examine the public-use file to ascertain whether their specific analytic objectives can be met using those data files.

1.5 Contents of Manual

This manual provides documentation for users of the 9-month public-use data of the ECLS-B. The manual contains information about the data collection instruments (chapter 2) and the psychometric properties of these instruments (chapter 3). For example, information about the design and content of the assessment measures (e.g., BSF-R) is found in chapter 2. Information about the reliability and validity of these measures is contained in chapter 3. The manual also describes the ECLS-B sample design and weighting procedures (chapter 4); data collection procedures and response rates (chapter 5); and data processing procedures (chapter 6). In addition, this manual shows how the 9-month data file is structured and provides definitions of composite or derived variables (chapter 7). Finally, this manual explains how to use the Electronic Code Book (chapter 8). Additional information about the ECLS program can be found on the World Wide Web at http://nces.ed.gov/ecls.

REFERENCE

U.S. Department of Education, National Center for Education Statistics. (2000). *America's Kindergartners: Findings from the Early Childhood Longitudinal Study, Kindergarten Class of 1998–99, Fall 1998* (NCES 2000–070), by Jerry West, Kristin Denton, and Elvira Germino Hausken. Washington, DC: U.S. Government Printing Office.

This page intentionally left blank.

2. DESCRIPTION OF DATA COLLECTION INSTRUMENTS

This chapter describes the survey instruments used during the 9-month data collection of the Early Childhood Longitudinal Study, Birth Cohort (ECLS-B).

The ECLS-B 9-month data collection took place from the fall of 2001 through the fall of 2002 (when children born in January through December 2001 turned 9 months of age). Data were collected by parent computer-assisted personal interview (CAPI) instruments and questionnaires, father questionnaires,[1] and direct child assessments during an in-person home visit. Data were also obtained from children's birth certificates and via field staff observation of the children's behavior and home setting during the home visit. Exhibit 2-1 lists all the sources of data in the 9-month data collection.

Exhibit 2-1. Sources of data and instruments in the ECLS-B 9-month data collection: 2001–02

Nine-month data sources and instruments
Direct child assessments
Parent CAPI instrument
Parent self-administered questionnaire
Resident father questionnaire
Nonresident father questionnaire
Child Observations and Interviewer Remarks
Birth certificate

SOURCE: U.S. Department of Education, National Center for Education Statistics, Early Childhood Longitudinal Study, Birth Cohort (ECLS-B) Nine-Month Data Collection, 2001–02.

Copies of the 9-month data collection instruments, except for the direct child assessments, are available on the ECLS-B CD-ROM, appendix A. The direct child assessment items are not available on the CD-ROM because they are largely copyright-protected materials and agreements with the test publishers restrict their distribution.

[1] During the home visit, self-administered questionnaires were left with the respondent for fathers to complete.

2.1 Direct Child Assessments

One-on-one direct child assessments were administered with ECLS-B sampled children during the 9-month data collection. Children participated in activities designed to measure important developmental skills in the cognitive, social, emotional, and physical domains. The assessments took about 52 minutes to administer. Most assessments were completed in a single home visit at the time of the parent interview. Only rarely (i.e., in less than 1 percent of cases) was it necessary to complete the child assessments over two home visits.

In general, the ECLS-B child assessments serve three purposes: (1) to describe children's developmental status at particular time points, (2) to examine the growth trajectory in children's development over time, and (3) to explore the relationship of early experiences to children's development (where assessment data are the outcomes). Additionally, the ECLS-B assessment enables researchers to accomplish several analytic goals. The data generated from the assessment can be used to explore the relationships between children's developmental outcomes and characteristics of their family, health care, child care, and community. One of the main purposes of assessment in the ECLS-B is to provide data that permit more direct examination of the variation in children's cognitive skills, socioemotional status, and physical well-being by characteristics of their backgrounds and in- and out-of-home experiences.

The ECLS-B uses two approaches to assessment—direct and indirect. Direct assessment includes those measures that involve the child's participation. The child's performance on these activities provides an indication of his or her current developmental level. In the ECLS-B, children's cognitive development and physical development are directly assessed. The indirect assessment consists of rating scales that evaluate the child's current behavior and skills and are completed by others (e.g., parents) who have knowledge of the child. Generally, indirect assessment is used for areas that are difficult to directly assess in young children.

At the 9-month data collection, information on children's development and experiences in the home, child care, and early education programs was gathered through interviews with parents. Data from the direct child assessments complemented these adult reports and provided important descriptive information such as children's achievement relative to key developmental milestones (e.g., sitting up, jabbering expressively) in the physical, cognitive and language, and social and emotional domains.

The ECLS-B 9-month direct child assessments consisted of three components: the Bayley Short Form–Research Edition (BSF-R), the Nursing Child Assessment Teaching Scale (NCATS), and physical measurements. Each of these components is discussed in sections to follow. Exhibit 2-2 displays the major domains measured during the direct child assessments by each component. Interviewers administered the components using a hard-copy booklet called the Child Activity Booklet. BSF-R item scores and physical measurements were recorded in the Child Activity Booklet. The instructions and activity list for the NCATS were also included in the Child Activity Booklet, which was available in both English and Spanish.

Exhibit 2-2. Components and domains covered in the ECLS-B 9-month direct child assessments: 2001–02

Direct child assessment component	Domain coverage
Bayley Short Form–Research Edition (BSF-R)	Cognitive (mental), physical (motor)
Nursing Child Assessment Teaching Scale (NCATS)	Social, emotional, cognitive
Physical measurements (height, weight, middle upper arm circumference, head circumference[1])	Physical

[1] Head circumference was measured only for ECLS-B sampled children who were very low birth weight.

SOURCE: U.S. Department of Education, National Center for Education Statistics, Early Childhood Longitudinal Study, Birth Cohort (ECLS-B) Nine-Month Data Collection, 2001–02.

2.1.1 Bayley Short Form–Research Edition (BSF-R)

In a national study of early childhood, a strong baseline measure of children’s development is key. The ECLS-B direct child assessments include a measure of children’s developmental accomplishments. At 9 months, the ECLS-B used the BSF-R. This instrument, designed specifically for the ECLS-B, uses a subset of items from the Bayley Scales of Infant Development, Second Edition (BSID-II; Bayley, 1993). The BSF-R assesses children’s cognitive and gross and fine motor skills as well as their receptive and expressive language skills. This section briefly describes this new assessment and the reasoning behind its creation. See section 3.1.1 for further information on the properties of and scores derived from this measure, chapter 5 for more detail on field staff training and the administration of the BSF-R, and *Early Childhood Longitudinal Study, Birth Cohort* (ECLS-B), *Methodology Report for the Nine-Month Data Collection, Volume 1: Psychometric Characteristics* (U.S. Department of Education, National Center for Education Statistics, forthcoming) for more extensive information on the development of the BSF-R.

The BSF-R is a reduced-item set of the original BSID-II. The BSID-II generally takes about 40 minutes to complete when administered in a clinical or laboratory setting by a trained professional. The BSF-R was developed for national use in a home setting with a traditional survey field staff (i.e., trained survey interviewers with basic child development knowledge). It was designed to retain the psychometric characteristics of the full BSID-II, a recognized measure in the field of infant/toddler assessment. The BSF-R (like the BSID-II) consisted of a set of items ranging in developmental difficulty. Items were selected based on both their operational ease and robust psychometric properties. While the full BSID-II features 178 mental items and 111 motor items (that can be used to assess children between the ages of one month and 42 months), the BSF-R for 9-month-olds consisted of 31 mental items and 35 motor items (that could be used to assess children between the ages of 8 and 11 months). Special efforts were made to identify and include items that could result in scoring information for multiple items from a single clustered administration. Therefore, in the BSF-R multiple items could be scored from one clustered administration, and, especially in the motor scale, several items could be scored from observation. For example, a cluster administration of the series of walking items (in which one holds the child's hands to assess standing with support and taking a step forward with support, and then lets go of the child's hand(s) to assess standing without support and taking steps without support) would result in several items being scored from one cluster administration. The number of item administrations was, therefore, fewer than the number of items scored. This also helped to reduce the amount of time to complete the BSF-R, which takes approximately 40 minutes to administer at this age by field staff with no training in child assessment.

The BSF-R for 9-month-olds consisted of three parts: a mental scale, a motor scale, and a set of spontaneous vocalization items. For most items, the children were presented with objects and verbal instructions, and their behavior was observed and recorded. Often the assessor also modeled the behavior required of the child. The BSF-R included stimuli such as blocks, cups, and a bell. The interviewer recorded children's responses in a hard-copy booklet as credit ("C") or no credit ("NC"). For example, one item required the child to put two round toys in a cup. If the child placed two or more toys in the cup he would receive credit ("C"); if he put in no toys or only one toy, he would receive no credit for the item ("NC"). More information on the scoring of items as credit/no credit can be found in chapter 5. The hard-copy assessment booklet (i.e., the Child Activity Booklet) included item-by-item administration steps and scoring instructions as well as score sheets. Assessments were shortened or discontinued if the interviewer perceived that the child was uncomfortable or distressed. Scores for each BSF-R scale (mental, motor) were computed only if at least two-thirds of the items were scored (i.e., had a “C” or “NC”).

Each BSF-R scale (i.e., mental and motor) featured a unique core set of items. The mental and motor core items were administered to all children in the ECLS-B. The BSF-R mental scale had 11 core items (plus 2 spontaneous vocalization items) and the motor scale had 14 core items. The 2 spontaneous vocalization items were scored solely by observation for all children and independently of the mental scale core items. Although the spontaneous vocalization items were not considered part of the mental scale core items in terms of administration, their scores were added into the mental scale scores when data were entered, for a total of 13 mental scale core item scores.

On the basis of their performance on the mental and motor core items, some children received a supplementary set of basal or ceiling items. Separate basal and ceiling sets were developed for each of the mental and motor scales. There were 9 items in the mental basal set and 9 items in the mental ceiling set. There were 11 items in the motor basal set and 10 items in the motor ceiling set. If the child only got a few mental core items correct (e.g., 3 or fewer out of 11 core items, excluding the spontaneous vocalization items), he or she was administered the set of 9 basal items. If the child received credit on almost all of the mental core items (e.g., 9 or more out of 11 items, excluding the spontaneous vocalization items), he or she completed the set of 9 ceiling items. Similarly, on the motor scale, if the child received credit for 3 or fewer core items, the set of 11 basal items were administered, and if the child received credit for 12 to 14 core items, the set of 10 ceiling items were administered.

The use of basal and ceiling item sets provided essential information on children’s developmental status. Without the basal items, if the child’s ability level was below the core item set difficulty level (i.e., getting very few answers correct), there would be no specific information on how low the child’s ability level was. Similarly, if a child got most of the core item set correct, then without the ceiling items it would be difficult to pinpoint the upper range of the child’s ability level from the core set alone. The basal and ceiling item sets described the infant’s developmental status more accurately and precisely. This is particularly important since the ECLS-B is a national sample of children. The population of infants demonstrates wide variation in early developmental abilities (from very advanced to delay in skills). A national sample represents diverse groups, and this variation in development may be even more prominent because the ECLS-B oversamples children with low birth weight, who may be at risk for developmental difficulties.

Mental Scale. The 9-month BSF-R mental scale included items designed to assess early cognitive and language ability as manifested in memory, habituation, preverbal communication, problem solving, and concept attainment. Children were presented with tasks such as putting blocks in a cup, ringing a bell, and responding to a parent's request (e.g., peek-a-boo). As mentioned previously, multiple items could sometimes be grouped together and scored from one administration. The mental scale contained 13 core item scores that were obtained from 9 administrations (with two administrations each scoring 2 items) and then two spontaneous vocalization items scored solely by observation. The basal item set for the mental scale consisted of 9 item scores from one item that was administered and from 8 items scored solely by observation. The mental scale ceiling item set contained 9 item scores of which 6 were administered and 3 were scored solely by observation. In the BSID-II, items are arranged into age sets by developmental difficulty. In the BSF-R mental scale, the core item set included items ranging in difficulty from 5 to 13 months of age;[2] some of the basal items were appropriate for children as young as 4 months; and some of the ceiling items were appropriate for children up to the age of 19 months. See chapter 3 for more details on the BSF-R mental scores.

Motor Scale. The 9-month BSF-R motor scale included items designed to assess gross and fine motor abilities (e.g., picking up objects, crawling, walking), perceptual-motor integration, and problem solving. Children participated in tasks such as picking up small objects, grasping a pencil, rolling over from being on their backs, sitting and standing unsupported, and walking with help. The BSF-R motor scale core set contained 14 item scores from 8 administered items. Similar to the mental scale, multiple items can be scored through one administration. The motor scale basal item set contained 11 item scores obtained through 4 administered items and 7 observed items. The motor scale ceiling item set had 10 item scores with 5 administered items and 5 observed items. In the BSID-II, items are arranged into age sets by developmental difficulty. In the BSF-R motor scale, the core item set included items ranging in difficulty from 5 to 13 months of age; some of the basal items were appropriate for children as young as 3 months; and some of the ceiling items were appropriate for children up to the age of 22 months. See chapter 3 for more details on the BSF-R motor scores.

---

[2] The age ranges of the mental and motor scales are based on the BSID-II age sets. An age range of 5 to 13 months indicates that across all of the BSF-R items some are included in the BSID-II 5-month age set and others are included in the BSID-II 13-month age set. It is not necessarily true that a particular item is found in all BSID-II age sets. For example, the mental scale may include two items: one with an age range of 5-7 months (meaning it is found in the BSID-II 5-, 6-, and 7-month age sets) and an item with an age range of 8-12 months (i.e., item included in the BSID-II 8-, 9-, 10-, 11-, and 12-month age sets). For details on the age coverage of individual items, see *Early Childhood Longitudinal Study, Birth Cohort (ECLS-B), Methodology Report for the Nine-Month Data Collection, Volume 1: Psychometric Characteristics* (U.S. Department of Education, National Center for Education Statistics, forthcoming).

Vocalization Items. While the interviewer administered items from the BSF-R mental and motor scales, she or he also scored several items related to children's spontaneous verbalizations and communication. Two of these contributed to the mental scale scores, but were not a part of the core mental scale administration (that is, they did not affect the decision for the interviewer to administer the basal or ceiling item sets). One spontaneous vocalization item was part of the basal item set and 2 were part of the ceiling item set. The vocalization items provided information on children's prelinguistic communication and early language development (e.g., production of vowel sounds, use of two distinct words).

2.1.2 Nursing Child Assessment Teaching Scale (NCATS)

The Nursing Child Assessment Teaching Scale (NCATS) is one part of the Nursing Child Assessment Satellite Training (NCAST), which is an entire package of instruments—including two scales (the teaching scale and a feeding scale) that assess parent-child interaction as well as an intervention component for case management (Sumner and Spietz, 1994). The NCATS assesses aspects of parent and child behavior in the context of a semi-standardized teaching interaction. The items assessing parent behaviors obtain information about such aspects as the parent's sensitivity to the child's cues, responsiveness to the child's distress, cognitive growth fostering, and socioemotional growth fostering. The items assessing child behaviors during the interaction obtain information about the clarity of the child's cues and the child's responsiveness to the parent. These aspects of parent-child interaction have been shown to be important to the development of social competence and security of attachment in children (Hartup, 1985; Karen, 1998), cognitive skills such as language acquisition (Snow, 1994), preliteracy (Bus and Van IJzendoorn, 1988; Whitehurst et al., 1994), and problem solving and exploration (Arend, Gove, and Sroufe, 1979; Matas, Arend, and Sroufe, 1978). The NCATS provides important information on the relationship between parent and child (the bi-directional "give and take" between them), as well as information on the child's home environment. At the 9-month data collection, the parent (the main respondent) selected an NCATS task that the ECLS-B sampled child did not yet know how to do and then taught the child that task. Several aspects of both the parent's and child's behavior were observed to assess the interaction. The NCATS gathered information on parental responsivity to cues (e.g., pausing as child initiates action), cognitive growth fostering (e.g., type of instructions and modeling), socioemotional growth fostering (e.g., cheerleading), and response to distress (e.g., rearranging toys). The NCATS also collected data on the child's clarity of cues (e.g., vocalizes) and responsiveness to the parent (e.g., smiles at parent, attempts eye contact).

Typically, a health or social service professional completes the NCATS, coding as the interaction occurs. While the interaction usually lasts less than 5 minutes, live coding in the ECLS-B would have required field staff, who were not all trained social service professionals, to observe and score 73 items of parent and child behavior. In addition, given the other tasks ECLS-B field staff had to complete, live coding would have limited the number of subscales that could have realistically been used, thereby reducing the amount of information gathered. Instead, the ECLS-B, with funding from its Administration on Children, Youth and Families cosponsor, videotaped the parent-child interaction at the 9-month data collection. Tapes were then coded along all subscales, providing information on early parent-child interactions. For more detail on the properties of the NCATS scales, see section 3.1.2. See chapter 5 for more information on the training and procedures for field staff to videotape and for coders to score these tapes.

2.1.3 Physical Measurements

To assess children's physical growth and development, the ECLS-B 9-month direct child assessment included several anthropometric measurements. These complemented the assessment of children's psychomotor skills in the motor scale of the BSF-R. Physical measurements were taken of children's length, weight, and middle upper arm circumference (MUAC) for all sampled children, and head circumference was obtained for very low birth weight (VLBW; 1,500 grams or less) children. These measurements took less than 5 minutes to obtain.

Children's length (or height) and weight are commonly used to chart children's growth, as indicators of general health and well-being, and, when they are older, to calculate a body mass index (BMI) that serves as an indicator of children's overall health and as an index to determine to what extent children are overweight. Children's middle upper arm circumference provides information on their nutritional state (e.g., malnourished). For very low birth weight children, an additional measurement was taken of their head circumference. Low birth weight children are at risk for small head size, which is linked to and predictive of problems with cognitive growth and development. Head circumference is often used in studies of young children as a measure of development and a reflection of brain size and growth. A measure mat (a pediatric length-measuring device) and a digital bathroom scale were used to obtain the length and weight[3] measurements, respectively. A specialized measurement tape was used to obtain the

[3] The child's weight was obtained by subtracting the weight of the respondent alone from the weight obtained by the respondent holding the child.

MUAC and head circumference measurements. All measurements were recorded on a recording form in the Child Activity Booklet. See chapter 5 for more information on the data collection procedures for measuring these attributes.

2.2 Parent Instruments

The 9-month parent data were collected using a parent CAPI instrument and a parent self-administered questionnaire (PSAQ). Parents or guardians were asked to provide important information about the sampled child, themselves, the home environment, parent attitudes, and family characteristics. Questions regarding family structure, child care use, household income, and community and social support were also included in the parent instruments. The interview was conducted as part of a home visit with the parent and child. See exhibit 2-3 for information on the topics and content of the 9-month parent instruments.

2.2.1 Parent CAPI Instrument

The first component of the ECLS-B parent interview was a 60-minute in-person interview in which the interviewer asked questions aloud and entered the responses in a computer (i.e., the parent CAPI instrument). The study design called for the child's mother to be the respondent for the parent instruments whenever possible; however, the respondent could be a father, step-parent, adoptive parent, foster parent, grandparent, another relative, or a nonrelative guardian. The respondent had to be knowledgeable about the child's care and education, 15 years of age or older at the time of the child's birth, and living in the household with the child. If the child's biological mother was under the age of 18 (and met the above conditions), she was still the primary respondent as she was considered an emancipated minor. Respondents for the parent interview were selected according to the following order of preference:

1. The child's biological mother;
2. The child's biological father;
3. Another parent or guardian; or
4. Another household member.

Exhibit 2-3. ECLS-B parent instruments by major content topics: 2001–02

Topic	Section	Content in parent interview
Household composition	FS	▪ Listing and updating household members and selected characteristics of members
Pregnancy, fertility, and neonatal hospitalization	PSAQ[1]	▪ Wantedness of pregnancy
	FH	▪ Receipt of prenatal care (e.g., place, payment, information conveyed during visits)
		▪ Pre-pregnancy weight and weight gained
		▪ Smoking, drinking, and supplement use pre-pregnancy and in last 3 months of pregnancy
		▪ If multiple birth, fertility help sought, and type of help received
		▪ If twins, whether identical/fraternal
	PSAQ[1]	▪ Hospitalization/intensive care as newborn
		▪ Age when first child born
Child nutrition/feeding	CD	▪ Breastfeeding
		▪ Early feeding (e.g., formula, cow's milk)
	HE	▪ Feeding practices (e.g., to sleep with bottle)
	CH	▪ Whether child able to feed self
Developmental milestones	CD	▪ Developmental milestones (e.g., crawling)
Child temperament	CD	▪ Distress to novel stimuli
		▪ Self-regulation
		▪ Level of difficulty to raise child
Socioemotional development	HE	▪ Frequency separated from child for a week or more
Home learning and language environment	HE	▪ Used any books/magazines on babies or parenting
		▪ Frequency respondent engages in selected activities with child (read books, tell stories, sing, errands)
		▪ Toys and other materials
		▪ Language(s) spoken in home
		▪ English literacy level
		▪ Frequency someone in household talks to child in language other than English
		▪ Frequency respondent engages in selected activities with child (e.g., go to a public place like zoo or museum with child)
Involvement of resident father[2]	HE	▪ Frequency engages in selected activities with child
		▪ Frequency takes care of child
Family routines and practices	HE	▪ Sleeping routines
Parental attitudes and childrearing knowledge	PA	▪ Parent authoritarianism
		▪ General knowledge of child development
Child care arrangements	CC	▪ Currently or ever in relative, nonrelative, or center care
		▪ Age first received care, by care type (e.g., relative, center)
		▪ Number of care arrangements
		▪ Characteristics of the care where spends most hours (for each care type)
		▪ Amount of time spent in care per week
		▪ Charges for the care and whether anyone helps financially
		▪ What parent looks for in selecting care
Child health	CH	▪ Rating of child's overall health status
		▪ Usual place of medical care

See notes at the end of this exhibit.

Exhibit 2-3. ECLS-B parent instruments by major content topics: 2001–02—Continued

Topic	Section	Content in parent interview
Child health (continued)	CH	▪ Number well-baby check-ups ▪ Selected conditions (asthma, respiratory illness, ear infection, gastrointestinal illness), treatment of most recent episode/incident, and emergency room or hospitalization ▪ Injuries (fall, choking, auto, burn, bite, poison) ▪ Other medical conditions, impairments, disabilities ▪ Use of ventilator, apnea monitor ▪ Medical specialist referral ▪ Services received, early intervention ▪ Health insurance
Family health	FH	▪ Any household members have a special need or disability ▪ Respondent's overall health ▪ Usual place of medical care ▪ Respondent's drinking behavior ▪ Respondent's current smoking behavior ▪ Anyone in household smokes ▪ Depression; mental health needs
	PSAQ[1]	▪ Stressful life events
Marital history	MH	▪ Marital status ▪ Marital history ▪ Ever married/lived with biological father and dates of first cohabitation and marriage ▪ Child's biological father living/dead
	PSAQ[1]	▪ Quality of current marriage/relationship ▪ Frequency and type of arguments
Social support	SS	▪ Emotional support ▪ Financial support ▪ Support in case of emergency ▪ Advice for parenting
	RI	▪ Respondent's parents living/dead ▪ Respondent's closeness to parents
Community support	CS	▪ Support from community agencies ▪ Frequency of social gatherings with friends ▪ Attendance at religious service in last year ▪ Participation in community service activities
Respondent's information	IN	▪ Month and year of birth
	RI	▪ Family composition growing up ▪ Educational attainment of respondent's parents ▪ Educational attainment of respondent ▪ School experiences
	HE	▪ Primary language
	RI	▪ Current participation in education or training ▪ Employment history 12 months before child's birth ▪ Current employment/unemployment, characteristics of job (e.g., duties, job title) ▪ Benefits through current job ▪ Maternity leave

See notes at the end of this exhibit.

Exhibit 2-3. ECLS-B parent instruments by major content topics: 2001–02—Continued

Topic	Section	Content in parent interview
Spouse information	SI	▪ How long spouse has lived with child ▪ Date of birth ▪ Educational attainment ▪ Current participation in education or training ▪ Current employment/unemployment, characteristics of job (e.g., duties, job title) ▪ Benefits through current job ▪ Paternity leave
Biological (nonresident) father information[3]	BF	▪ Whether father lived in household and length of time ▪ Date of birth ▪ Educational attainment ▪ Distance of residence from child and time between visits ▪ Frequency look after child ▪ Child support (any change and its effect) ▪ Other forms of support (clothes, medical, child care expenses) ▪ Legal paternity established ▪ Contact with respondent
Welfare and other public assistance	WP	▪ Food Stamps ▪ TANF/welfare ▪ Medicaid ▪ WIC
Household income and assets	HI	▪ Number of adults contributing to household income ▪ Household income in last year ▪ Type of housing (house, apt., rent, own, subsidy) ▪ Value of home ▪ Own automobile ▪ Own stocks, savings
Household food sufficiency	HF	▪ USDA questions on food sufficiency/shortages/availability

[1] Parent self-administered questionnaire.

[2] Because of widespread interest in fathers' involvement in their children's lives, information collected specifically from the father in a self-administered questionnaire is presented separately. The person being referred to in this section is the spouse/partner of the parent respondent. In most cases this person is the child's father who resides in the home, but could in some cases be a stepmother, grandfather, etc., if the parent respondent is not the child's mother. Conceptually, fathers' involvement is part of the home learning environment. These items ask the parent respondent about his/her spouse/partner's involvement.

[3] Information on the nonresident biological father is collected only when the parent respondent is the biological mother and the biological father does not reside with the child.

SOURCE: U.S. Department of Education, National Center for Education Statistics, Early Childhood Longitudinal Study, Birth Cohort (ECLS-B) Nine-Month Data Collection, 2001–02.

The parent interviews were conducted primarily in English, but provisions were made to interview parents who spoke other languages. Bilingual interviewers were trained to conduct the parent interview in either English or Spanish. A Spanish CAPI instrument was used when needed as the parent CAPI instrument was programmed in both English and Spanish. An interpreter, either a community or household member, was used for families who spoke languages other than English or Spanish. Fewer than 0.1 percent of the cases were not completed due to language difficulties (see section 5.4.1.4).

Topics addressed in the 9-month parent CAPI instrument included current household members, family sociodemographic characteristics, languages spoken in the child's home, child care arrangements, home activities, parent education and employment, receipt of public assistance, and household income and assets. The parent CAPI instrument also collected information on the parent's childrearing beliefs and knowledge of child development, the parent's psychological well-being and health, the community resources and social support available to the family, the household's food situation, and, if applicable, information on the child's nonresident biological father. In addition to information on these children's homes, families, and communities, the 9-month parent CAPI instrument also collected data on the children's development. At this young age, certain essential aspects of development are difficult to assess directly (e.g., temperament, past achievement of key milestones); thus, the 9-month parent CAPI instrument included questions on children's socioemotional, cognitive, and physical development (e.g., self-regulation, verbalizing, sitting unsupported, feeding self). This indirect information complements the data from the direct child assessment.

2.2.2 Parent Self-Administered Questionnaire

The ECLS-B 9-month parent interview included a short (about 5 to 10 minutes) self-administered questionnaire (PSAQ) with items that could be considered sensitive. For example, some items involved reporting incidents and opinions that a respondent may not have felt comfortable answering in front of her spouse, partner, or other children (e.g., marital happiness). In the interest of obtaining complete answers, maintaining confidentiality, and not creating any problems in the home, the questions were administered in a hard-copy self-administered format. The parent was given the questionnaire and an envelope and asked to complete and return the questionnaire to the interviewer during the home visit. The PSAQ was only given if the parent had sufficient privacy to complete it; that is, the spouse, partner, or another adult was not in the immediate interviewing area. Self-administered questionnaires were available in English and Spanish. The PSAQ was not administered in a household where the parent interview was conducted in a language other than English or Spanish.

2.3 Father Self-Administered Questionnaires

Because of the important role that fathers play in their children's lives, the ECLS-B collects data from sampled children's fathers. The ECLS-B is one of the few national U.S. studies to involve fathers through self-reporting, asking about their attitudes about being fathers and activities they do with their children during the early years of their children's lives. The father component of the ECLS-B provides an opportunity to examine the unique contributions fathers make to children's well-being and developmental outcomes over time.

At the 9-month data collection, data from fathers were collected directly through two separate father questionnaires: the resident father questionnaire and the nonresident biological father questionnaire. Both resident and nonresident fathers were identified during the parent interview. The resident father was the person who resided in the household and who was either the child's biological father or the person identified as the partner or spouse of the parent interview respondent (i.e., is the father or plays an important role in the child's life). The nonresident father questionnaire was administered to a biological father (not residing in the household with the child) who had contact with the child or mother at least once in the previous three months.[4] Contact was defined as a telephone call or an in-person visit. Therefore, an individual child in the ECLS-B could potentially have various types of father data available:

- No father data;
- Resident father questionnaire data only;
- Nonresident biological father questionnaire data only; or
- Both resident father and nonresident biological father questionnaire data.

The content of the two father questionnaires was similar in key areas (e.g., activities with the child, feelings about being a father, education, employment) and unique in others (e.g., separations from child for resident fathers, child support and legal arrangements for nonresident fathers). Fathers of sampled twin cases only completed one questionnaire. Exhibit 2-4 includes the topics included in the 9-month resident and nonresident father questionnaires.

[4] More specifically, a nonresident biological father was eligible for the nonresident father questionnaire only if the respondent to the parent interview was the child's biological mother, the mother granted permission to contact the father, and the nonresident biological father had (a) seen the child once in the last month or at least 7 days in the last 3 months or (b) the biological mother had been in touch with the biological father at least once a month in the past 3 months.

Exhibit 2-4. ECLS-B 9-month father self-administered questionnaires by major content topics: 2001–02

Topic	Content	Resident	Nonresident
Activities with and feelings about child	▪ Frequency engages in activities with child (e.g., sing songs)	✓	
	▪ Feelings about child/delight in child (e.g., talk about child, show pictures)	✓	✓
Time spent with child	▪ Last time saw child		✓
	▪ Frequency and amount of time spent with child		✓
Knowledge about child development	▪ General knowledge of child development	✓	
Parenting practices	▪ Frequency looks after child	✓	✓
	▪ Amount of influence in decisions (e.g., child care)	✓	✓
Separations from child	▪ Frequency separated from child overnight	✓	
	▪ Frequency separated from child for a week or more	✓	
Prenatal/neonatal experiences	▪ Wantedness of child	✓	✓
	▪ When first held child	✓	✓
	▪ Activities prior to baby's birth (e.g., see sonogram)	✓	
	▪ Present at birth	✓	
Attitudes about being a father	▪ Importance of different things fathers do for their children	✓	
	▪ Self-rating as a father	✓	✓
	▪ Attitudes towards demands of fatherhood	✓	
	▪ Attitude towards future involvement in raising child		✓
Current relationship with spouse/child's mother	▪ Quality of current marriage/relationship	✓	
	▪ Frequency and types of arguments	✓	✓
	▪ Conflict resolution style		✓
	▪ Nonresident father's quality of relationship with child's mother		✓
	▪ Frequency of conversations about child		✓
Father's childbearing and marital/partner history	▪ Number of times married	✓	
	▪ Number of biological children	✓	✓
	▪ Age when his first child was born	✓	✓
	▪ Number of biological children that live outside household	✓	
	▪ Child support for children living elsewhere	✓	
Nonresident father involvement for child and household	▪ Frequency of selected things done for child (e.g., buy clothes, diapers)		✓
	▪ Frequency of selected things done for child's household (e.g., help with repairs)		✓

See note at end of exhibit.

Exhibit 2-4. ECLS-B 9-month father self-administered questionnaires by major content topics: 2001–02—Continued

Topic	Content	Resident	Nonresident
Legal arrangements and child support	▪ Type of child support agreement (e.g., formal, informal)		✓
	▪ Estimated child support payment		✓
	▪ Actual child support payment		✓
Background information	▪ Month and year born	✓	✓
	▪ Country of origin	✓	✓
	▪ Citizenship status	✓	✓
	▪ Age first moved to the United States	✓	
	▪ First language learned to speak	✓	
	▪ English literacy level	✓	
Education and employment	▪ Educational attainment	✓	✓
	▪ School experiences	✓	✓
	▪ Current employment/unemployment, characteristics of job (e.g., duties, job title)	✓	✓
	▪ Current participation in education or training	✓	
	▪ Benefits through current job	✓	
Health	▪ Overall health	✓	
	▪ Depression	✓	✓
	▪ Mental health needs	✓	
	▪ Current smoking behavior	✓	
	▪ Current drinking behavior	✓	
	▪ Stressful life events	✓	
Social support network	▪ Father's closeness to his parents	✓	
	▪ Frequency attends religious services	✓	
	▪ Participation in community service activities	✓	
	▪ Frequency of social gatherings with friends	✓	
Family background	▪ Family composition growing up	✓	
	▪ Receipt of public assistance while growing up	✓	
	▪ Educational attainment of father's parents	✓	
Current living arrangement	▪ Number of other people living in household		✓
	▪ Total income of household		✓
Income	▪ Total income last year		✓

SOURCE: U.S. Department of Education, National Center for Education Statistics, Early Childhood Longitudinal Study, Birth Cohort (ECLS-B) Nine-Month Data Collection, 2001–02.

2.3.1 Resident Father Self-Administered Questionnaire

Historically, fathers were assumed to be on the periphery of children's lives and, therefore, of little direct importance to children's development. Research on fathers has focused mostly on their role as providers. Recently, however, researchers and policymakers are beginning to recognize that fathers influence their children's development in a variety of ways (Lamb, 1997).[5] Yet, little is known about what fathers think about being fathers or about their relationship with their children. What is known has often been obtained from mothers rather than fathers directly (Marsiglio and Day, 1997; The Working Group on the Methodology of Studying Fathers, 1997). For many topics, however, fathers are better sources of information than mothers.

The ECLS-B 9-month resident father questionnaire was designed to directly collect information from fathers about the role that they played in their children's lives. The resident father was the person identified as the partner or spouse of the parent interview respondent. In rare cases, this person could be female. As the partner or spouse of the parent respondent, the person completing the resident father questionnaire could be the child's grandfather if the child's grandmother was the parent respondent; or the male or female partner of the child's mother.[6] The first item in the questionnaire identified the relationship of the person completing the questionnaire to the sampled child (e.g., birth father). At 9 months, 98.1 percent of persons completing the resident father questionnaire were the child's biological father.

The resident father questionnaire is entitled "Questions for Fathers and Other Important People"; nevertheless, since in the majority of cases the person completing the questionnaire was the child's father, it is referred to throughout this document as the resident father questionnaire. During the parent interview, if the respondent had a partner or spouse residing in the household, then the parent CAPI instrument would prompt the interviewer to present the parent interview respondent (or the father if he was home) with the resident father questionnaire and a letter explaining the study and the instrument. The resident father questionnaire was available in English and Spanish. If an interpreter was used for the

---

[5] This is further seen in the effort by researchers and federal agencies to coordinate measures on father involvement. The Developing a Daddy Survey (DADS) consortium works to coordinate the design and development of measures of father involvement. See "Bridging research and policy: Including fathers of young children in national studies" (Cabrera et al., 2002) for more information.

[6] In developing the weights for father respondents, 15 completed father self-administered questionnaires were deleted. These cases were ones in which the biological father was the respondent to the parent CAPI questionnaire and the mother responded to the father questionnaire. Although the design allowed for this situation, there was no straightforward way given the design of the weights to weight those 15 mother respondents to the father questionnaire.

home visit, and the father was home, interpreters could translate and administer the English version of the resident father questionnaire to the father.

The 20-minute resident father questionnaire focused on those things unique to fathers and fathering that only fathers could answer (e.g., feelings about being a father, the father's influence on child care decisions). Where appropriate (e.g., constructs overlapped), the resident father questionnaire paralleled the content of the parent CAPI instrument to capture information such as activities with the child, knowledge about child development, and separations from the child as well as information on education, employment, and physical and mental health. In the few cases where the father was the main parent respondent (i.e., completed the parent CAPI instrument and PSAQ), questions unique to the resident father questionnaire were included as part of the parent CAPI instrument. See exhibit 2-4 for information on the constructs included in the 9-month resident father questionnaire.

2.3.2 Nonresident Biological Father Self-Administered Questionnaire

Many fathers do not live with their children. Nearly one-third of all children are born outside of marriage, and the majority of these children do not live with their fathers (Ventura and Bachrach, 2000). The high incidence of divorce and separation in this country leads to an increasing number of children living apart from their fathers as they grow older (U.S. Department of Education, National Center for Education Statistics, 1997). Although many fathers who do live near their children lose contact with them over time and tend to play a smaller role in their children's lives, a significant proportion of nonresident fathers remain involved. This involvement is important to children's lives (Le Menestrel, 1999; Liebman and Abell, 2000; U.S. Department of Education, National Center for Education Statistics, 1997; 2001). To understand children's development, it is, therefore, important to learn more about fathers who live apart from their children.

The ECLS-B 9-month nonresident biological father questionnaire (nonresident father questionnaire) was designed to collect information about the role that fathers play in their children's lives, even if they do not live with their children. As mentioned earlier, the nonresident father questionnaire was administered to the child's biological father (who did not reside in the household with the child) who had contact with the child or mother in the previous three months. Contact was defined as a telephone call or an in-person visit. For the nonresident biological father to be eligible for the nonresident father questionnaire, the biological mother had to be the respondent for the parent interview and she had to grant permission for the nonresident biological father to be contacted.

The questionnaire is formally entitled “Questions for Fathers.” During the parent interview, it was determined whether the sampled child’s biological father lived in the household, and if not, the level of contact he had with the mother and/or child. If the father was eligible[7] for the nonresident father questionnaire, permission to contact the child’s father and contact information was obtained from the mother. The nonresident father questionnaire was available in English and Spanish. Chapter 5 includes information on administration of the nonresident father questionnaire.

The nonresident father questionnaire was a 10-minute questionnaire that included items about child support, time spent with the child, activities with child, relationship with the child’s mother, feelings about himself as a father, and how he helped the family and child as well as information on education, employment, mental health, and income. Many of the items are the same as those found in the resident father questionnaire. Exhibit 2-4 contains information on the constructs included in the 9-month nonresident father questionnaire.

2.4 Child Observations and Interviewer Remarks

After the home visit was completed, the interviewer completed the Child Observations, which were two sets of questions about the child’s behavior during the child assessment and about the home environment. The Child Observations collected additional information about the children’s home environment and the children’s behavior without increasing burden on the family. They also provided another perspective of the child and home, as the interviewer served as an external observer, as opposed to solely parent-report.

The first set of questions was seven items selected from the Behavior Rating Scale (BRS) of the Bayley Scales of Infant Development, Second Edition (Bayley, 1993) that evaluated the sampled child’s behavior during the direct child assessment. They provided information on children’s positive and negative affect (or emotional behaviors, e.g., smiling, laughing, fussing, crying) in a testing situation. The

[7] A nonresident biological father was eligible for the nonresident father questionnaire if he had a certain level of contact. This was defined as seeing the child in the last month or at least 7 days in the last 3 months *or* contacting the biological mother at least once a month in the past 3 months. The nonresident father information was only collected when the respondent to the parent interview was the child’s biological mother and she granted permission for the father to be contacted.

items also assessed children's adaptation to a change in test materials (e.g., relinquishing items willingly), their interest in materials, and their attention to tasks. Information on children's social engagement with the interviewer or parent and their control of muscle movement was gathered as well, thus providing additional data on children's social and physical development.

The second set of questions included eight items from the Home Observation for Measurement of the Environment (HOME) Inventory (Caldwell and Bradley, 1984) that assessed the quality of various characteristics of children's home environments. These items considered aspects of the home including parent behavior toward the child (e.g., talking with the child, caressing the child, spanking), the parent's structuring of the home environment (e.g., allowing exploration, providing toys), and the safety of the home environment.

Finally, the interviewer remarks questionnaire was a self-administered, computer-assisted instrument that included questions about the order of the home visit tasks and other aspects of the data collection. These questions were designed to help in the evaluation of the home visit. Both the Child Observations and Interviewer Remarks components were completed on the interviewer's laptop computer as soon as possible after the interviewer left the child's home.

2.5 Birth Certificate Data

The ECLS-B used birth certificates of live births in 2001 as the sampling frame (see chapter 4 for more information). This sampling procedure provided access to data on prenatal and neonatal characteristics at the time of the child's birth. Children's birth certificates contained a variety of useful data about the children, their mothers, and their families. They contained information on the date of birth, birth weight, gender, plurality (e.g., if multiple birth and birth order), and prematurity as well as on parents' education, parents' race and ethnicity (including Hispanic origin), and mother's marital status. Birth certificates also included information on the mother's pregnancy history, prenatal care, medical and other risk factors during this pregnancy, and complications during labor and birth. Information on the obstetric procedures and delivery method was also available. Health characteristics of children, such as congenital anomalies, abnormal conditions of the baby, and the baby's APGAR score, were included. Selected variables from the birth certificates were included in the data set, and the birth certificate data were appended to the ECLS-B 9-month child record.

REFERENCES

Arend, R., Gove, F.L., and Sroufe, L.A. (1979). Continuity of individual adaptation from infancy to kindergarten: A predictive study of ego-resiliency and curiosity in preschoolers. *Child Development*, 50, 950-959.

Bayley, N. (1993). *Bayley Scales of Infant Development, Second Edition Manual*. San Antonio, TX: The Psychological Corporation.

Bus, A.G., and Van IJzendoorn, M.H. (1988). Mother-child interactions, attachment, and emergent literacy: A cross-sectional study. *Child Development*, 59, 1262-1272.

Cabrera, N., Brooks-Gunn, J., Moore, K., West, J., Boller, K., and Tamis-LeMonda, C. S. (2002). Bridging research and policy: Including fathers of young children in national studies in C.S. Tamis-LeMonda and N. Cabrera (eds.) *Handbook of Father Involvement: Multidisciplinary Perspectives* (pp. 489–523). Mahwah, NJ: Lawrence Erlbaum Associates.

Caldwell, B. M. and Bradley, R. H. (1984). *Home Observation for Measurement of the Environment*. Little Rock, AR: University of Arkansas Little Rock.

Hartup, W. (1985). Relationships and their significance in cognitive development. In R. Hinde and A. Perret-Clermont (Eds.), *Relationships and Cognitive Development* (pp. 67-82). Oxford: Oxford University Press.

Karen, R. (1998). *Becoming Attached.* Oxford: Oxford University Press.

Lamb, M. (Ed.) (1997). *The Role of the Father in Child Development (Third Edition)*. New York: Wiley.

Le Menestrel, S. (1999). *What Do Fathers Contribute to Children's Well-Being?* Washington DC: Child Trends.

Liebman, S. J. and Abell, S. C. (2000). The forgotten parent no more: A psychoanalytic reconsideration of fatherhood. *Psychoanalytic Psychology,* 17 (1), 88–105.

Marsiglio, W. and Day, R. (March, 1997). *Social Fatherhood and Paternal Involvement: Conceptual, Data, and Policymaking Issues*. Paper presented at the NICHD-sponsored Conference on Fathering and Male Fertility: Improving Data and Research.

Matas, L., Arend, R. A., and Sroufe, L. A. (1978). Continuity and adaptation in the second year: The relationship between quality of attachment and later competence. *Child Development*, 49, 547-566.

Snow, C.E. (1994). Beginning from baby talk: Twenty years of research on input and interaction. In C. Galloway and B. Richards (Eds.), *Input and Interaction in Language Acquisition* (pp. 3-12). London: Cambridge University Press

Sumner, G. and Spietz, A. (1994). *NCATS Caregiver/Parent-Child Interaction Teaching Manual*. Seattle, WA: NCAST Publications, University of Washington, School of Nursing.

U.S. Department of Education, National Center for Education Statistics. (1997). *Fathers' Involvement in Their Children's Schools* (NCES 98–091), by C.W. Nord, D. Brimhall, and J. West. Washington, DC: U.S. Department of Education.

U.S. Department of Education, National Center for Education Statistics. (2001). *Fathers' and Mothers' Involvement in Their Children's Schools by Family Type and Resident Status* (NCES 2001–032), by C.W. Nord and J. West. Washington, DC: U.S. Department of Education.

U.S. Department of Education, National Center for Education Statistics. (Forthcoming). *Early Childhood Longitudinal Study, Birth Cohort (ECLS-B), Methodology Report for the Nine-Month Data Collection, Volume 1: Psychometric Characteristics.* Washington, DC: U.S. Department of Education.

Ventura, S. J. and Bachrach, C. A. (2000). Nonmarital Childbearing in the United States, 1940-1999. *National Vital Statistics Reports*, 48 (16). Hyattsville, MD: National Center for Health Statistics.

Whitehurst, G. J., Arnold, D.S., Epstein, J.N., Angell, A.L., et al. (1994). A picture book reading intervention in day care and home from children from low-income families. *Developmental Psychology*, 30, 679-689.

The Working Group on the Methodology of Studying Fathers. (March, 1997). *Methodological Issues in Improving Data on Father.* Paper prepared for the NICHD-sponsored Conference on Fathering and Male Fertility: Improving Data and Research.

3. ASSESSMENTS, SCALES, AND OTHER MEASURES IN THE NINE-MONTH DATA COLLECTION

In the Early Childhood Longitudinal Study, Birth Cohort (ECLS-B), 9-month wave, information about children's cognitive, socioemotional, psychomotor, and physical development was obtained through both direct and indirect assessments. This chapter describes the scores developed from the direct assessments and the scales derived from the indirect assessments. Information on the direct child assessments is presented first, including descriptions of the scores, descriptive statistics, and guidelines for using the scores. Then, the items comprising the indirect assessments in the parent computer-assisted personal interviewing (CAPI) instrument are summarized.

The 9-month direct child assessments include the following:

- The Bayley Short Form–Research Edition (BSF-R);
- The Nursing Child Assessment Teaching Scale (NCATS);
- The Child Observations during the Bayley Short Form–Research Edition; and
- Physical measurements of children's length, weight, middle upper arm circumference (MUAC), and head circumference for only very low birth weight children (less than 1,500 grams).

The indirect child assessments (in the parent CAPI instrument) include the following:

- Items adapted from the Minnesota Child Development Inventory (MN-CDI) and
- Selected items from the Infant/Toddler Symptom Checklist (ITSC).

In addition, the ECLS-B contains selected items drawn from existing scales that measure the home environment and parent attitudes and knowledge about childrearing. These measures include the following:

- Knowledge of Infant Development Inventory (KIDI);
- Home Observation for Measurement of the Environment (HOME) Inventory;
- Family Practices Scale;
- Father's Perspective on Role of Fathers in Children's Lives;

- How Father Rates Himself as a Father (single item);[1]
- Parent Investment in Children;
- Early Head Start Involvement with Child; and
- Center for Epidemiologic Studies Depression Scale (CES-D).

Descriptions of the scores for the direct assessments and items for the indirect child assessments follow, along with variable names, variable descriptions, and descriptive statistics from the ECLS-B data file. Guidelines for using these scores are also provided in this chapter. Information on the selected items that obtained information about the respondent's parenting attitudes, beliefs and knowledge, family routines, and child's home environment is presented in section 3.3.2.

3.l Direct Child Assessments

The following section describes the scores that were obtained from the direct assessments of the child. This includes the BSF-R, the NCATS, and children's physical measurements (length, weight, MUAC, and, for very low birth weight babies only, head circumference).

3.1.1 Bayley Short Form–Research Edition

The Bayley Short Form–Research Edition[2] is derived from the Bayley Scales of Infant Development, Second Edition (BSID-II; Bayley, 1993), which is a standardized assessment of developmental status for children from birth to 42 months of age. A brief comparison of the BSF-R with the BSID-II is provided in this section. A more detailed summary of item selection criteria and procedures and the statistical analyses used to develop the BSF-R is presented in section 3.1.1.2. A full description of the development of the BSF-R can be found in *Early Childhood Longitudinal Study, Birth Cohort (ECLS-*

[1] Adapted and reproduced by special permission of the publisher, Psychological Assessment Resources, Inc., 16204 North Florida Avenue, Lutz, FL 33549, from the Parenting Stress Index–Short Form by Richard R. Abidin, Ed.D., Copyright 1990 by PAR, Inc. Further reproduction is prohibited without permission from PAR, Inc.
[2] Bayley Short Form–Research Edition. © Copyright 2001 by The Psychological Corporation, a Harcourt Assessment company. Adapted from the "*Bayley Scales of Infant Development: Second Edition*" © Copyright 1993 by The Psychological Corporation. Adapted and reproduced by permission of the publisher. All rights reserved.
"*Bayley Scales of Infant Development*" is a trademark of The Psychological Corporation, registered in the United States of America and /or other jurisdictions.

B), Methodology Report for the Nine-Month Data Collection, Volume 1: Psychometric Characteristics (U.S. Department of Education, National Center for Education Statistics, forthcoming [b]).

The BSID-II contains two main scales: a mental scale and a motor scale. The mental scale of the BSID-II measures children's cognitive development, for example, memory, means-end behavior, exploratory competence, object permanence, expressive communication, and receptive communication, among others. Although the BSID-II mental scale contains items that come from a range of domains, factor analytic studies (e.g., see Burns, Burns, and Kabacoff, 1992), primarily carried out using the first edition of the BSID (Bayley, 1969) with selected groups of children, suggest predominance of a single general factor—mental ability—through 2 years of age, with some evidence suggesting the presence of weak subfactors. Beyond 2 years, presumably with the acquisition of language, factor analytic studies provide evidence of a more complex factor structure.

The factor analytic studies were replicated by Westat analysts using BSID-II standardization data. These analyses confirmed a single factor of mental ability from birth through about 28 months of age on the mental scale of the BSID-II. The presence of a strong principal factor of mental ability implies that items can be selected at roughly equal intervals from across the ability distribution of the BSID-II (in the target age range) to represent the BSID-II mental scale as a whole. In addition, every effort was made to include items from as many content domains as possible, but only if the item possessed strong psychometric properties. For example, although puzzle completion items formed their own content domain, they were not retained because their psychometric properties were considered inadequate.

The motor scale of the BSID-II measures children's psychomotor development, including fine motor development and gross motor development. Fine motor development includes skills that use small muscle groups (e.g., fingers and hands), such as reaching and grasping, manipulating small objects, and using a pencil. Gross motor development uses the large muscle groups and includes such skills as sitting, standing, walking, and balance.

The BSID-II items follow a developmental progression (i.e., beginning with easier items for babies and extending through harder items for preschoolers). The test is organized in age item sets, groups of items that psychometric analyses have shown to apply within certain age ranges (based on chronological age). When the BSID-II is administered to children at 9 months of age, the mental scale 9-month age set includes 25 items and the motor scale 9-month age set includes 14 items. However, the child's performance determines whether the mental scale or the motor scale 9-month age sets need to be supplemented with one or more earlier (basal) age sets or with one or more older (ceiling) age sets.

The full BSID-II was administered during the first ECLS-B field test conducted in fall 1999.[3] Because administration of the BSID-II during the ECLS-B home visit was too burdensome, the decision was made to develop a shortened version of the BSID-II, called the Bayley Short Form–Research Edition (BSF-R; The Psychological Corporation, 2001). Items selected for the BSF-R mental and motor scales were organized into a core set of items that would be administered to all the children in the ECLS-B, analogous to the 9-month age set of the BSID-II. Children's performance on this core set determines whether any supplementary sets of items should be administered. The core set for the BSF-R mental scale includes 11 items (excluding 2 additional spontaneous vocalization items that are scored solely by observation and independently of the mental core items for all children). In addition, a single basal set of 9 items was created that covered a broad range of younger ages, as well as a single ceiling set of 9 items that covered a broad range of older ages. The mental scale core items range in age from 5 months to 13 months, the basal items reach down to 4 months, and ceiling items extend up to 19 months. (Because the spontaneous vocalization items were scored entirely by observation, they were not considered part of the administered mental scale; however their scores were added to the mental scale core items for a total of 13 core scores.)

The core set for the BSF-R motor scale includes 14 items, with 11 additional items in the basal set and 10 additional items in the ceiling set. (The motor scale core items range in age from 5 to 13 months, the basal items reach down to 3 months, and ceiling items extend as high as 22 months). On the motor scale, many items could be scored by direct observation of the child's motor abilities or by clustering similar items into a single administration, bringing down the number of administrations to 8 in the core set, 4 administrations in the basal set, and 5 administrations in the ceiling set. The rules used to determine whether to administer either the basal set or the ceiling set were simplified from those used in the BSID-II so that all decisions were made on the basis of the number of items receiving credit. Table 3-1 compares the content and structure of the BSF-R with the BSID-II.

[3] A test of the 9-month design was conducted during the 1999 field test with a sample of approximately 1,500 cases. A key feature of the 9-month data collection was the administration of the full BSID-II by field staff during the home visit. Problems implementing the design were soon apparent. Interviewer attrition occurred at more than twice the expected rate, and production was less than half the expected rate. In addition, hours per completed interview were about 50 percent higher than expected. The BSID-II was obtained on only about 50 percent of the field test sample. Because the 9-month data collection time-in-home and the total time per completed case were so much higher than expected, and because it proved too difficult for many interviewers to complete the work, it was decided to revise the design to make it shorter and simpler. Therefore, it was decided to develop a shortened form of the BSID-II that would include only a reduced set of items derived from the full BSID-II.

Table 3-1. Comparison of numbers of items per set of the Bayley Short Form–Research Edition and the Bayley Scales of Infant Development, Second Edition: 2001–02

Item set	Mental scale BSF-R	Mental scale BSID-II	Motor scale BSF-R	Motor scale BSID-II
Core or 9-month age set	13 items	25 items	14 items	14 items
Basal set				
1st set	9 items	3 items	11 items	7 items
2nd set	†	5 items	†	7 items
Total core + basal sets	22 items	28-33 items	25 items	21-28 items
Ceiling set				
1st set	9 items	1 item	10 items	2 items
2nd set	†	5 items	†	5 items
Total core + ceiling sets	22 items	26-31 items	24 items	16-21 items

† Not applicable.

NOTE: The BSF-R only had one set of basal items and one set of ceiling items. In the BSID-II additional sets would be administered until the appropriate criterion was met: On the mental scale, the tester would administer additional basal sets until the child received “credit” for five or more items within a set, or additional ceiling sets until the child reached “no credit” for three or more items. On the motor scale, the tester would administer additional basal sets until the child received “credit” for four or more items, or additional ceiling sets until the child received “no credit” for two or more items. To match the age span of the BSF-R, from basal to ceiling, the BSID-II mental scale would include 12 age sets and the BSID-II motor scale would include 14 age sets.

SOURCE: U.S. Department of Education, National Center for Education Statistics, Early Childhood Longitudinal Study, Birth Cohort, Nine-Month Data Collection, 2001–02.

Interviewers recorded the scores for each item in the Child Activity Booklet, a hard-copy booklet with score sheets, administration instructions, and scoring instructions. The item-level scores (credit/no credit) were data-entered at Westat. Item response theory (IRT) calibration and scoring were used to develop the scores that are described below.

The ECLS-B 9-month data file contains three basic types of IRT scores based on the BSF-R mental scale and the BSF-R motor scale: scale scores, standardized T-scores, and proficiency probabilities. Each of these scores is provided for both the mental scale and the motor scale. These scores are described below; the appropriate use of these scores is discussed in section 3.1.1.1. Table 3-2 presents the variable name, label, and descriptive statistics for all the BSF-R scores.

The BSF-R was designed to assess the development of children 7 months 16 days (which rounds to 8 months using BSID-II test administration rules)[4] to 11 months 15 days (which rounds to 11

[4] To comply with BSID-II administration procedures, the BSF-R was designed to be administered to children from 8 to 11 months, rounded to the nearest the half-month. That is, 7 months, 16 days through 8 months, 15 days is considered 8 months, etc. This is because the BSID-II is based

months) of age. Most of the children (83 percent) were assessed within this age range. The majority of out-of-range children were older than the intended age because of the delayed start of the field period in 2001 and the time that was required to locate families and to schedule and conduct the home visit. As a result, children ranged in age from approximately 6 months to 22 months at the time of the assessment although, as noted above, the majority (83 percent) were within the target range of 8–11 months.

Scale Scores. As mentioned earlier, the BSF-R is an adaptation of the BSID-II. The ECLS-B used a subset of items from the full BSID-II to form the BSF-R. IRT true-score equating was used to report ECLS-B results on the same scale used by the full BSID-II. This makes it possible to compare ECLS-B results with those obtained using the full set of BSID-II items. IRT uses item difficulty and discrimination parameters and the pattern of correct and incorrect responses to determine an ability estimate for each child (using maximum likelihood). This estimate is then used to determine the number-right raw score the child would have obtained had the full BSID-II been administered.

In theory, the ECLS-B mental scale score could range from 0 to 178, and the ECLS-B motor scale score could range from 0 to 111, the same raw score ranges as found in the BSID-II. At 9 months, the raw score for the BSF-R mental scale ranges from 0 to 13 for the core set of items administered to all children, and the motor scale raw score ranges from 0 to 14 for the core set. (On the full BSID-II, the 9-month age set includes 25 mental items and 14 motor items so that raw scores could range from 0 to 25 on the mental scale and 0 to 14 on the motor scale on condition that only the 9-month age set items were required.)

Standard Errors. The standard error of measurement (SEM) represents the precision of the IRT true-score. The smaller the standard error, the greater the precision of measurement. The true-score standard error is obtained by translating the standard error of BSF-R IRT ability estimates into the corresponding BSID-II raw score metric.

Standard errors for IRT ability estimates are based on test information. An important postulate of IRT models is that test information depends entirely on the independent and additive contributions of individual items. The contribution of an individual item is the square of the derivative of the item characteristic curve (ICC), divided by its variance at a given level of ability θ. This contribution is known as the *item information function*:

---

on item age sets, i.e., the “8-month” item age set is intended for children from 7 months, 16 days through 8 months, 15 days, etc. Therefore, the target age range for the BSF-R is 7 months, 16 days through 11 months, 15 days.

$$I(\theta, x_i) = \frac{P_i'^2}{P_i Q_i}.$$

Test information is the sum of item information over all of the *n* items in a test:

$$I(\theta, \mathbf{x}) = \sum_i^n I(\theta, x_i) = \sum_i^n \frac{P_i'^2}{P_i Q_i}.$$

The standard error for the maximum likelihood estimator of ability is the square root reciprocal of the test information function at $\hat{\theta}$:

$$\text{se}(\hat{\theta}) = \sqrt{1/I(\theta, \mathbf{x})}\ .$$

In the ECLS-B, IRT true-scores are used to report results in publisher raw-score metric for comparison purposes. Since these scores are very nearly linear in the case of the BSID-II over the age range studied in the ECLS-B, standard errors for IRT true-scores are obtained by projecting $\text{se}(\hat{\theta})$ onto the same scale metric. For BSF-R reporting at 9 months, IRT true-scores have a number-right mean score of 77.5 items, with an average standard error of ±3.7 items for the sample as a whole. The user should expect to find larger standard errors at the tails of the ability distribution, where standard errors reach to ±4.3 items among the youngest and least able children, while rising to ±7.6 items among the oldest and most able. For the motor scale, IRT true-scores have a number-right mean score of 56.8 items, with an average standard error of ±1.9 items for the sample as a whole. At the extremes of the distribution, these errors reach ±3.1 items among the youngest and least able, while rising to ±6.6 items among the oldest and most able.

Standard errors can be used to determine the probability that any given child has surpassed a specific threshold. For software such as Hierarchical Linear Modeling (HLM) that permits the use of plausible values, plausible values can be generated by taking a series of random draws from within the normal distribution defined by the number-right raw score and the standard error.

Some children in the ECLS-B sample were administered the BSF-R beyond the recommended age range. The standard error of measurement may be high in these cases. To limit the impact of these cases in an analysis, the analyst may choose to weight all individual cases that are part of an analysis by the inverse of the error variance (i.e., the square of the standard error of measurement) so that the variation in measurement precision will be acknowledged when calculating point estimates.

Standardized Scores (T-Scores). Standardized scores are designed to enable individual comparisons with the corresponding age-reference population. The T-score is based on biological age, adjusted for prematurity as necessary. In the ECLS-B, the standardized T-score represents the position of the child in the corresponding age group of the reference population. The BSF-R mental and motor T-scores have a mean of 50 and a standard deviation of 10. Consequently, a T-score in excess of 50 implies that the child's performance level exceeds the average for children of the same age. This does not mean that the child has mastered a particular set of skills, only that his or her performance exceeds the average for children of the same age. The same line of reasoning applies to groups of children. The average T-score for a group of children can be used to determine whether they are above or below the average level of performance for children of similar age.

BSF-R mental and motor T-scores provided in the ECLS-B data file are scaled to have a mean of 50 and a standard deviation of 10. Mental and motor age-normed developmental index scores in N (50, 10) metric are normed to the ECLS-B population. Age-adjusted developmental index scores are provided for children of all ages. A –99 is used to indicate a missing T-score because –9, which is typically used to indicate a missing composite, can be a valid T-score. A –99, however, cannot be a valid T-score.

Users will note an interesting pattern if they examine the standardized scores by children's age. Children assessed at progressively older ages have lower T-scores than those assessed at younger ages. This pattern does not indicate that the progressively older children are less able. On average, they are much more able than younger children. What the standardized scores show is that at progressively older ages, the children in the ECLS-B sample are less developed than other children of the same age (i.e., the children in the publisher's standardization sample). It is not possible to define a particular age beyond which the scores decline, as if there were a threshold. Rather, the scores show a general, steady decline from 8 months and beyond. Reasons why children were assessed at older ages include the following: difficult to locate; initial refusal; initial language difficulty; initial out of country, state, or normal residence; and illness. Some of these factors do make an independent contribution to lower standardized scores, but they do not explain the age-ability relationship observed. At present, the study cannot explain why children who were interviewed late are less developed mentally than children of the same age.

Proficiency Probabilities. Proficiency probabilities represent mastery of a specific skill or ability within a developmental domain (i.e., mental and motor) and provide information about whether children have reached key milestones thought to be of interest to analysts. As with the scale scores and T-scores, the proficiency probabilities are based on information from the IRT model that equates the BSF-R to the full BSID-II. Proficiency probabilities can take on any values between 0.0 and 1.0 and represent the likelihood that a child has surpassed a given developmental milestone. They provide a means of distinguishing status or gain in specific skills. These probabilities are computed using the overall performance on each respective test.

Milestones are hierarchical in the sense that later developing skills are built upon, and surpass, earlier acquired skills. The proficiency probabilities represent a succession of milestones that describe a developmental sequence during the course of development. This is necessarily so because the proficiency probabilities are built on clusters of items at different difficulty levels in the IRT model. This does not necessarily imply, however, that the milestones are placed at equal developmental intervals. Item clusters, typically of four to six items with similar content and difficulty to represent a specific skill, were selected at several points along the scale score continuum. Each cluster was then treated as a single item in order to estimate the probability of mastery of each skill. The hierarchical nature of skill item sets justifies using the IRT model in this fashion. The items follow a Guttman model, where a child who is able to complete a given task is expected to have mastered tasks at lower levels of ability; a failure to complete a given task implies nonmastery of items at higher levels of ability.

Because the items in each cluster range from easier to harder, if a child receives a score of .25 it means that the child has only just begun to acquire this skill. On the other hand, a child who receives a score of .75 has mastered most of the related items and is likely to be skilled in this cluster. For example, on the sitting cluster, if a child's score implies that she can successfully complete *only the first item*, then she would receive a score of .17, implying that the child is unlikely to have mastered this skill level. On the other hand, if a child's score indicates that she can *complete four out of six* of the related items, then she would obtain a proficiency probability of .67, indicating that she is more likely to have attained this skill level. This child may sit alone well, although not necessarily for long periods of time. Probabilities that are exactly 0 or 1 are rare, and imply that the child's developmental level is far removed from the milestone.

Since the proficiency probabilities are based on overall performance, they can be calculated for all children with scale scores, even though they have not necessarily been given an opportunity to attempt each of the tasks in the cluster. Item parameters estimated from the publisher standardization data set have been used in this way to calculate proficiency probabilities for the ECLS-B sample.

The proficiency probabilities in the 9-month data file are based on four, five, or six items. It is important to point out that the proficiency probabilities do not depend on the number of items. Rather, they represent the probability that a certain milestone has been reached or surpassed. The weighted sum of these probabilities represents the number of children in a group who have reached or surpassed a given milestone. The average of these probabilities represents the proportion of children in a group who have reached or surpassed a given milestone.

Proficiency probability scores can be averaged to produce estimates of mastery rates within population subgroups. The weighted sum of these probabilities represents the number of children in a group who have reached or surpassed a given milestone. The average of these probabilities represents the proportion of children in a group who have reached or surpassed a given milestone. From a longitudinal perspective, these continuous measures provide a close look at the change in individual status over time. Gains in the probability of mastery at each proficiency level allow researchers to study gains in specific outcomes over time. Thus, the exposure and experiences of children can be related to improvements in specific skills.

The five proficiency levels created for the BSF-R mental scale correspond to: exploration of objects; purposeful exploration; early problem solving; and two communications proficiency probabilities, one for babbling and one for word use.

The motor scale items fall into two broad domains: fine motor development and gross motor development. In addition, the motor scale includes items from clearly defined skill areas, such as sitting, pre-walking, and walking. These skill areas also follow the developmental course of psychomotor ability. These skill areas are reflected in the five proficiency levels for the motor scale: eye-hand coordination, sitting, pre-walking, walking, and balancing. It should be noted that the first two motor proficiency levels, eye-hand coordination and sitting, are intercorrelated, indicating that these milestones are at about the same developmental level.

Exhibit 3-1 describes the BSF-R mental scale and motor scale proficiencies.

Exhibit 3-1. BSF-R mental scale and motor scale proficiencies, 9-month data collection: 2001–02

BSF-R Mental Scale Proficiencies:

- **Explores objects.** This proficiency can be characterized as the ability to explore objects, for example, reaching for and holding objects, manipulating objects like a cup or a string, and banging objects in play. The child may have no specific purpose or goal except to play or discover. (6 items)

- **Explores purposefully.** This proficiency refers to children’s purposeful exploration of objects, that is, the child now touches and works with the objects for a reason. For example, the child may explore a bell to understand the source of the sound, engage in means-end behavior such as using a string to obtain an object, hold multiple blocks at one time to gather, or look at the pictures in a book. (5 items)

- **Babbles.** This proficiency assesses communication through diverse nonverbal sounds and gestures, for example, vowel and vowel-consonant sounds, gesturing for an object, babbling, and jabbering. (4 items)

- **Early problem solving.** This proficiency can be characterized as engaging in early problem solving types of reasoning, for example, using a tool to retrieve an out-of-reach toy, locating a hidden toy, or knowing how to use a pencil to make marks on paper. (4 items)

- **Uses words.** This proficiency measures communication using words, both receptive (knowing an object by its name) and expressive (saying the name of an object) communication. (4 items)

BSF-R Motor Scale Proficiencies:

- **Eye-hand coordination.** This proficiency refers to children’s ability to use visual tracking to guide hand movements to pick up a small object. This is a fine motor skill. (4 items)

- **Sitting.** This proficiency reflects children’s ability to maintain control of the muscles used in sitting with and without support. (6 items)

- **Pre-walking.** This proficiency measures children’s ability to engage in various pre-walking types of mobility, with and without support, such as shifting weight from one foot to another, making alternating stepping movements, or walking when holding onto furniture. (6 items)

- **Independent walking.** This proficiency measures children’s ability to walk and stand independently, without help from others or from holding onto anything. (5 items)

- **Balance.** This proficiency refers to children’s ability to maintain balance when changing position, for example, when squatting, shifting weight from side to side while standing, or standing on one foot. (4 items)

SOURCE: U.S. Department of Education, National Center for Education Statistics, Early Childhood Longitudinal Study, Birth Cohort, Nine Month Data Collection, 2001–02.

By reporting continuous probabilities, the analyst is free to determine what cut-off probability to use to establish mastery. Although some analysts prefer to use a 50 percent probability to represent a developmental threshold, the public generally expects mastery to imply that an overwhelming majority of children can successfully perform the tasks representing the developmental milestone. The ECLS-B reports the actual probabilities so that alternative mastery thresholds of 50, 60, or 80 percent can be defined by the analyst to represent mastery of the developmental milestone.

Table 3-2 summarizes the BSF-R variables and presents distributional information for each.

Researchers with an interest in children with disabilities may find the last four variables in table 3-2 useful in their research (X1MTLMOD, X1MTRMOD, X1MTLSKP, X1MTRSKP). These variables summarize the modifications and adjustments interviewers made when administering the BSF-R to children with known diagnoses of disabilities or developmental delays, e.g., Down syndrome, cerebral palsy. The BSID-II is often used as part of the diagnostic procedure for identifying children with disabilities. Although there are no standardized instructions for adapting the administration to take into consideration the special needs that children with disabilities may have, it is possible to modify items on a case-by-case basis to provide children an even chance to respond while adhering to the administration rules.

During training, interviewers were instructed about how to modify BSF-R items when working with a child with a disability or developmental delay, such as Down syndrome, blindness, or deafness. They were also trained to skip items when modifications were not feasible or appropriate. The purpose of the modifications was to give such children an even chance with other children to respond on each item. Approved modifications were provided in an item-by-item table, and interviewers were instructed to note any modifications in the Child Activity Booklet. When the Child Activity Booklets were returned to Westat, these modifications were reviewed to ensure their appropriateness. Four composites were created to indicate the number of core items modified or the number of core items skipped (and thus scored as “no credit”) for each scale (i.e., mental and motor). See chapter 3 of the *Early Childhood Longitudinal Study, Birth Cohort (ECLS-B) Design and Operations Report for the Nine-Month Data Collection* (U.S. Department of Education, National Center for Education Statistics, forthcoming [a]) for more information on both modifications made to accommodate special needs children and on training field staff to administer the BSF-R to special needs children.

Table 3-2. Bayley Short Form–Research Edition scores: 2001–02

Variable name	Variable label (description)	Range	Weighted mean[1]	Standard deviation
X1MTLSCL	X1 MENTAL SCALE SCORE	54.55-112.48	77.32	7.54
X1MTLSSE	X1 STAND ERR IRT TRUE MTL SCR	2.06-7.68	3.72	0.39
X1MTRSCL	X1 MOTOR SCALE SCORE	31.95-79.55	56.47	6.64
X1MTRSSE	X1 STAND ERR IRT TRUE MTR SCR	1.35-7.05	1.99	0.54
X1MTLTSC	X1 MENTAL T-SCORE	-36.11-117.57	50.00	10.00
X1MTRTSC	X1 MOTOR T-SCORE	-29.56-100.46	50.00	10.00
X1MTL1	X1 MENTAL PROB1: EXPLORES OBJECTS	0.73-1.00	0.99	0.01
X1MTL2	X1 MENTAL PROB2: EXPLORES PURPOSEFULLY	0.05-1.00	0.91	0.12
X1MTL3	X1 MENTAL PROB3: BABBLES	0.09-1.00	0.55	0.19
X1MTL4	X1 MENTAL PROB4: EARLY PROBLEM SOLVING	0.00-0.97	0.09	0.15
X1MTL5	X1 MENTAL PROB5: USES WORDS	0.00-0.93	0.03	0.09
X1MTR1	X1 MOTOR PROB1: EYE-HAND COORDINATION	0.28-1.00	0.92	0.07
X1MTR2	X1 MOTOR PROB2: SITTING	0.31-1.00	0.95	0.07
X1MTR3	X1 MOTOR PROB3: PRE-WALKING	0.05-1.00	0.79	0.20
X1MTR4	X1 MOTOR PROB4: INDEPENDENT WALKING	0.00-1.00	0.31	0.30
X1MTR5	X1 MOTOR PROB5: BALANCE	0.00-0.91	0.06	0.15
X1MTLMOD	X1 MENTAL: NUMBER OF ITEMS MODIFIED	0.00-11.00	0.00	0.09
X1MTRMOD	X1 MOTOR: NUMBER OF ITEMS MODIFIED	0.00-14.00	0.00	0.08
X1MTLSKP	X1 MENTAL: NUMBER OF ITEMS SKIPPED	0.00-10.00	0.01	0.18
X1MTRSKP	X1 MOTOR: NUMBER OF ITEMS SKIPPED	0.00-14.00	0.01	0.23

[1] The full sample respondent-child weight W1C0 was used to obtain the weighted mean.

SOURCE: U.S. Department of Education, National Center for Education Statistics, Early Childhood Longitudinal Study, Birth Cohort, Nine-Month Data Collection, 2001–02.

There is also a status variable for both the BSF-R mental scale and the BSF-R motor scale that indicates whether the assessment was administered with or without modifications, was not administered, or was unscoreable (X1STBMTL and X1STBMTR, respectively). At least 66 percent of core items on the BSF-R mental scale needed to be administered in order to be considered scoreable and thereby receive a scale score. The same is true for the BSF-R motor scale core items.

3.1.1.1 Choosing the Appropriate BSF-R Score for Analysis

Each of the scores described above presents information about children’s mental and motor developmental status from a slightly different perspective. The choice of the most appropriate score for analytic purposes should be driven by the context in which it is to be used and by the research question to be addressed in the analyses:

- Scale scores are measures of overall mental or overall motor ability;
- Standardized T-scores are indicators of relative ability at a single point in time, in relation to other children of the same age; and
- Proficiency probabilities are criterion-referenced indicators of status for a specific skill.

The **mental and motor scale scores** are overall measures of children’s developmental ability at a single point in time. Scale scores represent children’s absolute positions along a developmental continuum. Older children will naturally tend to have higher scale scores than younger children, if only because of their age. This makes scale scores ideal for studies involving longitudinal growth over time, where age is used as a measure of time in order to compare developmental growth rates.

Scale scores are useful to answer the question: At what ability level is the child functioning? Or, alternatively, what is the child’s general developmental level? The scale scores provide a summary measure of children’s mental and motor development and are useful in regression analyses where development is the dependent variable and characteristics such as family demographics (e.g., mother’s education) or family processes (e.g., playing with child) are the independent or explanatory variables. For example, the scale score would be an appropriate dependent variable when examining the relationship between prenatal care or nutritional status and development, or between home environment and development. Eventually, when subsequent waves of the ECLS-B data are released, the scale scores would also be used in longitudinal analyses of children's developmental growth.

Standardized T-scores are overall indicators of relative ability at a single point in time. They are norm-referenced within each age group to have a mean of 50 and a standard deviation of 10. T-scores are relative measures of development based on age-group norms. Older and younger children will have similar T-score values when they occupy the same relative position in their respective age-groups. Since age has no further bearing once the norming takes place, T-scores are ideal for cross-sectional comparisons examining the relative development of children at different ages.

The T-scores are good for answering the question: How does a particular group of children compare with their peers and with the general population of children born in the year 2001 and having similar demographic characteristics? For example, if an analyst is curious about how urban children's motor development compares to rural children's motor development or how both groups compare to children in the general population, the T-score would be a good choice for a comparison of means test. Since the mean of the general population's T-score will always be 50, the analyst can produce mean T-scores for urban children and rural children separately and then compare these means with the overall population mean. Because it is age-normed, the T-score is also useful for comparing scores between individuals of different ages. The T-distribution (i.e., mean of 50, standard deviation of 10) provides a standard "ruler" for comparing scores between individuals and between subgroups. With a standard deviation of 10, a 5-point difference in scores would mean a difference of one half of a standard deviation between two individuals, or between two subgroups. For example, a mean of 45 for one group and 50 for another group would be the equivalent of one-half of a standard deviation difference.

When subsequent waves of the ECLS-B data are released, the T-scores, because they are age-normed, could be used in a longitudinal fashion to illustrate the increase or decrease in gaps among subgroups over time, but in a somewhat limited manner. They are not recommended for calculation of a change or "gain" score. (The scale score is recommended for this.).

The **proficiency probabilities** represent status in relation to specific skills included in the BSF-R mental and motor scales. They are ideal for studying the stages of mental and motor development, such as children's exploration of objects or their early problem solving. Proficiency probabilities enable the analyst to gain a better understanding of where (e.g., in what skill sets) the child's developmental status gains are taking place, as opposed to the overall scale score. A comparison of the mean scores may also be used to address more specific questions, such as: How many children in a certain age or demographic group are exploring objects? What percentage of children is walking independently?

In addition, these scores can be used in non-linear regression analyses to investigate the relationship between the attainment of a particular milestone and, for example, the child's health status or home environment. When the data from later waves become available it will also be possible to investigate the longitudinal relationship between early achievement of milestones and subsequent cognitive outcomes. Although proficiency probabilities are continuous variables, as proportions they are by their very nature non-linear. Proficiency probabilities should be transformed into normal deviates or logits prior to their use in regression and correlation analyses.

3.1.1.2 Additional Information on the Computation and Reliabilities of the BSF-R Scores

In the context of the ECLS-B 9-month home visit with administration by lay assessors, the full BSID-II with its complicated age-set rules would not be feasible. Therefore, the BSF-R was developed in consultation with a panel of experts in educational and developmental assessment. (Please see section 3.1.1 for a comparison of the BSF-R and BSID-II.) IRT analyses were used to assist the development of a BSF-R that is as comparable as possible with the BSID-II. In addition to the IRT analyses, items were also reviewed for their feasibility of administration in a field setting. Ideally, an item selected for inclusion in the BSF-R would have strong psychometric properties (see below), would be easy to administer, would require minimal administration materials, and would be as objective as possible to score. This section summarizes the steps that were taken to select the items for the BSF-R, describes the IRT analyses that were conducted, and presents the reliabilities of the mental scale and motor scale.

In order to select the items for the BSF-R, the standardization data set for the BSID-II was obtained from the Psychological Corporation, publisher of the BSID-II. Item calibrations were conducted to determine the IRT difficulty and discrimination parameters for all items in the BSID-II. The difficulty parameter locates each item on a continuum of ability extending from 0 to 42 months of age. In essence, all items are now arranged in developmental order. The discrimination parameter reveals how useful an item is at distinguishing between children who are above or below the level of difficulty represented by the item. All items in the appropriate difficulty range for the ECLS-B target population (i.e., 8–11 months) were reviewed, and a pool of candidate items was identified. The discrimination parameter was used to eliminate those items that were weak discriminators of ability. In general, the goal was to try to keep those items with a discrimination parameter of 0.7 or greater.

Once a list of items in the appropriate ability range with good discrimination was created, items were reviewed again for content and feasibility of administration in the field by a national field staff. In particular, items were selected for their ease of administration, clarity of scoring criteria, number of materials, and appropriate representation of content. Additionally, to prioritize content coverage, factor analyses were conducted by Westat analysts using the standardization data set of the BSID-II from the publisher. Similar to earlier studies (see Burns, Burns, and Kabacoff, 1992) with the BSID (first edition, Bayley, 1969), it was found that the mental scale demonstrates one main factor—ability. Although item content and breadth were considered, appropriate psychometric properties and ease of administration and scoring were the key criteria used in item selection. The BSF-R instrument was field tested, and IRT analyses were again conducted to review the psychometric properties of the selected items. Revisions were made to the BSF-R, replacing items not performing satisfactorily or reformatting administration and scoring procedures for greater clarity. See *ECLS-B Methodology Report for the Nine-Month Data Collection, Volume 1: Psychometric Characteristics* (U.S. Department of Education, National Center for Education Statistics, forthcoming [b]) for more details on the creation of the BSF-R.

The final version of the BSF-R used in the ECLS-B 9-month national data collection has also been analyzed for its psychometric properties. Initially, the complete sets of 178 mental and 111 motor items from the BSID-II were calibrated using an IRT two-parameter model (i.e., ability and discrimination parameters). The 9-month BSF-R mental and motor scales were then calibrated with an identical IRT model using the item scores from the national data collection, and true-score equating was then carried out. Based on these analyses, several items did not equate with publisher items as expected. These items were not included in the true-score equating of the BSF-R, but were still used in scoring. Four mental items and four motor items were disregarded during equating but were subsequently used in scoring.

The overall success of the BSF-R when used with the ECLS-B population is reflected in the reliability coefficient, an indicator of how well the items measure a construct consistently. The IRT reliability coefficient obtained for the 9-month BSF-R mental scale in the ECLS-B is $r_{xx} = 0.80$. The IRT reliability coefficient obtained for the BSF-R motor scale is $r_{xx} = 0.93$. These coefficients represent the reliability of the BSF-R mental scale and motor scale for the target population of the ECLS-B as a whole. The corresponding reliability coefficient for the same set of mental items in the BSID-II is $r_{xx} = 0.87$. The reliability coefficient for the same set of motor items in the BSID-II is also $r_{xx} = 0.87$. The BSF-R was created to measure the developmental status of children for an age range of 8–11 months. Some children in the ECLS-B were outside this age range when the BSF-R was administered. See *ECLS-B Methodology Report for the Nine-Month Data Collection, Volume 1: Psychometric Characteristics* (U.S. Department of Education, National Center for Education Statistics, forthcoming [b]) for more information on the reliabilities by age.

3.1.2 Nursing Child Assessment Teaching Scale

The Nursing Child Assessment Teaching Scale (NCATS)[5] is an observational coding system for rating caregiver-child interaction patterns that is widely used in clinical practice and research to screen mothers and infants for early intervention programs as well as in maternal and child health research (Sumner and Spietz, 1994). The individual NCATS items are in the file and are clearly identified with the NCATS item number and a brief label that paraphrases the content of the item; e.g., for variable C1CHRCH, the variable description is "C1 NCATS 2 CG PUT CHILD REACH MATERIAL." For the exact wording of all the items, please refer to the NCATS manual (Sumner and Spietz, 1994). To obtain further information about the NCATS, or to place an order for a manual, the analyst is referred to the Nursing Child Assessment Satellite Training (NCAST) web site: http://www.NCAST.org.

The NCATS scale is composed of 73 binary (yes/no) items that are scored by trained observers, usually from direct observation, but also from videotaped interactions. In the ECLS-B, the interviewer videotaped the interaction during the home visit. (Approximately 90 percent of cases required only a single trial to complete the videotape.) When the videotape was returned to Westat, a trained coder viewed the videotaped interaction and scored each item as a "Yes" if the behavior was observed and a "No" if the behavior was not observed.

The NCATS coding system has two main scales: the parent scale, which has 50 items, and the child scale, which has 23 items. The items in the parent scale are grouped into four conceptual categories for ease of scoring: Sensitivity to Cues, Response to Child's Distress, Cognitive Growth Fostering, and Socioemotional Growth Fostering. The child items are grouped into two categories: Clarity of Cues and Responsiveness to Caregiver. In addition, the items measuring contingency, that is, items that have both caregiver and infant components, are singled out and appear in both the subscales and two subscales of their own, the Parent Contingency Scale and the Child Contingency Scale.

[5] The NCATS is used with permission from its developer, Dr. Kathryn Barnard (agreement dated January 28, 2003) who granted NCES permission to include 9-month NCATS item-level scores, the Total Scale Score, Total Parent Scale Score, and the Total Child Scale Score in the public-use files of the Early Childhood Longitudinal Study, Birth Cohort.

The NCATS manual presents alpha reliability coefficients for each of these groups and researchers often use these groups as subscales. However, in the survey setting of the ECLS-B, there was little psychometric support for creating these six subscales as many of the alphas are quite low. [Please see *ECLS-B Methodology Report for the Nine-Month Data Collection, Volume 1: Psychometric Characteristics* (U.S. Department of Education, National Center for Education Statistics, forthcoming [b]) for a detailed discussion of the NCATS subscale reliabilities.] The total parent scale, the total child scale, and the overall total NCATS scale (parent items plus child items), nevertheless, all have high internal consistency as measured by alpha. Therefore, on the ECLS-B data file there are item level data, the total parent scale score (sum of the "Yes" responses to the 50 parent items), the total child scale score (sum of the "Yes" responses to the 23 child items), and the total NCATS scale score (sum of the "Yes" responses to all 73 items).

According to the NCATS manual (Sumner and Spietz, 1994), the NCATS can be used to screen mothers and children who can be considered "at risk" (described in the NCATS manual as "worrisome"). The suggested criterion is that mothers, children, or mother-child pairs scoring in the lowest 10 percent of NCATS scores should be targeted for early intervention programs. The manual further suggests that the 10 percent criterion should also be based on the ethnic grouping of the sample, yielding separate scores for the major subgroups. Analysts interested in screening at-risk mothers, children, and mother-child pairs are referred to the NCATS manual for a table that presents suggested cut-off scores to identify at-risk dyads (Sumner and Spietz, 1994, p. 111). However, as described in the manual, these values refer only to children of normal birthweight and to mothers at least 20 years of age with, at minimum, a high school diploma. No information is provided in the manual for cases that do not meet these criteria. Table 3-3 provides the variable names and descriptions of the three NCATS total scale scores that are included in the ECLS-B data file, along with their distributions. Each of these scale scores is described below along with its appropriate use.

Total Parent Scale Score. The focus of the parent scale of the NCATS is on the caregiver's use of a "teaching loop," which consists of four components: (1) alerting the child by getting the child's attention, and setting up the expectations for what is about to be done; (2) instruction, during which the mother teaches the child by asking questions, making suggestions, giving directives, and so forth; (3) performance, which involves letting the child respond to the caregiver's suggestions or teaching; and (4) feedback, in which the mother comments on or responds to the child's attempts to do the task (Sumner and Spietz, 1994). The total parent scale score is useful for descriptive purposes to summarize the parent's teaching behaviors during the NCATS interaction. A low score would indicate that the parent did not engage in many teaching-oriented behaviors, and a high score would indicate that the parent was more likely to engage in a broader range of teaching behaviors that include use of the teaching loop.

The total parent scale score is the sum of the “Yes” responses for the 50 parent items. All items in this scale are unidirectional. That is, a “Yes” indicates the presence of a desirable behavioral characteristic (or absence of an undesirable behavior) and a “No” indicates the absence of a positive behavior. For example, if the item “Caregiver positions child so child is safely supported” is checked as “Yes,” it means that the caregiver has taken steps to make sure the child is in a safe position; when that item is checked as “No,” it means that the caregiver has not taken any steps to keep the child safe. Likewise, if the item “Caregiver avoids yelling at the child” is checked as “Yes,” it indicates the absence of a negative behavior; if it is checked as “No,” it means that the caregiver yelled at the child.

Table 3-3. NCATS scores in the ECLS-B 9-month data file: 2001–02

Variable name	Variable label (description)	Range of values	Weighted mean[1]	Standard deviation
X1NCATTP	X1 NCATS – Total Parent Score	15.00-49.00	34.68	4.51
X1NCATTC	X1 NCATS – Total Child Score	4.00-23.00	15.50	2.70
X1NCATTS	X1 NCATS – Total Score	23.00-70.00	50.18	5.77

[1] The full sample respondent-child weight W1C0 was used to obtain the weighted mean.
SOURCE: U.S. Department of Education, National Center for Education Statistics, Early Childhood Longitudinal Study, Birth Cohort, Nine-Month Data Collection, 2001–02.

Individual items are coded as either “Yes” for presence of a positive behavior (or absence of a negative behavior) or “No” (for absence of the behavior in question). These are scored on the file as “Yes” = 1, and “No” = 2. All total scale scores for the NCATS, the total NCATS scale score, the total parent scale score, and the total child scale score are all the sum of the “yeses.” Therefore, potentially the total parent score can range from 0 to 50, although the actual range is from 15 through 49.

It is possible to have missing data for the individual items in the Parent Scale. For example, if the video camera was too far away from the dyad, the audio level may be too low to hear the content of the speech but adequate enough to detect changes in tone and inflection. In such a case, it would be possible to code the items focused on warmth of tone, intonation, and inflection but not possible to code the items that assess the cognitive content of the caregiver’s speech. If there are missing data that do not exceed 25 percent of the items, then the score is prorated based on the number of items scored and converted into the 0–50 range using the following simple formula:

Formula for Prorating NCATS Scale Scores:

$$\frac{X}{\text{Total number of items}} = \frac{\text{Number of “Yes” responses}}{\text{Number of items coded}}$$

For example, if 10 parent items were missing for any reason, then the score is based on the 40 scored items but converted into the 0-50 range. That is, if the case had 35 “Yes” responses out of 40 scored items, the total parent score would be 43.75 (35/40 x 50). However, if 16 or more parent items were not scored (more than 25 percent missing data), then the score is determined to be missing (-9). This occurred in less than 0.5 percent of the cases.

The total parent scale score would be the variable of choice for analysts who are interested in the associations between, say, key demographic variables such as family socioeconomic status, caregiver education, maternal age at child’s birth, caregiver parenting attitudes, and knowledge of infant development with observed caregiver teaching behaviors on the NCATS. As the data from later waves become available, the analyst might consider examining the longitudinal associations between the total parent scale score and such child outcome measures as language ability and adjustment to school.

Total Child Scale Score. This is the sum of the “Yes” responses for the 23 child items. To participate in a mutual relationship with the caregiver, the child must send clear cues to the caregiver and be responsive to the caregiver. If the child is not responsive to the behavioral cues of the caregiver, and the caregiver is not responsive to the cues of the child, then mutual adaptation will not be possible. The total child scale score is useful for descriptive purposes to summarize the child’s ability to communicate clear cues to the parent (e.g., “Child grimaces or frowns during the teaching episode”) and to respond in turn to the caregiver (e.g., “Child vocalizes or babbles within 5 seconds after caregiver’s verbalization”). A low score would indicate that the child did not communicate clear cues and did not respond easily to the caregiver’s attempts to engage the child. A high score would indicate that the child was communicating clearly with the caregiver and responded adaptively to the caregiver’s cues.

Again, individual items in the child scale are scored as either “1” (“Yes”) or “2” (“No”). The total child scale score is the sum of the “yeses.” Therefore, the potential range for the total child scale score is from 0 to 23. As with the parent scale, all items are unidirectional. However, this unidirectionality is subtle in the case of the child. One subset of items assesses the clarity of the child's cues. The cues themselves may be either positive or negative. At issue is whether those cues were clear or not. A child who communicates clear cues would receive a higher score on this set of items than a child who communicates ambiguous cues. This is also true of the child items that are typically grouped in the responsiveness to caregiver subset of items. A child who responds to the caregiver clearly, either positively or negatively, would receive a higher score on this set of items than a child who responds ambiguously.

If there are missing data that do not exceed 25 percent of the items, then an adjusted score is obtained based on the number of items scored, using the same formula shown earlier. For example, if 3 child items were not scored, then the total child scale score is based on the 20 scored items. That is, if the case had 17 "Yes" responses out of 20 scored items, the total child scale score would be 19.55 (17 divided by 20 and then that ratio multiplied by 23). If 4 or more child items were not scored (i.e., more than 25 percent of child items were not scored), then the total child scale score is determined to be missing (−9). This occurred in less than 0.5 percent of the cases.

The total child scale score would be useful when investigating the relationships between child characteristics (e.g., fussiness, excitability, irritability, health status, or developmental status) and children's clarity of communication to the caregiver and his or her responsiveness to the caregiver. When the data from subsequent rounds become available, the total child scale score would be useful when studying children's later development and adjustment to preschool.

Total NCATS Scale Score. This is the sum of the "Yes" responses for the entire set of items in the NCATS (i.e., 73 items), including both the parent items and the child items. The Total NCATS Scale score is the same as the sum of the parent and child scale scores, provided that at least 75% each of the parent items and of the child items are scored. It is useful for descriptive purposes to summarize the caregiver's use of the teaching loop and the child's responsiveness to the caregiver. A low score would indicate that a dyad (i.e., parent-child) did not engage in many behaviors associated with the teaching loop, whereas a high score would indicate increased use of the teaching loop by both the mother and child. It would not be considered solely a frequency count of teaching loop behaviors, because the NCATS is not a frequency-count measure, per se. Rather, it combines elements of quality of interaction and breadth of behaviors. The total NCATS scale score could be considered a count of the different types of behaviors that make up the teaching loop. For example, a score of 30 would describe an interaction that likely includes only some of the behaviors associated with the teaching loop, whereas a high score—60, for example—would indicate that both caregiver and child participated to a greater extent in the teaching loop (Sumner and Spietz, 1994).

The analyst might choose the total NCATS scale score as the dependent variable to investigate the associations between, for example, maternal education, parenting attitudes, or knowledge of infant development and use of the teaching loop during the NCATS. The total NCATS scale score would also be useful as the independent variable in analyses that examine use of the teaching loop with various child outcomes, both concurrent and longitudinally when future data become available. Analysts interested in more specific aspects, for example, children's responsiveness within the teaching loop, may optionally choose to recreate the subscales or explore alternative factor structures. Analysts interested in recreating the subscales typically used by researchers are cautioned about the low alphas for some of the subscales in the ECLS-B, as discussed above, and are referred to the NCATS manual (Sumner and Spietz, 1994), and to the NCAST web site presented earlier.

Analysts who are interested in the associations between more specific aspects of the NCATS subscales and various outcome measures may refer to the NCATS manual and use the item-level data to recreate the subscales or to conduct factor analyses to determine whether an alternative factor structure does a better job of identifying interpretable subscales. The analyst is referred to *ECLS-B Methodology Report for the Nine-Month Data Collection, Volume 1: Psychometric Characteristics* (U.S. Department of Education, National Center for Education Statistics, forthcoming [b]) for further information about the NCATS items and subscales. For this reason, individually scored item-level data (i.e., the 73 individual items scored as "Yes" or "No" are included on the data file). These item-level data are provided so that analysts may optionally recreate the subscales or conduct factor analyses to investigate alternative factor structures.

3.1.2.1 NCATS Supplemental Contextual Variables

The data file also includes several additional NCATS-related variables that may be of interest to analysts. These variables include the length of time in minutes (MM) and seconds (SS) of the NCATS interaction, the presence of others during the interaction, the child's state of alertness, whether or not the child displayed any potent disengagement cues during the NCATS (e.g., "tray pounding," back arching, crying), and the caregiver's language to the child during the NCATS.

These supplemental variables are presented in table 3-4, which lists the variable name, its variable description, and its range of values.

Table 3-4. NCATS supplemental contextual variables in the ECLS-B 9-month data file: 2001–02

Variable name	Variable label (description)	Range of values
C1LENGTH	C1 NCATS LENGTH OF TIME TEACHING (MM:SS)	0:45-36:00[1] mm:ss
C1OTHERS	C1 NCATS OTHERS PRESENCE	Yes (1), No (2)
C1CHSTTE	C1 NCATS CHILD'S STATE AT TEACHNG START	Quiet Sleep (1) Active Sleep (2) Drowsy (3) Quiet Alert (4) Active Alert (5) Crying (6)
C1PDC	C1 NCATS POTENT DISENGGMNT CUES OBSERVED	Yes (1), No (2)
X1LNGNCT	X1 LANGUAGE IN WHICH NCATS CONDUCTED	English (1) Spanish (2) Chinese (3) Other (4)

[1] *Tapes that continued past the 5 mm and 30 ss required for reliable NCATS scoring were only coded for the first 5 mm and 30 ss of interaction.*
SOURCE: U.S. Department of Education, National Center for Education Statistics, Early Childhood Longitudinal Study, Birth Cohort, Nine-Month Data Collection, 2001–02.

3.1.2.2 Reliability of NCATS Coding

To ensure the quality of the ECLS-B data, the reliability of the NCATS coding was established. The reliability of a scale refers, on one hand, to the stability of scores across different coders, and, on the other hand, to the ability of the items to measure a construct consistently. Reliability was established in two ways for the ECLS-B. First of all, a total of 171 videotaped ECLS-B NCATS interactions (the "reliability tapes") (representing 1.8 percent of the total of 9,437 codeable NCATS tapes), selected at random, were sent to NCAST at the University of Washington to be coded by a certified coder.

Coding of ECLS-B videotapes by NCAST staff was done to establish inter-lab reliability, the standard against which the ECLS-B coders were measured. ECLS-B coders were required to complete one reliability tape per week to establish coder percent agreement with the University of Washington.

(Please see section 5.2.4 for a summary of the NCATS training procedure and requirements for coding certification, as well as for a discussion of the reliability procedure.) Overall, coders averaged 86 percent agreement with the certified NCAST coder at the University of Washington, which exceeded the target of 85 percent agreement between ECLS-B coders and NCAST.

Secondly, reliability coefficients were separately obtained for all ECLS-B NCATS tapes, including the subset of 173 videotapes that were coded at the University of Washington. This was done to establish the internal consistency of the subscales and to supplement the percent agreement inter-lab reliabilities that were obtained on a weekly basis. If Westat coders were in agreement with the coder at the University of Washington, then the alphas would be similar.

If the items of a single subscale are equivalent, then those items should be correlated with each other. Cronbach's alpha, which is based on the split-half reliability correlation coefficients, was obtained for the videotapes that were coded at the University of Washington and for the ECLS-B videotapes that were coded at Westat. Cronbach's alphas for the NCATS Total Parent Scale, Total Child Scale, and Total NCATS Scale are presented in table 3-5 for the total ECLS-B codeable tapes, for the ECLS-B reliability tapes coded at the University of Washington, and, for comparison, for the publisher data set in the NCATS manual.

Table 3-5. Cronbach's alphas for Total Parent, Total Child, and Total NCATS Scales: 2001–02

NCATS scales	9-month ECLS-B NCATS tapes (N = 8520)	U. of Washington reliability tapes (n=171)	U. of Washington (N=1,887, publisher data set)
Total Parent Scale	.68	.74	.87
Total Child Scale	.62	.63	.80
Total NCATS Scale	.72	.77	.87

SOURCE: U.S. Department of Education, National Center for Education Statistics, Early Childhood Longitudinal Study, Birth Cohort, Nine-Month Data Collection, 2001–02.

3.1.3 The Child Observations During the Bayley Short Form–Research Edition

The Interviewer Observations of Child Behavior during the Bayley Short Form–Research Edition (Child Observations) consist of a short set of items selected from the Behavior Rating Scale (BRS), a supplementary component of the BSID-II. Only 9 out of the 30 BRS items are included in the ECLS-B. These items were selected to be representative of different domains of behavior. During the national training, interviewers received extensive training on observing these behaviors using videotapes that included examples of the full range of behaviors in each item. Please see the *Early Childhood Longitudinal Study, Birth Cohort (ECLS-B) Design and Operations Report for the Nine-Month Data Collection* (U.S. Department of Education, National Center for Education Statistics, forthcoming[a]).

The analyst is advised that the items were not intended to be scaled or factor analyzed. Nor are the psychometric properties of the BRS being summarized here because this subset should not really be considered the “BRS.” Seven items were observed by the interviewer during the administration of the BSF-R. Two items are parent reports of the child’s behavior and performance on the BSF-R. These nine items provide discrete information about children’s interest, engagement, and behavior during the BSF-R that may help the analyst interpret individual mental and motor scale scores.

Table 3-6 presents the variable names and descriptions of the set of Child Observations items in the ECLS-B.

Table 3-6. Child Observations items in the data file: 2001–02

Variable name	Variable label (description)	Range of values
Two Interviewer Observation questions asked of parent respondent:[1]		
C1BBHAV1	C1BSF CG RATE BEHAVIOR TYPICAL VISIT 1 (How typical was your child's behavior? Did {CHILD} play the way {he/she} usually does? Was {he/she} as happy or upset as usual? As alert and active as usual?)	Very atypical (1); Mostly atypical (2); Somewhat typical (3); Typical (4); Very typical (5).
C1BBHAV2	C1BSF CG RATE BEHAVIOR TYPICAL VISIT 2 (How typical was your child's behavior? Did {CHILD} play the way {he/she} usually does? Was {he/she} as happy or upset as usual? As alert and active as usual?)	Very atypical (1); Mostly atypical (2); Somewhat typical (3); Typical (4); Very typical (5).
C1PRFM1	C1 BSF CG RATE PERFORMANCE VISIT 1 (Do you think {CHILD} did as well as {he/she} could? Have you seen {CHILD} do better or worse on the type of things we worked on?)	Poor indicator (1); Barely adequate (2); Adequate (3); Good (4); Excellent (5).
C1PRFM2	C1 BSF CG RATE PERFORMANCE VISIT 2 (Do you think {CHILD} did as well as {he/she} could? Have you seen {CHILD} do better or worse on the type of things we worked on?)	Poor indicator (1); Barely adequate (2); Adequate (3); Good (4); Excellent (5).

See notes at end of table.

Table 3-6. Child Observations items in the data file: 2001–02—Continued

Variable name	Variable label (description)	Range of values
Subset of BRS Questions Completed by Interviewer:[1]		
R1POSAFF	R1 CO035 CHILD DISPLAYS POSITIVE AFFECT	No positive affect displayed (1); 1 or 2 brief displays of positive affect (2); 3 or more brief displays of positive affect (3); 1 or 2 intense, heightened, or prolonged displays of positive affect (4); 3 or more intense, heightened, or prolonged displays of positive affect (5).
R1NEGAFF	R1 CO040 CHILD DISPLAYS NEGATIVE AFFECT	3 or more intense, heightened, or prolonged displays of negative affect (1); 1 or 2 intense, heightened, or prolonged displays of negative affect (2); 3 or more brief displays of negative affect (3); 1 or 2 brief displays of negative affect (4); No negative affect displayed (5).
R1ADAPT	R1 CO050 CHILD ADAPT CHANGE IN MATERIAL	Consistently resists relinquishing materials and/or refused to accept new materials (1); Typically resists relinquishing materials and/or refused to accept new materials, makes 1 or 2 transitions easily (2); Makes poor transitions half the time, makes good transitions half the time (3); Typically relinquishes materials and accepts new materials, 1 or 2 poor transitions (4); Consistently relinquishes materials and accepts new materials (5).
R1INTRST	R1 CO065 CH SHOWS INTEREST IN MATERIAL	No interest (1); 1 or 2 displays of interest (2); Moderate interest (3); Much interest (4); Constant interest (5).

See notes at end of table.

Table 3-6. Child Observations items in the data file: 2001–02—Continued

Variable name	Variable label (description)	Range of values
Subset of BRS Questions Completed by Interviewer[1]—Continued		
R1ATNTSK	R1 CO080 CHILD PAYS ATTENTION TO TASKS	Constantly off task, does not attend (1); Typically off task, attends in 1 or 2 instances (2); Off task half the time (3); Typically attends, attention wanders in 1 or 2 instances (4); Constantly attends (5).
R1SOCIAL	R1 CO110 CH DISPLAYS SOCIAL ENGAGEMENT	No attempts to interact socially (1); 1 or 2 attempts to interact socially (2); Several attempts to interact socially (3); Many attempts to interact socially (4); Constant attempts to interact socially (5).
R1CNTLMV	R1 CO130 CH SHOWS CONTROL OF MOVEMENTS	Consistently jerky or clumsy (1); Typically jerky or clumsy (2); Jerky or clumsy half the time, smooth or coordinated half the time (3); Typically smooth or coordinated (4); Consistently smooth or coordinated (5).

[1] Interviewer Observation materials developed for the Early Childhood Longitudinal Study, Birth Cohort (ECLS-B). Adapted from the *Bayley Scales of Infant Development: Second Edition, Behavior Rating Scale*. Copyright © 1993 by The Psychological Corporation, a Harcourt Assessment Company. Computer adapted and reproduced by permission of the publisher. All rights reserved.
SOURCE: U.S. Department of Education, National Center for Education Statistics, Early Childhood Longitudinal Study, Birth Cohort, Nine-Month Data Collection, 2001–02.

3.1.4 Physical Measurements

The ECLS-B measures children’s weight, length, and middle upper arm circumference (MUAC). For children born with very low birth weight (<1,500 grams or <3 pounds, 4 ounces), head circumference is also measured. Standard procedures were followed to obtain the physical measurements to ensure that all measurements were obtained in the same way for all children. To obtain child weight, the mother first stood on the scale and her weight was recorded. The interviewer then handed the child to the mother and the scale’s tare feature (in which the child’s weight is automatically subtracted from the combined weight of the mother and child) then displayed the child’s weight alone.

Occasionally, another individual other than the mother held the child as the child was weighed, for example, if the mother was not available or refused to be weighed. In most cases, this “stand-in” was the interviewer, although other adults may have been involved, such as the father, or a grandmother. Whenever this happened, the interviewer recorded on the record sheet that someone other than the mother was tbe adult weighed. If anyone other than the mother (defined as the household mother) was weighed, then X1HMWGHT = −1. If the weight of the mother was obtained, then the value of X1HMWGHT will be the average of the mother’s two weights (i.e., C1ADWGT1 and C2ADWGT2) or will be the value of C1ADWGT1 if the second weight is missing, or if someone other than the mother was weighed for the second obtained weight.

In addition, there are approximately 90 cases that have child weight (C1CHWGT1 and C1CHWGT2) information that do not have adult weight information (C1ADWGT1 and C1ADWGT2). It is possible in such cases that the child weight was obtained directly by placing the child on the scale. The scale all interviewers used is sensitive enough to obtain accurate weights to within 0.2 kilogram for weights under approximately 13 kilograms. Therefore, weights obtained from the child alone can be considered accurate. In addition, in one case where the scale malfunctioned, the child weight was recorded by parent report of a well-baby check at a medical facility the day before. The values for this case were retained as reported.

As is standard practice in all major health studies, each of these physical measurements is obtained twice. Therefore, for each measurement there are three variables on the data file, the first measurement, the second measurement, and the average of the two. For further information about procedures used to average the values, please see section 7.5.1.5.

In those cases where only one measurement was obtained (instead of two), the average is the single value that was obtained. In addition, some physical measurements for some cases may have been adjusted at the home office at the time of data entry to correct for errors in data recording made in the field. Adjustments to physical measurement data occurred in approximately 2 percent of cases, mostly early in data collection. Common reasons for adjustments were misalignment of the decimal point with the space provided on the record form, which was corrected by moving the decimal point; not resetting the scale’s tare feature such that the child’s weight appeared to be greater than the caregiver’s weight, which was corrected by subtracting the caregiver’s weight; recording the child’s length in inches instead of centimeters, both of which were available on the measuring device (resulting in a value that was far out of range, e.g., 25 instead of 63), which was corrected by converting to centimeters; and recording head circumference that was well below range, possible if one read from the wrong side of the SECA tape, which was corrected by reading the value on the correct side of the tape.

When the first cases were received at the home office, physical measurements with any problems (e.g., a mispositioned decimal point) were held aside for review by Child Development staff members, who made appropriate corrections (e.g., correctly positioned the decimal point). A flag, C1ADJSTD, was created to indicate whether or not it was necessary to adjust a physical measurement for any of the above reasons. It soon became apparent that a single flag indicating that a physical measurement had been adjusted was not sufficient, as multiple measurements may have been adjusted. Therefore, additional flags that are specific to the physical measurements were subsequently created. C1ADJSTD indicates to the analyst that one of the physical measurements was adjusted; the analyst should then review the following additional variables to identify what physical measurement was adjusted: C1PMXAW1 and C1PMXAW2; C1PMXCW1 and C1PMXCW2; C1PMXUL1 and C1PMXUL2; C1PMXUP1 and C1PMXUP2; C1PMXUC1 and C1PMXUC2; P1PMXLG1 and P1PMXLG2; and C1PMXHC1 and C1PMXHC2. Each of these variables will have a value of 1 if an adjustment was made. A value of 2 for any of these variables indicates that this variable was not adjusted because an adjustment was not necessary for this variable. It should be noted that C1ADJSTD was created at the beginning of data entry, as the physical adjustments were first being made. However, the flags for the specific physical measurements were added shortly thereafter so that specific information about adjusted values may be missing for a very small number of cases.

In using the MUAC data, analysts should note the occurrence of “out-of-range” data with the upper arm length and midpoint measurements (i.e., C1MUAM1, C1MUAM2, C1MUAL1, and C1MUAL2). This may have occurred due to miscalculation of the midpoint location or errors in measurement by using the wrong side of the SECA tape. The concern here is that if the midpoint is not located correctly, the MUAC measurement may also be incorrect. This occurs in less than 2 percent of ECLS-B cases. These data were retained and, if they were consistent with the criteria (see below for a discussion of how these composites were created), were used in creating the composite average for the child’s MUAC (X1CHMUAC). In cases where the MUAC value is 6.5 cm or less, the analyst may want to consider not using these cases in analysis, using the composite MUAC (X1CHMUAC), or using only the measurement that does not fall out of range.

The reliabilities of the physical measurements were quite high, as measured by the correlation between the first recording and the second recording of each set of measurements. For the child’s weight, the correlation between the first and second recordings was $r = .81$; for the child’s length, this correlation was r = .81; for the MUAC, the correlation was .85; and for head circumference, it was r = .84.

Table 3-7 presents the physical measurements variable names, range, and mean and standard deviation for the sample as a whole.

Table 3-7. Physical measurement composites in the ECLS-B 9-month data file: 2001–02

Variable name	Variable label (description)	Range of values	Weighted mean[1]	Standard deviation
X1CHMUAC	X1 PM – CHILD'S MIDDLE UPPER ARM CIRCUMFERENCE (CM)	9.00-23.00 cm	15.87 cm	1.96
X1CHLENG	X1 PM – CHILD'S LENGTH (CM)	47.30-95.00 cm	73.10 cm	4.00
X1CHWGHT	X1 PM – CHILD'S WEIGHT (KG)	2.00-20.00 kg	9.57 kg	1.67
X1CHCRFM	X1 PM – CHILD'S HEAD CIRCUMFERENCE (CM)	35.00-52.00 cm	44.16 cm	2.69

[1]The full sample respondent-child weight W1C0 was used to obtain the weighted mean.
SOURCE: U.S. Department of Education, National Center for Education Statistics, Early Childhood Longitudinal Study, Birth Cohort, Nine-Month Data Collection, 2001–02.

The Centers for Disease Control and Prevention (CDC) maintains a web site that provides information about the norms for children's physical measurements. To compare the children in the ECLS-B, separately by gender and by age, with the CDC norms, the analyst is referred to that web site: www.cdc.gov/nchs/about/major/nhanes/growthcharts/clinical_charts.htm#Clin%201. However, it should be noted that the size of the ECLS-B sample is much larger than the size of the sample for the CDC norms. It is possible that the ECLS-B will obtain measurements that will differ in some cases from the CDC norms.

3.2 Indirect Child Assessments

The ECLS-B uses two approaches to assessment—direct and indirect. Direct assessment includes those measures that involve the child's participation, and were described above. The child's performance on these direct assessments provides information about his or her current developmental level. For example, the motor items of the BSF-R measure whether the child is able to walk or grasp objects at the time the BSF-R was administered. In essence, the BSF-R obtains a snapshot of the child's ability at a single point in time.

Generally, indirect assessment is used for areas that are difficult to assess directly or observe directly in young children. Indirect assessment can be characterized as obtaining information about the more enduring abilities, traits, and characteristics of the child over time from a person (or persons) who knows the child well, such as the primary caregiver. At 9 months, the ECLS-B indirect assessments consisted of rating scales that evaluated the child's current behavior and skills and were completed by the child's parents or guardians. The indirect assessments asked the parent to provide information about what the child was generally like, on a daily basis, in a wide range of circumstances and different conditions. In addition, the questions to the parent about the child's acquisition of various developmental milestones addressed the issue of whether the early acquisition of developmental milestones is particularly predictive of developmental status, in general. By comparing and contrasting the information obtained in the direct and indirect assessments, a fuller picture of children's developmental paths results.

Two indirect assessments of children's development were included in the ECLS-B 9-month data collection. Information was obtained from the primary caregiver during the parent interview via

- A subset of questions that obtain information about the age at which children achieved selected developmental milestones and
- A subset of questions that obtain information about difficult child behaviors that could be considered symptomatic of a potential self-regulatory problem.

The set of questions about developmental milestones was derived from the Infant Development Inventory, an age-specific application of the Child Development Inventory. The Child Development Inventory was developed at the University of Minnesota and is generally known as the Minnesota-CDI, or MN-CDI. Although the subset of questions included in the ECLS-B is from the infant version of that inventory, it is still referred to as the MN-CDI, its most common referent. The subset of questions about self-regulatory problem behaviors was selected from the Infant/Toddler Symptom Checklist (ITSC).

3.2.1 Selected Items from the Child Development Inventory (MN-CDI)

The Child Development Inventory (Ireton, 1997) was designed as a parent-report measure to screen young children whose development may not be optimal. Its appropriate age range is from birth through 6 years of age. The age-appropriate infant version, the Infant Development Inventory, measures development in five areas: social, self-help, gross motor, fine motor, and language. According to the standard administration, the parent is asked to read a series of statements that describe young children's behaviors. These statements are ordered according to the ages at which the behaviors are achieved, on average. The parent is then asked to respond "Yes" or "No" to each statement and to report whether he or she has seen the child behave in this way. A "Yes" indicates that the statement describes the child's present behavior or describes behavior that the child used to do but has outgrown. A "No" indicates that the statement does not describe the child's behavior or that it is a behavior the child is just beginning to do or only does sometimes. However, for the ECLS-B parent CAPI instrument, the emphasis of the items was switched from whether the child could do the behavior to when the child first became able to do it. This required that the MN-CDI items be modified to specify the age at which the child acquired an ability. Thus estimates based on the items in the ECLS-B may not be directly comparable with estimates based on the original MN-CDI. For the gross motor skills (sit alone, crawl, pull self to standing position, and walk while holding on to something), if a child was not yet able to perform the skill, the interviewer, on the basis of observations during the BSF-R, entered a value of "95" to indicate that the child did not yet have the skill.

The MN-CDI was used to identify key developmental milestones that should be included in the ECLS-B parent CAPI instrument. In addition, because the information obtained was retrospective, the salience of the milestone was an important consideration. The milestones were selected because they were clearly observable by parents. Presumably, because the milestone was recognizable, the parent was more likely to remember the age at which the child achieved the milestone. Although the MN-CDI was used as the source for the behaviors, the response categories have changed so much that the items should be considered derived from the MN-CDI rather than actual MN-CDI items, per se.

In the ECLS-B the intention for the developmental milestone items was to provide a descriptive measure of children's acquisition of key developmental milestones that would characterize America's children. It was generally assumed that children who pass these milestones early are in some way more developmentally advantaged than children who pass these milestones later, and that this early advantage is sustained. Information about age of milestone acquisition obtained in the ECLS-B presumably could be compared with research in early childhood education and in child development to examine whether the early acquisition of milestones is associated with positive outcomes in the preschool and early school years. This is the first chance to examine these issues with a national data set.

The developmental milestones included in the ECLS-B parent CAPI instrument derived from the MN-CDI are summarized in table 3-8.

Table 3-8. Developmental milestone questions in the parent CAPI instrument: 2001–02

Variable name	Variable label (description)	Range of values	Weighted mean[1]	Standard deviation
P1AGFSLF	P1 CD075 AGE STARTED FEEDING SELF (MTHS)	1-14 mos.	8.33	2.05
P1AGSIT	P1 CD080 AGE SITTING WITHOUT SPPRT (MTHS)	2-14 mos.	6.07	1.34
P1AGCRWL	P1 CD090 AGE WHEN STARTED CRAWLING (MTHS)	1-14 mos.	7.16	1.46
P1AGSTND	P1 CD100 AGE PULLED SELF TO STAND (MTHS)	2-14 mos.	7.97	1.41
P1AGWALK	P1 CD110 AGE WALKING WITH HELP (MTHS)	3-14 mos.	8.57	1.39

[1] The full sample respondent-child weight W1C0 was used to obtain the weighted mean.
SOURCE: U.S. Department of Education, National Center for Education Statistics, Early Childhood Longitudinal Study, Birth Cohort, Nine-Month Data Collection, 2001–02.

3.2.2 Selected Items from the Infant/Toddler Symptom Checklist

The Infant/Toddler Symptom Checklist[6] (ITSC; DeGangi, Poisson, Sickel, and Wiener, 1995) is a screening instrument designed to be completed by nonprofessionals, including parents. Its purpose is to identify infants and toddlers with regulatory disorders who, therefore, have the potential to be demanding of their caregivers or unpredictably fussy, or have problems with sleep, feeding, or regulating their behavior and mood. Presumably, such problems may be manifestations of an underlying sensory-integrative disorder (e.g., balance problems, poor motor planning, hypersensitivity to touch), attentional deficit, or behavior problem that may emerge later, in the preschool years. It has been estimated, for example, that up to 70 percent of children with learning disabilities have an underlying sensory-integration disorder (Carte, Morrison, Sublett, Venura, and Setrakian, 1984).

Until recently, the assessment of behaviors related to self-regulation, oversensitivity to sensory stimulation, and attentional problems in young children has fallen under the domain of temperament scales. Temperament scales have as their focus such characteristics as fussiness, adaptability to distress, persistence, and impulsivity, but do not include the assessment of functional behaviors that may be related to behavioral, developmental, or learning difficulties. That is, temperament scales do not examine emotion regulation, self-calming, attention regulation, or self-regulation.

The ITSC, on the other hand, focuses on the infant's responses in the domains of self-regulation, attention, sleep, feeding, sensitivity to tactile stimulation (e.g., when being dressed, bathed, or touched), tolerance of movement, sensitivity to auditory stimulation and visual stimulation, and socioemotional functioning.

The ITSC is used from 7 to 30 months of age and has five age-appropriate versions (e.g., 7–9 months, 10–12 months). At 9 months, the full ITSC includes 19 items, which was too lengthy for the ECLS-B's purpose. Therefore, a subset of 8 items was selected for the ECLS-B parent CAPI instrument on the basis of the items' ability to identify children with regulatory disorders. The items selected cover the domains of self-regulation, sleep-wake regulation, and attending. The parent could respond *never*, *used to be*, *sometimes*, or *most times*. The ECLS-B rating for these items could, therefore, range from 0 to 3. One item, which asks the parent about how difficult the child was to raise, could range from 0 to 5. The items included in the ECLS-B are summarized in table 3-9.

Table 3-9. Infant/Toddler Symptom Checklist items in the ECLS-B 9-month data file: 2001–02

Variable name[1]	Variable label (description)	Range of values
P1FUSSY	P1 CD120A IS FUSSY OR IRRITABLE	Never (0)
		Used to be (1)
		Sometimes (2)
		Most times (3)
P1WHMPR	P1 CD120B GOES FROM WHIMPER TO CRYING	Never (0)
		Used to be (1)
		Sometimes (2)
		Most times (3)
P1ATTN	P1 CD120C DEMANDS ATTENTION AND COMPANY	Never (0)
		Used to be (1)
		Sometimes (2)
		Most times (3)
P1WAKES	P1 CD120D WAKES UP 3 OR MORE TIMES	Never (0)
		Used to be (1)
		Sometimes (2)
		Most times (3)
P1HLPSLP	P1 CD120E NEEDS HELP TO FALL ASLEEP	Never (0)
		Used to be (1)
		Sometimes (2)
		Most times (3)
P1STRTL	P1 CD120F STARTLED BY LOUD SOUNDS	Never (0)
		Used to be (1)
		Sometimes (2)
		Most times (3)
P1NOWAIT	P1 CD120G CRIES FOR FOOD OR TOYS	Never (0)
		Used to be (1)
		Sometimes (2)
		Most times (3)
P1RAISE	P1 CD125 DIFFICULTY TO RAISE ON AVERAGE	Not at all difficult (1)
		Not very difficult (2)
		About average (3)
		Somewhat difficult (4)
		Very difficult (5)

[1] Parent CAPI instrument items CD 120a through 120g (P1FUSSY, P1WHMPR, P1ATTN, P1WAKES, P1HLPSLP, P1STRTL, and P1NOWAIT) are from the *Infant/Toddler Symptom Checklist*. Copyright © 1995 by The Psychological Corporation. Reproduced by permission. All rights reserved.
"*Infant/Toddler Symptom Checklist*" is a trademark of The Psychological Corporation registered in the United States of America and/or other jurisdictions.
SOURCE: U.S. Department of Education, National Center for Education Statistics, Early Childhood Longitudinal Study, Birth Cohort, Nine-Month Data Collection, 2001–02.

The items selected for inclusion in the ECLS-B all successfully differentiated, on the basis of t-tests presented in the ITSC manual (DeGangi, et al., 1995, pp. 46-47), those children with regulatory disorders and those without such disorders. The results of the published t-tests are for the items included in the ECLS-B and are summarized in table 3-10. These data were obtained in a validity study, conducted by the authors of the measure, that recruited children through pediatric practices and involved two groups of children, one a group of normal children and the other a group of children clinically diagnosed with a regulatory disorder. The value of the t-test is in the second column from the right. In addition, Cronbach's alpha for the set of items was .63, which is adequate and suggests that the items form a scale. Analysts with an interest in the etiology of attention and other learning disorders may consider establishing cut-off scores to identify children at risk for such problems in later rounds.

Table 3-10. Differentiation of regulatory disordered children by items of the Infant/Toddler Symptom Checklist in the ECLS-B: 2001–02

Variable name	Description	Mean normal	Mean regulatory disorder	T-test	Degrees of freedom
P1FUSSY	Self-regulation: Child is frequently irritable, fussy	.34	1.28	- 5.06*	48
P1WHMPR	Self-regulation: Child goes easily from a whimper to an intense cry	.09	1.44	-8.16*	48
P1ATTN	Self-regulation: Child demands your attention and company constantly	.36	1.64	-4.84*	34
P1WAKES	Sleep: Child wakes up 3 or more times in the night and is unable to go back to sleep	.03	1.22	-8.67*	47
P1HLPSLP	Sleep: Child needs a lot of help to fall asleep	.41	1.00	-2.48*	48
P1STRTL	Listening: Child startles or is upset by loud sounds such as a vacuum, doorbell, or barking dog	.06	.39	-2.04*	48
P1NOWAIT	Self-regulation: Child is unable to wait for food or toys without crying or whining	.10	1.36	-5.27*	32

* $p < .05$

NOTE: These results are published in the manual for the ITSC (DeGangi et al., 1995), and are not from ECLS-B data. Mean normal represents the ITSC value for children not diagnosed with a regulatory disorder. Mean regulatory disorder refers to the ITSC value for children diagnosed with a regulatory disorder. Values range from 0 (Child is never like this) to 3 (Child is like this most times). The t-test statistic presented here is the significance of the ITSC item to differentiate children's behavior by regulatory disorder status.

SOURCE: DeGangi, G. A., Poisson, S., Sickel, R Z., & Weiner, A. S. (1995). Infant/Toddler Symptom Checklist: A Screening Tool for Parents. San Antonio, Texas: Therapy Skill Builders, a division of The Psychological Corporation.

3.3 Parent and Home Environment Measures

This section summarizes the constructs that are measured as part of the assessment of the child's home environment. In most cases, a subset of items was selected from an original, well-known instrument (see below) and has been included in the ECLS-B with no modifications. In some cases, however, the items were changed substantially so that they are no longer representative of the original instrument. For this reason, items are categorized as (a) identical to the original scale; (b) slightly modified from the original scale; and (c) no longer bearing much similarity to the original scale. The user should consider whether changes have been made to items and whether and how to create a scale from the items. In most cases, only a few items were selected from a much larger source and some of the items were modified from the original, so the ECLS-B does not claim to be using the scale. Below, two subsets of questions are discussed in detail because they are more complicated in their application or in their scoring—the Knowledge of Infant Development Inventory (KIDI) and the Home Observation for Measurement of the Environment (HOME).

3.3.1 Items from the Knowledge of Infant Development Inventory (KIDI)

The first subset to be described consists of 11 items selected from the Knowledge of Infant Development Inventory[7] (KIDI; MacPhee, 1981), which was designed to assess knowledge of parental practices, developmental processes, and infant norms of behavior. The items included were recommended by the measure's author as being the most successful for differentiating high versus low parenting knowledge. Each of the 75 items on the KIDI questionnaire describes either typical infant behaviors or aspects of parenting that could affect an infant's growth and behavior. Items include knowledge about infants' sleep needs, when children learn right from wrong, when infants typically say their first word, and so forth. The parent indicates where he or she "agrees," "disagrees," or "is not sure." Some of the items also ask the respondent to specify whether the behavior in question can be expected of an older or a younger child. The 11 items used in the ECLS-B were identical to the original KIDI.

Parents are asked to answer on the basis of their knowledge of infants in general and not with regard to their own children or how their children behaved at a given age. Because this is a measure

[7] The KIDI items are used with the permission of its developer, Dr. David MacPhee (agreement dated January 28, 2003) who recommended the subset of KIDI items for use in the Early Childhood Longitudinal Study, Birth Cohort.

of an individual's knowledge of child development, there are correct and incorrect answers. With the full KIDI, a score is obtained by summing the number of correct answers. In the ECLS-B, the individual items are presented with the parent response, not whether the response is correct or incorrect. These items were included in both the parent CAPI instrument (PA005-PA050) and in the resident father questionnaire (Q6 and Q7). In the parent CAPI instrument, the first four KIDI items are straightforward to score as the answer depends on only one question, for example, "All infants need the same amount of sleep, do you agree or disagree?" The rest of the KIDI items in the parent CAPI instrument are two-part questions with the correct score being dependent on the combination of the two parts. For example, "A 1-year-old knows right from wrong. Do you agree or disagree?" is followed by "Would a child be younger or older than 1 year when she first knows right from wrong? Or are you not sure?" The correct answer to the first question is "disagree," and the correct answer to the second question is "older." Together, these two questions would be credited with 1 point. In the 9-month resident father questionnaire, the 2-part KIDI items were simplified because it is a paper and pencil measure. In this situation, the father checked a box to indicate whether he agreed with the statement or whether it was a characteristic of an older or a younger child.

The "answer key" for the set of KIDI items in the 9-month parent CAPI instrument and in the Questions for Fathers is presented in table 3-11, along with the points credited. The analyst then simply sums the correct answers for each case.

Table 3-11. Correct responses for KIDI items, 9-month data collection: 2001–02

Variable name	Description	Correct answer
Parent CAPI Instrument		
P1AMTSLP	P1 PA015A INFANTS NEED SAME AMOUNT OF SLEEP	(2) Disagree (1 point)
P1SIBWET	P1PA015B YOUNG SIBLINGS MAY WET BED	(1) Agree (1 point)
P1CHSPCR	P1 PA015C CHILD THINKS SPEAKS CORRECTLY	(1) Agree (1 point)
P1CHLGCP	P1 PA015D CH LEARN ALL LANG BY COPYING	(2) Disagree (1 point)
P1RIGHTWR	P1 PA020A 1 YEAR OLDS KNOW RIGHT/WRONG	(2) Disagree
P1OYRGWR	P1PA020B OLDER/YOUNGER LEARN RIGHT/WRONG	(2) Older (1 point)
P1KNWSNM	P1 PA025A BABY RESPONDS NAME AT 10 MONTHS	(2) Disagree
P1OYKNNM	P1PA025B OLDER/YOUNGER BABY KNOWS NAME	(1) Younger (1 point)
P1TLTTRN	P1 PA030A CH RDY FOR TOILET TRNG AT 1YR	(2) Disagree
P1OYTLTR	P1 PA030B OLDER/YOUNGER 1 YR.-TOILET TRNG	(2) Older (1 point)
P1REMTOY	P1 PA035A 1 YR-CAN REMEMBER HIDDEN TOYS	(1) Agree (1 point)
P1OYRMTY	P1 PA035B OLDER/YOUNGER 1 HR-HIDDEN TOYS	
P1SHRPLY	P1 PA040A 1 YR OLD SHARE/PLAY TOGETHER	(2) Disagree
P1OYSHPL	P1 PA040B OLDER/YOUNGER 1 YR SHARE/PLAY	(2) Older (1 point)
P1GBRRCH	P1 PA045A BABY CAN GRAB/REACH AT 7 MONTHS	(2) Disagree
P1OYGRRC	P1 PA045B OLDER/YOUNGER 7 MTHS GRAB/RCH	(1) Younger (1 point)
P1WRD1ST	P1 PA050A BABY SAYS 1ST WORD BY 6 MTHS	(2) Disagree
P1OY1SWD	P1 PA050B OLDER/YOUNGER 1ST WRD AT 6 MTH	(2) Older (1 point)
Resident father questionnaire		
F1RGHTWR	F1 Q7A 1 YEAR OLDS KNOW RIGHT/WRONG	(2) Older (1 point)
F1KNWSNM	F1 Q7B BABY RESPONDS TO NAME AT 10 MTHS	(3) Younger (1 point)
F1TLTTRN	F1Q7C RDY FOR TOILET TRNG AT 1 YR	(2) Older (1 point)
F1REMTOY	F1 Q7D 1 YR-CAN REMEMBR HIDDEN TOYS	(1) Agree (1 point)
F1SHRPLY	F1 Q7E 1 YR OLD SHARE/PLAY TOGETHER	(2) Older (1 point)
F1GRBRCH	F1 Q7F BABY CAN GRAB/REACH AT 7 MTHS	(3) Younger (1 point)
F1WRD1ST	F1 Q7G BABY SAYS 1ST WORD AT 6 MTH	(2) Older (1 point)

SOURCE: U.S. Department of Education, National Center for Education Statistics, Early Childhood Longitudinal Study, Birth Cohort, Nine-Month Data Collection, 2001–02.

3.3.2 Quality of the Home Environment: Items from the National Household Education Survey and the Home Observation for Measurement of the Environment

There are two sets of items in the ECLS-B that obtained information about the characteristics of the child's home environment. One set of questions was asked of the parent respondent and has been used in previous National Household Education Survey (NHES) data collection rounds. The other set of items was selected from the Short Form of the Home Observation for Measurement of the Environment (HOME) scale (Caldwell and Bradley, 2001; Caldwell and Bradley, 1979). The following paragraphs describe the home environment questions that were adapted from the NHES. A discussion of the eight items selected from the HOME Short Form concludes this subsection.

A comparison of selected home environment items included in both the ECLS-B and the NHES is presented in table 3-12. The NHES is a household-based data collection system designed to address a wide range of education-related issues. The NHES collects timely data about the educational activities of the U.S. population. NHES surveys were conducted in 1991, 1993, 1995, 1996, 1999, and 2001. Of the items shown in table 3-12, only the question on how often the family reads to the child was asked of infants in the NHES (in both the 1999 and the 2001 NHES). The other items were asked of children between the ages of 3 and 6. For the comparisons shown in table 3-12, the age of the NHES children was restricted to 3 years for those items not asked of infants to make the estimates more comparable to those from the ECLS-B.

The estimates for how often the family reads to the child are quite similar for the ECLS-B and the NHES. Approximately a third of parents reported reading to their infants every day and 22 percent reported reading 3 to 6 times per week. Although the remaining items weren't strictly comparable because of the age differences between the children in the ECLS-B and the children in the NHES, the comparisons are suggestive. It appears that as the children got a bit older, parents were more likely to tell them stories. In the ECLS-B, 44 percent of the infants had family members tell them stories three or more times per week, while more than half of the 3-year-olds in the NHES were likely to have parents report that they told their children stories three or more times per week. On the other hand, as the children got older, the parents appeared less likely to take them out on errands (85 percent of the ECLS-B infants were taken on errands three or more times per week compared to 74 percent of the 3-year-olds in the NHES). In spite of the age differences, the estimates from both surveys indicate that parents were actively engaged in their young children's lives.

Table 3-12. Comparison of frequencies between the ECLS-B and the 1999 and 2001 National Household Education Survey (NHES) for selected items on activities families share with their children

Activities family engages in with child in a typical week	ECLS-B (percent)	NHES (percent)
How often read to child[1]		
Not at all	13	18
Once or twice	34	25
3 to 6 times	22	22
Every day	31	35
How often tell stories to child[2]		
Not at all	23	17
Once or twice	32	30
3 + times per week	44	53
How often sing songs with child[2,3]		
Not at all	2	17
Once or twice	10	27
3 + times per week	89	57
How often take child on errands[2]		
Not at all	2	6
Once or twice	13	20
3 + times per week	85	74

[1] The estimates for this item in the ECLS-B were restricted to children less than 1 year old to make them comparable to the estimates from the 2001 National Household Education Survey (NHES). The Early Childhood Program Participation file from the *National Household Education Survey, 2001 Electronic Codebook and Data File* (NCES 2003-078) was used to obtain estimates of children read to in the past week. The NHES question asks about how often someone read to the child in the past week while the ECLS-B asks about a typical week. NHES estimates are for children who were less than 1 year old as of December 31, 2000.

[2] The ECLS-B and NHES estimates are not strictly comparable for these items. The ECLS-B estimates are for all children in the ECLS-B, most of whom were about 9 months of age at the time of the interview. The Parent and Family Involvement in Education File of the 1999 NHES were used to obtain these estimates. The NHES estimates are for children who were 3 years old as of December 31, 1998. These items were not asked of younger children in the 1999 NHES or in the 2001 NHES.

[3] The questions asked are not identical. ELCS-B asks about how often the family sings songs with the child. The 1999 NHES asks how often the family taught songs to the child.

SOURCE: U.S. Department of Education, National Center for Education Statistics, *Early Childhood Longitudinal Study, Birth Cohort (ECLS-B), 9-Month Public-Use Data File and Electronic Code Book* (NCES 2005–014), forthcoming; U.S. Department of Education, National Center for Education Statistics; *National Household Education Surveys Program 2001 Electronic Codebook and Data File,* Early Childhood Program Participation file (NCES 2003-078), April 2003; U.S. Department of Education, National Center for Education Statistics, *National Household Education Survey 1991 to 1999 Electronic Codebook and Data Files* (1999 Parent and Family Involvement in Education file) (NCES 2002-005), March 2002.

The second set of questions was selected from the Home Observation for Measurement of the Environment (HOME), which has been widely used in research. The HOME is an instrument designed to measure the quality and extent of stimulation in the home environment available to a child.

The HOME serves as a screener to identify environments that are not stimulating to children and to assist in the development of interventions that benefit both the caregiver and the child. Because measurement of the environment needs to take developmental appropriateness into consideration, the HOME has separate inventories for infants and toddlers (birth to 3 years old), early childhood (ages 3 to 6), and middle childhood (ages 6 to 10).

The complete HOME was not feasible for the ECLS-B because it has a large number of items that span 6 subscales. There is a Short Form of the HOME (HOME-SF) that contains only 21 items. Not all of the HOME-SF items were used, however. Many of these items were not appropriate for the ECLS-B because they require extended observation of the child's environment beyond the limits of the home visit. In addition, several of the HOME-SF items are quite similar to the above-mentioned items that were adapted from NHES. For example, the two HOME-SF items, "Child is taken to grocery store at least once a week," and "Child gets out of the house at least 4 times a week," are similar to the NHES item, "How often take child on errands." Therefore, these two items were collapsed into the NHES "How often take child on errands" item in order to maximize the content coverage and to enable the comparability of data across the NHES and ECLS-B data sets.

Accordingly, a subset of 8 HOME-SF items was selected on the basis of interviewers' ability to see the target behavior during a routine home visit, maximize content coverage, as well as to ensure comparability with NHES. Typically, the interviewer would usually either see the behavior in question or not see it. However, for the ECLS-B the option "no opportunity to observe" was added for those uncommon situations where, perhaps, the interview was not done with the child present, or took place at some location other than the home setting. These 8 items could be supplemented by the above-listed NHES items that are answered by the parent respondent. These items (reads to child and takes child on errands) did "double duty" with the home environment questions, as described in the paragraphs above, and with the HOME items completed by the interviewer. These NHES items were also included in the resident father questionnaire. Table 3-13 documents the HOME items in the ECLS-B and summarizes the slight changes in wording.

Table 3-13. Comparison of HOME items in NHES and in the ECLS-B Child Observations, parent CAPI instrument, and resident father questionnaire: 2001–02

Variable name[1]	Variable label (description)	Range of values	Similarity to NHES
R1RSPKCH	R1CO165 RESP SPOKE SPONTANEOUS 2 CHILD	Yes (1) No (2) No opportunity (3)	Added “no opportunity to observe”
R1IORSVB	R1 COI070 RESP VERBAL RESPOND TO CHILD	Yes (1) No (2) No opportunity (3)	Added “no opportunity to observe”
R1IOCRSS	R1 CO175 RESP CARESS/ KISS/HUG CHILD	Yes (1) No (2) No opportunity (3)	Added “no opportunity to observe”
R1IORSHT	R1 C0180 RESP SLAPPED/SPANKED CHILD	Yes (1) No (2) No opportunity (3)	Added “no opportunity to observe”
R1IOINTF	R1 CO185 RESP INTERFERED WITH CH ACTION	Yes (1) No (2) No opportunity (3)	Added “no opportunity to observe”
R1IOPTYS	R1 CO190 RESP PROVIDED TOYS TO CHILD	Yes (1) No (2) No opportunity (3)	Added “no opportunity to observe”
R1IOINVW	R1 CO195 RESP KEPT CHILD IN VIEW	Yes (1) No (2) No opportunity (3)	Added “no opportunity to observe”
R1IOENVS	R1 CO200 PLAY ENVIRONMENT WAS SAFE	Yes (1) No (2) No opportunity (3)	Added “no opportunity to observe”
P1READBO	P1 HE102A HOW OFTEN YOU READ TO CHILD	Not at all (1) Once or twice (2) 3-6 times (3) Every day (4)	Time frame—“in the past week” changed to “in a typical week”
F1READBO	F1 Q2A HOW OFTEN YOU READ TO CHILD	Not at all (1) Once or twice (2) 3-6 times (3) Every day (4)	Time frame—“in the past week” changed to “in a typical week”
P1ERRAND	P1 HE102D HOW OFTN TAKE CHILD ON ERRANDS	Not at all (1) Once or twice (2) 3-6 times (3) Every day (4)	Time frame—“in the past week” changed to “in a typical week.” The category “Every day” was added to the ECLS-B.

See notes at end of table.

Table 3-13. Comparison of HOME items in NHES and in the ECLS-B Child Observations, parent CAPI instrument, and resident father questionnaire: 2001–02—Continued

Variable name	Variable label (description)	Range of values	Similarity to original
F1ERRAND	F1 Q2D HOW OFTEN TAKE CHILD ON ERRANDS	Not at all (1) Once or twice (2) 3-6 times (3) Every day (4)	Time frame—"in the past week "changed to "in a typical week." The category "Every day" was added to the ECLS-B.
P1TELLST	P1 HE102B HOW OFTN YOU TELL CH STORIES	Not at all (1) Once or twice(2) 3-6 times (3) Every day (4)	Time frame—"in the past week" changed to "in a typical week." The category "Every day" was added to the ECLS-B."
F1TELLST	F1 Q2B HOW OFTEN YOU TELL CH STORIES	Not at all (1) Once or twice (2) 3-6 times (3) Every day (4)	Time frame—"in the past week" changed to "in a typical week." The category "Every day" was added to the ECLS-B.
P1SINGSO	P1 HE102C HOW OFTEN YOU ALL SING SONGS	Not at all (1) Once or twice (2) 3-6 times (3) Every day (4)	Time frame—in the past week changed to "in a typical week." The category "Every day" was added to the ECLS-B. The question wording in NHES is slightly different.
F1SINGSO	F1 Q2C HOW OFTEN YOU ALL SING SONGS	Not at all (1) Once or twice (2) 3-6 times (3) Every day (4)	Time frame—"in the past week" changed to "in a typical week." The category "Every day" was added to the ECLS-B. The question wording in NHES is slightly different.

[1] Variable names with the prefix "R1" are from the interviewer observations of the child (Child Observations) during the home visit. Variable names with the prefix "P1" are from the parent CAPI instrument. Variable names with the prefix "F1" are from the resident father self-administered questionnaire.

SOURCE: U.S. Department of Education, National Center for Education Statistics, Early Childhood Longitudinal Study, Birth Cohort, Nine-Month Data Collection, 2001–02.

3.3.3 Additional Subsets of Questions

Additional questions were obtained from many sources. Certain sets of questions of particular analytic interest are noted here. Similar to the KIDI and HOME, they are parent and home environment scales; however, they are more straightforward in terms of modifications and scoring. In most instances, only a few items were selected from each source and some of the items were modified from the original, so the ECLS-B does not purport to be using the original scale. A summary of these items, the construct measured, and how the item was modified is presented in table 3-14.

In the case of the depression questions, a modified version of the Center for Epidemiologic Studies–Depression Scale (CES-D) is used. This modified version was used in the Family and Child Experiences Survey (FACES) and information about the items (FACES items Q2a-g) and their modifications can be found at the FACES web site: http://www.acf.hhs.gov/programs/core/ongoing_research/faces/faces_instruments_parent.html. Then just select the Head Start Parent Interview for 1999.

ECLS-B findings can therefore be compared with the FACES findings. The item scoring is the same as that used for FACES, and a CES-D score can be created as follows. To create the CES-D score:

1. Subtract 1 from the response category for each item so that the range for a particular item is 0-3 instead of 1-4.

2. Count the number of missing or invalid responses for each individual. Code the CES-D score as missing for anyone who has more than 3 invalid (missing or refused) responses on the 12 items.

3. Sum the responses for each individual having a sufficient number of valid responses.

4. Using the following cut points, classify the respondents as nondepressed, mildly depressed, moderately depressed, and severely depressed. (Alternatively, the analyst may want to consider using the CES-D as a continuous variable with higher scores indicating increased severity of depression.)

The cut points are as follows (exhibit 3-2):

Exhibit 3-2. Cut points for the Center for Epidemiologic Studies–Depression (CES-D) scale scores in the 9-month data collection: 2001–02

Classification	Score
Nondepressed	0–4
Mildly depressed	5–9
Moderately depressed	10–14
Severely depressed	15 or higher

SOURCE: U.S. Department of Education, National Center for Education Statistics, Early Childhood Longitudinal Study, Birth Cohort, Nine-Month Data Collection, 2001–02.

Table 3-14. Summary of items from additional parent involvement, parenting attitudes, and home environment scales in the ECLS-B 9-month data file: 2001–02

Name and Source	ECLS-B instrument(s)	Variable name	Variable label (description)	Construct	Degree of modification
Center for Epidemiologic Studies–Depression **(**Radloff, 1977)	Parent self-administered questionnaire	P1BOTHER	P1 SAQ3A RESP UNUSUALLY BOTHERED	Depression	(a) Modified version as used in FACES. See website: http://www.acf.hhs.gov/programs/core/ongoing_research/faces/faces_instruments_parent.html
		P1APPETI	P1 SAQ3B RESP POOR APPETITE		
		P1BLUE	P1 SAQ3C RESP CAN'T SHAKE BLUES		
		P1KPMIND	P1 SAQ3D RESP TROUBLE KEEPING FOCUS		
		P1DEPRES	P1 SAQ3E RESP FELT DEPRESSED		
		P1EFFORT	P1 SAQ3F RESP EVERYTHING AN EFFORT		
		P1FEARFL	P1 SAQ3G RESP FELT FEARFUL		
		P1RESTLS	P1 SAQ3H RESP SLEEP RESTLESS		
		P1TALKLS	P1 SAQ3I RESP TALKED LESS THAN USUAL		
		P1LONELY	P1 SAQ3J RESP FELT LONELY		
		P1SAD	P1 SAQ3K RESP FELT SAD		
		P1NOTGO	P1 SAQ3L RESP COULD NOT GET GOING		

See notes at end of table.

Table 3-14. Summary of items from additional parent involvement, parenting attitudes, and home environment scales in the ECLS-B 9-month data file: 2001–02—Continued

Name and source	ECLS-B instrument(s)	Variable name	Variable label (description)	Construct	Degree of modification
Center for Epidemiologic Studies–Depression (Radloff, 1977)	Resident father questionnaire	F1BOTHRD	F1 Q65A HOW OFTEN FELT BOTHERED	Depression	(a) Modified version as used in FACES. See website:. http://www.acf.hhs.gov/programs/core/ongoing_research/faces/faces_instruments_parent.html
		F1NOEAT	F1 Q65B HOW OFTEN HAD POOR APPETITE		
		F1BLUE	F1 Q65C HOW OFTEN FELT BLUE		
		F1TRBFOC	F1 Q65D HOW OFTEN TROUBLE KEEPING FOCUS		
		F1DEPRSD	F1 Q65E HOW OFTEN FELT DEPRESSED		
		F1EFFORT	F1 Q65F HOW OFTEN FELT EFFORT		
		F1FEARFL	F1 Q65G HOW OFTEN FELT FEARFUL		
		F1RESTLS	F1 Q65H HOW OFTEN RESTLESS SLEEP		
		F1TALKLS	F1 Q65I HOW OFTEN TALKED LESS		
		F1LONELY	F1 Q65J HOW OFTEN FELT LONELY		
		F1SAD	F1 Q65K HOW OFTEN FELT SAD		
		F1NOGO	F1 Q65L HOW OFTEN COULD NOT GET GOING		

See notes at end of table.

Table 3-14. Summary of items from additional parent involvement, parenting attitudes, and home environment scales in the ECLS-B 9-month data file: 2001–02—Continued

Name and source	ECLS-B instrument(s)	Variable name	Variable label (description)	Construct	Degree of modification
Center for Epidemiologic Studies–Depression (Radloff, 1977)	Nonresident father questionnaire	N1BOTHRD	N1 Q36A HOW OFTEN FELT BOTHERED	Depression	(a) Modified version as used in FACES. See website: http://www.acf.hhs.gov/programs/core/ongoing_research/faces/faces_instruments_parent.html
		N1NOEAT	N1 Q36B HOW OFTEN HAD POOR APPETITE		
		N1BLUE	N1 Q36CHOW OFTEN FELT BLUE		
		N1TRBFOC	N1 Q36D HOW OFTEN TROUBLE KEEPING FOCUS		
		N1DEPRSD	N1 Q36EHOW OFTEN FELT DEPRESSED		
		N1EFFORT	N1 Q36F HOW OFTEN FELT EFFORT		
		N1FEARFL	N1 Q36G HOW OFTEN FELT FEARFUL		
		N1RESTLS	N1 Q36H HOW OFTEN RESTLESS SLEEP		
		N1TALKLS	N1 Q36I HOW OFTEN TALKED LESS		
		N1LONELY	N1 Q36J HOW OFTEN FELT LONELY		
		N1SAD	N1 Q36K HOW OFTEN FELT SAD		
		N1NOGO	N1 Q36L HOW OFTEN COULD NOT GET GOING		

See notes at end of table.

Table 3-14. Summary of items from additional parent involvement, parenting attitudes, and home environment scales in the ECLS-B 9-month data file: 2001–02—Continued

Name and source	ECLS-B instrument(s)	Variable name	Variable label (description)	Construct	Degree of modification
Delight Subscale From: Parental Investment in Children (Bradley et al., 1997)	Resident father questionnaire	F1FTHTLK F1CCRYPX F1FTHKCH F1HLDFUN F1FTHGFT	N1 Q5A HOW OFTEN TALK ABOUT CHILD N1 Q5B HOW OFTEN CARRY PICTURES OF CHILD N1 Q5C HOW OFTEN THINK ABOUT CHILD N1 Q5D HOW OFTEN THINK HOLDING CH IS FUN N1 Q5E HOW OFTEN THINK BUYING FOR CH FUN	Paternal delight in child	(b) Some rewording of items Used 5 out of 7 items on the Delight subscale; Delight subscale is 1 out of 4 scales on the 24-item Parent Investment in Children Scale
Delight Subscale From: Parental Investment in Children (Bradley et al., 1997)	Nonresident father questionnaire	N1FTHTLK N1CCRYPX N1FTHKCH N1HLDFUN N1FTHGFT	N1 Q5A HOW OFTEN TALK ABOUT CHILD N1 Q5B HOW OFTEN CARRY PICTURES OF CHILD N1 Q5C HOW OFTEN THINK ABOUT CHILD N1 Q5D HOW OFTEN THINK HOLDING CH IS FUN N1 Q5E HOW OFTEN THINK BUYING FOR CH FUN	Paternal delight in child	(b) Some rewording of items Used 5 out of 7 items on the Delight subscale; Delight subscale is 1 out of 4 scales on the 24-item Parent Investment in Children Scale

See notes at end of table.

Table 3-14. Summary of items from additional parent involvement, parenting attitudes, and home environment scales in the ECLS-B 9-month data file: 2001–02—Continued

Name and source	ECLS-B instrument(s)	Variable name	Variable label (description)	Construct	Degree of modification
Family Practices Scale From: *Authoritarian Family Ideology* (Ernhart and Loevinger, 1969)	Parent CAPI instrument	P1PCKCRY	P1 PA005A SPOIL BABY BY PICK UP WHEN CRY	Authoritarian vs. authoritative parenting attitudes	(a) No modifications Used 5 out of 49 items
		P1TOITRN	P1 PA005B TOILET TRAINING BEFORE 1 YEAR		
		P1FEEDSC	P1 PA005C FEED BABY ON SCHED OR HUNGRY		
		P1HBVSHP	P1 PA005D CHILD BAD HABIT VS HAPPINESS		
		P1MOMVAL	P1 PA005E MOMS TOO PERMISSIVE OR NOT		
Father's Perspective on Role of Fathers in Children's Lives From: Inventory of Father Involvement (Palkovitz, 1984; Hawkins et al., 2002)	Resident father questionnaire	F1FTHPLY	F1 Q16A AGR/DSG FTH SHOULD PLAY W/ CHILD	Attitude about role of fathers in children's lives	(a) No modifications Used 7 out of 43 items
		F1MENFLG	F1 Q16B AGR/DSG SHOW CH LOVE HARD FOR M		
		F1FTHINV	F1 Q16C AGR/DSG FTH INVOLVED IN CH CARE		
		F1FTHEFF	F1 Q16D AGR/DSG FTH TREAT BBY LONG EFFCT		
		F1FTHACT	F1 Q16E AGR/DSG FTH'S ROLE IS TO PROVIDE		
		F1FSPRTM	F1 Q16F AGR/DSG FTH SHOULD SUPPORT MTH		
		F1FTHREW	F1 Q16G AGR/DSG FATHERHOOD VERY RWARDING		

See notes at end of table.

Table 3-14. Summary of items from additional parent involvement, parenting attitudes, and home environment scales in the ECLS-B 9-month data file: 2001–02—Continued

Name and source	ECLS-B instrument(s)	Variable name	Variable label (description)	Construct	Degree of modification
How Father Rates Self as a Father Adapted from the *Parenting Stress Index* with permission of the publisher	Resident and nonresident father questionnaires	F1FTHSLF N1FTHSLF	F1 Q18 HOW YOU FEEL ABOUT BEING A FATHER N1 Q12 HOW YOU FEEL ABOUT BEING A FATHER	Parenting stress and competence	(a) No modification of single item Adapted and reproduced by special permission of the publisher, Psychological Assessment Resources, Inc., 16204 North Florida Avenue, Lutz, FL 33549, from the *Parenting Stress Index–Short Form* by Richard R. Abidin, Ed. D. Copyright 1990 by PAR, Inc. Further reproduction is prohibited without permission from PAR, Inc.

See notes at end of table.

Table 3-14. Summary of items from additional parent involvement, parenting attitudes, and home environment scales in the ECLS-B 9-month data file: 2001–02—Continued

Name and source	ECLS-B instrument(s)	Variable name	Variable label (description)	Construct	Degree of modification
Involvement with Child Early Head Start Research and Evaluation Project (http://www.acf.hhs.gov/programs/core/ongoing_research/ehs/ehs_intro.html)	Parent CAPI instrument (includes path for father respondents)	P1PEEKA	P1 HE105A PLAY PEEKABOO	Caregiving activities	(a) No modifications
		P1TCKL	P1 HE105B TICKLE CHILD		
		P1PLAY	P1 HE105C OUTSIDE TO WALK OR PLAY		
		P1CGDIAP	P1 HE106A DAD RSP[1] CHANGE DIAPER		
		P1PREPFD	P1 HE106B DAD RSP-PREPARE MEALS/BOTTLES		
		P1FEEDBT	P1 HE106C DAD RSP-FEED/GIVE CHILD BOTTLE		
		P1DPEEKC	P1 HE106D DAD RSP-PLAY PEEKABOO W/ CHILD		
		P1HOLDCH	P1 HE106E DAD RSP-HOLD CHILD		
		P1DTCKL	P1 HE106F DAD RSP-TICKLE CHILD		
		P1PTSLEP	P1 HE106G DAD RSP-PUT CHILD TO SLEEP		
		P1WASHCH	P1 HE106H DAD RSP-BATHE CHILD		
		P1DPLAY	P1 HE106I DAD RSP-TAKE CHILD OUTSIDE		
		P1DRSSCH	P1 HE106J DAD RSP-DRESSES CHILD		

See notes at end of table.

.

Table 3-14. Summary of items from additional parent involvement, parenting attitudes, and home environment scales in the ECLS-B 9-month data file: 2001–02—Continued

Name and source	ECLS-B instrument(s)	Variable name	Variable label (description)	Construct	Degree of modification
Involvement with Child Early Head Start Research and Evaluation Project (http://www.acf.hhs.gov/programs/core/ongoing_research/ehs/ehs_intro.html)	Resident father questionnaire	F1TAKECH	F1 Q2D TAKE YOUR CHILD ALONG ON ERRANDS	Caregiving activities	(a) No modification
		F1CGDPR	F1 Q3A HOW OFTEN CHANGE CHILD'S DIAPER		
		F1PREPML	F1 Q3B HOW OFTEN PREPARE MEALS/BOTTLES		
		F1FEEDCH	F1 Q3C HOW OFTEN FEED CHILD		
		F1PKABO	F1 Q3D HOW OFTEN PLAY PEEK-A-BOO W/CHILD		
		F1HOLDCH	F1 Q3E HOW OFTEN HOLD CHILD		
		F1TCKLCH	F1 Q3F HOW OFTEN TICKLE CHILD		
		F1SLPCH	F1 Q3G HOW OFTEN PUT CHILD TO SLEEP		
		F1WASHCH	F1 Q3H HOW OFTEN WASH OR BATHE CHILD		
		F1WALKCH	F1 Q3I HOW OFTEN TAKE CHILD FOR WALK		
		F1DRSSCH	F1 Q3J HOW OFTEN DRESS YOUR CHILD		
		F1WAKECH	F1 Q4A GET UP WITH CHILD DURING NIGHT		
		F1STHCH	F1 Q4B HOW OFTEN SOOTHE CHILD WHEN UPSET		
		F1DCTRCH	F1 Q4C HOW OFTEN TAKE CHILD TO DOCTOR		
		F1CARECH	F1 Q4D HOW OFTEN STAY HOME CARE FOR CH		
		F1TKCCCH	F1 Q4E HOW OFTEN TAKE CHILD TO/FROM CARE		

[1] "DAD RSP" indicates that these items are only asked if the child's father is the primary caregiver and therefore the parent respondent.

SOURCE: U.S. Department of Education, National Center for Education Statistics, Early Childhood Longitudinal Study, Birth Cohort, Nine-Month Data Collection, 2001–02.

REFERENCES

Bayley, N. (1969). *Manual: Bayley Scales of Infant Development*. San Antonio, TX: The Psychological Corporation.

Bayley, N. (1993). *Manual: Bayley Scales of Infant Development, Second Edition*. San Antonio, TX: The Psychological Corporation.

Bradley, R.H., Whiteside-Mansell, L., Brisby, J.A., and Caldwell, B.M. (1997). Parent's socioemotional investment in children. *Journal of Marriage & the Family, 59,* 77-90.

Burns, W.J., Burns, K.A., and Kabacoff, R.I. (1992). Item and factor analyses of the Bayley Scales of Infant Development. In C. Rovee-Collier & L.P. Lipsitt (Eds.), *Advances in Infancy Research, 7,* 199-214.

Caldwell, B. and Bradley R.H. (2003). *Administration Manual: Home Observation for Measurement of the Environment.* Little Rock, AR: University of Arkansas at Little Rock.

Caldwell, B. and Bradley R.H. (1979). *Home Observation for Measurement of the Environment.* Little Rock, AR: University of Arkansas.

Carte, E., Morrison, D., Sublett, J., Venura, A., and Setrakian, W. (1984). Sensory integration therapy: A trial of a specific neurodevelopmental therapy for the remediation of learning disabilities. *Developmental and Behavioral Pediatrics, 5(4),* 189-194.

DeGangi, G.A., Poisson, S., Sickel, R.Z., and Wiener, A.S. (1995). *Infant/Toddler Symptom Checklist.* San Antonio, TX: Therapy Skill Builders, The Psychological Corporation.

Ernhart, C.B. and Loevinger, J. (1969). *Authoritarian Family Ideology: A Measure, Its Correlates, and Its Robustness.* Austin, TX: Society of Multivariate Experimental Psychology.

Hawkins, A.M., Bradford, K.P., Palkovitz, R., Christiansen, S.L., Day, R.D., and Call, V.R.A. (2002). The Inventory of Father Involvement: A pilot study of a new measure of father involvement. *The Journal of Men's Studies, 10,* 183-196.

Ireton, H. (1997). *Child Development Inventories in Education and Health Care: Screening and Assessing Young Children.* Minneapolis, MN: Behavior Science Systems, Inc.

MacPhee, D. (1981). *Manual: Knowledge of Infant Development Inventory.* Chapel Hill, NC: Department of Psychology, unpublished manuscript.

Palkovitz, R. (1984). Role of Father Questionnaire. *Developmental Psychology* 20(6), 1054–60. (Available: University Microfilms Order Number 8105035.)

The Psychological Corporation (2001). *Bayley Short Form–Research Edition.* San Antonio, TX: Author.

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

Sumner, G. and Spietz, A. (1994). *NCAST Caregiver/Parent-Child Interaction Teaching Manual*. Seattle, WA: NCAST Publications, University of Washington.

U.S. Department of Education, National Center for Education Statistics (forthcoming [a]). *Early Childhood Longitudinal Study, Birth Cohort (ECLS-B), Design and Operations Report for the Nine-Month Data Collection.* Washington, D.C.: U.S. Department of Education.

U.S. Department of Education, National Center for Education Statistics (forthcoming [b]). *Early Childhood Longitudinal Study, Birth Cohort (ECLS-B), Methodology Report for the Nine-Month Data Collection, Volume 1: Psychometric Characteristics.* Washington, D.C.: U.S. Department of Education.

This page intentionally left blank.

4. SAMPLE DESIGN, WEIGHTING, AND VARIANCE ESTIMATION

4.1 Sample Design

The Early Childhood Longitudinal Study, Birth Cohort (ECLS-B) selected a nationally representative probability sample of children born in 2001. Children were mostly sampled via registered births from the National Center for Health Statistics (NCHS) vital statistics system (U.S. Department of Health and Human Services, 1997).[1] Births were sampled within a set of primary sampling units (PSUs) and in some cases secondary sampling units (SSUs) in order to control data collection costs.[2] This clustered, list frame design allowed for maximum, efficient coverage of the target population and was considered preferable to other approaches that would involve large screening efforts or coverage errors (e.g., sampling hospitals and other birthing places, augmenting household surveys, or sampling birth certificates available at a variety of levels: NCHS, state registrars, and county and local records offices).[3] Approximately 14,000 births were sampled and fielded and yielded 10,688 9-month completed cases. This sample size was designed to produce survey estimates with specified precision targets both overall and for specific domains (see section 4.1.1). More details on the sample design and precision targets can be found in *Early Childhood Longitudinal Study, Birth Cohort (ECLS-B), Methodology Report for the Nine-Month Data Collection, Volume 2: Sampling* (U.S. Department of Education, National Center for Education Statistics, forthcoming [a]).

The ECLS-B target population consists of all children born in the United States in the year 2001 except the following:

- Children born to mothers less than 15 years of age;
- Children who died before the 9-month assessment; and
- Children who were adopted prior to the 9-month assessment.

Children born to mothers less than 15 years of age were excluded from the sampling frame in response to state confidentiality and sensitivity concerns. Sampled children whom the states identified as having died before or been adopted prior to the 9-month assessment were removed from the study in field operations.

---

[1] In two PSUs, births were selected from lists provided by hospitals.
[2] A PSU is an individual county or group of contiguous counties. An SSU is a subdivision of a PSU.
[3] See Levine and Bryant (1997).

4.1.1 Analytic Domains and Effective Sample Sizes

The ECLS-B was designed to support statistical analyses in the following analytic domains:

1. Child's race/ethnicity (as derived from birth certificate data)

- American Indian;[4]
- Chinese;
- Other Asian or Pacific Islander (i.e., excluding Chinese);
- Hispanic;
- Black, non-Hispanic; or
- White, non-Hispanic.

2. Birth weight

- Very low (less than 1,500 grams);
- Moderately low (1,500 to 2,500 grams); or
- Normal (greater than 2,500 grams).

3. Plurality

- Twin; or
- Non-twin (single birth and other multiple births).

When combined, these analytic domains result in 36 distinct birth certificate sampling strata (hereafter called case strata).

The race/ethnicity domains are mutually exclusive and exhaustive. The American Indian, Chinese, and Other Asian/Pacific Islander groups are rare and required sampling at higher rates than the other race/ethnicity groups. The race/ethnicity of the infant was assigned based on hierarchical rules that used the mother and father's race/ethnicity as they appeared on the birth certificate. Note that birth certificate data do not specify the child's race/ethnicity individually. The hierarchical rules worked to implement specific definitions desired for certain subgroups (i.e., for the American Indian [mother <u>or</u> father] and Chinese subgroups [mother <u>and</u> father]) and to assign cases with missing or unknown

[4] American Indian includes Alaska Natives.

race/ethnicity to subgroups with larger sampling rates (an approach that benefits smaller subgroups analytically). Exhibit 4-1 shows how a child's race/ethnicity for sampling was derived from the mother's and father's race/ethnicity as reported on the birth certificate. The hierarchical logic was as follows:

1. If the mother or father was reported as American Indian, then the child was classified as American Indian;

2. Otherwise, if the mother and father were reported as Chinese, then the child was classified as Chinese;

3. Otherwise, if the mother was reported as Chinese or Asian/Pacific Islander or the father was reported as Chinese or Asian/Pacific Islander, then the child was classified as Other Asian/Pacific Islander;

4. Otherwise, if the mother was reported as Hispanic, then the child was classified as Hispanic;

5. Otherwise, if the mother was reported as Black, non-Hispanic, then the child was classified as Black, non-Hispanic;

6. Otherwise, if the father was reported as Hispanic, then the child was classified as Hispanic;

7. Otherwise, if the father was reported as Black, non-Hispanic, then the child was classified as Black, non-Hispanic;

8. Otherwise, if the mother was reported as White, non-Hispanic or the father was reported as White, non-Hispanic, then the child was classified as White, non-Hispanic;

9. All remaining cases (both mother and father's race/ethnicity was reported as missing): the child was classified as Other Asian/Pacific Islander.

In some cases, the infant's race/ethnicity reported in the survey and appearing on the data file differs from that assigned for sampling. This can occur because of the differences between birth certificate and ECLS-B data collection techniques; differences between parent reports and the hierarchical logic described above; and the fact that the current multi-race classification was not in operation for birth certificate data at the time of sampling. Thus, 2001 birth certificates did not allow for multi-race responses, and the sample was therefore selected based on a single race classification. In contrast, the ECLS-B classification employs the current NCES standards that allow for multi-race responses.

Exhibit 4-1. Rules for assigning infant's race/ethnicity based on parents' race/ethnicity from birth certificates, in 9-month data collection: 2001–02

Mother's race/ethnicity[1]	Father's race/ethnicity[1]						
	American Indian	Chinese	Other Asian/Pacific Islander (PI)	Hispanic	Black, non-Hispanic	White, non-Hispanic	Missing
American Indian	American Indian	American Indian	American Indian	American Indian	American Indian	American Indian	American Indian
Chinese	American Indian	Chinese	Other Asian/PI	Other Asian/PI	Other Asian/PI	Other Asian/PI	Other Asian/PI
Other Asian/Pacific Islander (PI)	American Indian	Other Asian/PI	Other Asian/PI	Other Asian/PI	Other Asian/PI	Other Asian/PI	Other Asian/PI
Hispanic	American Indian	Other Asian/PI	Other Asian/PI	Hispanic	Hispanic	Hispanic	Hispanic
Black, non-Hispanic	American Indian	Other Asian/PI	Other Asian/PI	Black, non-Hispanic	Black, non-Hispanic	Black, non-Hispanic	Black, non-Hispanic
White, non-Hispanic	American Indian	Other Asian/PI	Other Asian/PI	Hispanic	Black, non-Hispanic	White, non-Hispanic	White, non-Hispanic
Missing	American Indian	Other Asian/PI	Other Asian/PI	Hispanic	Black, non-Hispanic	White, non-Hispanic	Other Asian/PI

[1] In cases of multiple races, assignment to race/ethnicity is hierarchical from top to bottom for mother and from left to right for father. Thus, the child of a mother who is American Indian and Hispanic would be assigned race/ethnicity based on the first row of the table, i.e., "American Indian."
SOURCE: U.S. Department of Education, National Center for Education Statistics, Early Childhood Longitudinal Study, Birth Cohort (ECLS-B), Nine-Month Data Collection, 2001–02.

Furthermore, the classification of a multi-race response into a single race category is not straightforward and would not necessarily agree with the response to a single race question. The two methods of classifying race/ethnicity also differ in the classification of cases as Hispanic. Table 4-1 gives the cross-classification of the child's sampling (W1CRAC) and parent report (X1CHRACE) of race/ethnicity.

For any given analysis, users will need to decide how they want to define race/ethnicity. Using the American Indian population as an example, 873 completed cases were classified as American Indian for sampling purposes, but only 286 were classified as American Indian for the composite of child's race (X1CHRACE). Recall that for sampling purposes, a case was classified as American Indian if either parent reported that she or he was American Indian. For the composite X1CHRACE, a much more stringent definition was used. To be classified as American Indian for X1CHRACE, three conditions needed to be met: (1) the parent reported the child was American Indian; (2) the parent did not report any other race for the child (e.g., American Indian and White or American Indian and Black); and (3) the parent reported that the child was not Hispanic. As can be seen in table 4-1, 286 cases met this more stringent definition. The file also contains dichotomous race variables (X1CHAMIN, X1CHASN, X1CHPIC, X1CHBLCK, X1CHWT, X1CHHISP, and X1CHMLRC) that can be used to subset the sample by race. These composites take values of "1" whenever a particular race/ethnicity is reported regardless of whether other races or ethnicities are also reported. For the dichotomous American Indian composite (X1CHAMIN), 749 cases are classified as American Indian, much closer to the number obtained using the sampling definition (873). Users will need to carefully consider which of these available race/ethnicity variables are best suited for their purposes or whether they want to create their own categorizations of race/ethnicity.

The two analytic groups of special interest in the birth weight domain are those with very low birth weight (under 1,500 grams) and those with moderately low birth weight (between 1,500 and 2,500 grams). The analytic group of special interest in the plurality domain is twins. The last groups listed for each of the last two domains (normal birth weight and single births and other non-twins) are included to provide an exhaustive classification of the population, but they are not subgroups that required oversampling. The small proportions of births in the low and very low birth weight and the twin groups were such that each required appreciable oversampling in order to obtain required sample sizes.

Table 4-1. Cross-classification of child's sampling and parent report of race/ethnicity, 9-month data collection: 2001–02

Child's race/ethnicity from sampling (W1CRAC)	Total	Parent report of child's race/ethnicity								
		American Indian or Alaska Native, non Hispanic	Asian, non-Hispanic	Native Hawaiian or other Pacific Islander, non-Hispanic	Hispanic, race specified	Hispanic, no race specified	Black, non-Hispanic	White, non-Hispanic	More than 1 race	Not ascertained
Total	10,688	293	1,195	49	1,514	679	1,698	4,441	780	39
American Indian	873	286	14	2	148	24	16	127	250	6
Chinese	466	1	458	1	0	0	0	0	1	5
Other Asian/ Pacific Islander	1,345	2	712	40	122	21	22	108	305	13
Hispanic	1,581	0	0	0	974	560	8	34	3	2
Black, non-Hispanic	1,771	0	2	3	53	12	1,619	13	65	4
White, non-Hispanic	4,652	4	9	3	217	62	33	4,159	156	9

SOURCE: U.S. Department of Education, National Center for Education Statistics, Early Childhood Longitudinal Study, Birth Cohort (ECLS-B), Nine-Month Data Collection, 2001–02.

Before reporting counts of completed cases, the complexity of the ECLS-B instruments must be acknowledged. The 9-month data collection consisted of three separate components: data were collected from the parent, using computer-assisted personal interviewing (CAPI) and a self-administered questionnaire, from child assessments and observations, and from the father (resident father, nonresident father or both) with a self-administered questionnaire. The parent CAPI instrument was required for all cases included in the respondent data file. Table 4-2 details the number of cases with 9-month data, by component (Parent, Child and Father), overall, and for each analytic subgroup. See section 4.2.1 for a more thorough discussion of component combinations and sets of weights. Note also that these are actual sample sizes. See section 4.3.3 for discussion of design effects, which work to reduce these actual sample sizes to effective sample sizes. Note that some of the cases included in table 4-2 are ineligible due to deaths and adoptions. Such cases could not be identified at the time of sampling and excluded, unlike children born to mothers less than 15 years of age. (See section 4.1 for the description of the ECLS-B target population.)

Note that other analytic subgroups were considered but were not included explicitly in the ECLS-B sample design. For example, children at risk for disability were of particular interest, but this subgroup could not be oversampled effectively via birth certificate data. Children in poverty were also of interest. [5] Although this subgroup could not be oversampled effectively, it was estimated that there would be a reasonable sample size of such children in the ECLS-B sample, as the percentage of children in poverty has hovered around 20 percent since 1981 (Federal Interagency Forum on Child and Family Statistics, 1999).

Table 4-3 gives the actual yields for children in poverty (less than 100 percent of poverty) by children's sampling race/ethnicity, birth weight, and plurality.

[5] Children are considered to be poor if they live in families with incomes less than 100 percent of the poverty threshold for a family of their size. The ECLS-B estimated childhood poverty by comparing parental reports of income against the 2000 weighted poverty thresholds published by the U.S. Bureau of the Census.

Table 4-2. Number of cases with component data, overall and by domain, in the 9-month data collection: 2001–02

Domain	Sample	Target parent completes[3]	Completed interviews: Parent	Parent-Child	Parent-Father	Parent-Child-Father
Total	14,197	10,870	10,688	10,221	6,988	6,816
Race/ethnicity[1]						
American Indian	1,086	880	873	835	551	542
Chinese	732	637	466	418	352	324
Other Asian/Pacific Islander	1,832	1,209	1,345	1,287	885	871
Hispanic	2,096	1,517	1,581	1,501	918	900
Black, non-Hispanic	2,161	1,746	1,771	1,715	817	802
White, non-Hispanic	6,014	4,881	4,652	4,465	3,465	3,377
Ineligible[2]	276	†	†	†	†	†
Birth weight						
Very low	1,473	1,268	1,151	1,093	709	685
Moderately low	2,055	1,512	1,647	1,585	1,037	1,013
Normal	10,393	8,090	7,890	7,543	5,242	5,118
Ineligible[2]	276	†	†	†	†	†
Plurality						
Twin	2,023	1,357	1,658	1,585	1,141	1,113
Non-twin (Single birth and other multiple births)	11,898	9,513	9,030	8,636	5,847	5,703
Ineligible[2]	276	†	†	†	†	†

† Not applicable.
[1] Race/ethnicity information was determined from the birth certificate as defined in exhibit 4-1.
[2] Ineligible cases include deaths and adoptions.
[3] The target parent completes represent the expected yield, at the parent component level, of the sample after sample reduction.
SOURCE: U.S. Department of Education, National Center for Education Statistics, Early Childhood Longitudinal Study, Birth Cohort (ECLS-B), Nine-Month Data Collection, 2001–02.

Table 4-3. Number of 9-month parent completes with children in poverty, by race/ethnicity, birth weight, and plurality: 2001–02

Children's sampling race/ethnicity	Total 9-month parent completes	9-month parent completes with children in poverty (less than 100 percent of poverty)
Total	10,688	2,603
Race/ethnicity		
American Indian	873	316
Chinese	466	31
Other Asian/Pacific Islander	1,345	208
Hispanic	1,581	591
Black, non-Hispanic	1,771	807
White, non-Hispanic	4,652	650
Birth weight		
Very low (less than 1,500 grams)	1,151	321
Moderately low (1,500 to 2,500 grams)	1,647	446
Normal (greater than 2,500 grams)	7,890	1,836
Plurality		
Twin	1,658	378
Non-twin (Single birth and other multiple births)	9,030	2,225

SOURCE: U.S. Department of Education, National Center for Education Statistics, Early Childhood Longitudinal Study, Birth Cohort, (ECLS-B), Nine-Month Data Collection, 2001–02.

4.1.2 Core Sample Design

The core ECLS-B sample consists of births sampled within 96 PSUs and represents all infants born in the U.S. in the year 2001. In addition, to provide adequate precision in the American Indian analysis domain, an additional sample of 18 PSUs was selected from a supplemental frame consisting of areas where the population has a higher proportion of American Indian births. These PSUs were located mainly in the western region of the U.S.

Births are registered in the state of occurrence, although birth certificates include the state and county of both occurrence and residence. Membership in an ECLS-B PSU could, therefore, be determined based on either occurrence (where the birth occurred) or residence (where the child's mother lived as reported on the birth certificate). Each of these approaches had advantages and disadvantages. The major advantage to sampling by occurrence was that it simplified the sample selection. States were asked to provide the birth certificates only for births occurring in the selected PSUs in their own state, and only states that contained sampled PSUs needed to provide certificates for sampling. In contrast, sampling by residence would have required that all states participate in the ECLS-B, regardless of the states in which the ECLS-B PSUs happen to fall, since the resident of a given state and county may give birth in any state and county.

On the other hand, sampling by occurrence had the disadvantage of a wider geographical spread of births, since the sample could not be clustered into small geographic units by residence. However, an evaluation of the geographical and temporal spread of the 9-month sample and consideration of the increased geographical spread in later rounds due to mobility led to the conclusion that the somewhat wider geographical spread of an occurrence-based sample would not be a serious disadvantage. Therefore, the ECLS-B sample was drawn using birth certificates by occurrence.

As discussed in section 4.1.1, a set of analysis domains were defined based on race/ethnicity (American Indian, Chinese, Other Asian/Pacific Islander, Hispanic, Black non-Hispanic, and White non-Hispanic); birth weight (very low, moderately low, and normal); and plurality (twins versus other births). These subgroups were crossed to define a set of 36 mutually exclusive case strata. A composite measure of size (MOS) that utilized the analysis domains was computed for each PSU using the formula given in Folsom, Potter, and Williams (1987).

First, let

N_k = average number of births in 1994-1996 in the k-th subgroup;
n_k = required sample size for the k-th subgroup; and
f_k = n_k / N_k..

Let M_{ik} denote the number of births in the k-th subgroup within the i-th PSU. The adjusted MOS for the i-th PSU is computed as

$$M_i = \sum_k f_k M_{ik} .$$

Thus, the measure of size is computed as a weighted sum over the analysis domains, with the weight for a given domain being computed as the required sample size divided by overall population size for the domain. The objective of weighting the MOS was to obtain equal workloads within PSUs to meet the sample size requirements discussed in section 4.1.1.

Contiguous counties were combined to form PSUs for the ECLS-B. Westat used metropolitan statistical area (MSA) definitions in large metropolitan areas and the NCHS health service areas[6] (U.S. Department of Health and Human Services, National Center for Health Statistics, 1991) as a guide for combining counties to form PSUs in other areas. The health service areas are relatively self-contained in terms of health service supply and demand; they were especially useful in forming PSUs in rural areas, where it was often necessary to combine several counties because of the relatively low incidence of births. In developing the frame, counts of births were averaged over the years 1994, 1995, and 1996 to obtain annualized estimates for each county.

Twenty-four PSUs were large enough to be selected with certainty. The remaining PSUs were stratified on the following variables:

- Census region (Northeast, Midwest, South, West);
- MSA status (metro area, non-metro area);
- Minority status (high, low);
- Median income (high, low); and
- Size (continuous sort on MOS—see above).

Census region was determined by the Federal Information Processing Standards (FIPS) state and county codes of counties making up a given PSU. PSUs were constructed within census region boundaries. MSA status was also determined by the FIPS state and county codes of the counties making up the PSU, with the entire PSU considered as an MSA PSU when one or more counties within the PSU was a member of an MSA. PSUs were constructed to maintain existing MSA definitions, with the exception of MSAs that involved more than one noncertainty state (i.e., a state where all PSUs within the

[6] The National Center for Health Statistics health service areas were created for the National Health Interview Survey.

state had probabilities of selection less than 1.0). Minority status was determined within the two stratification variables above based on the distribution of non-White births. The high and low categories were stratum-specific and divided the stratum approximately in half with respect to MOS. Median income stratification was done similarly. Stratification by size was carried out using a serpentine (i.e., alternatingly ascending or descending) sort on the MOS within the first four stratification variables given above. Two PSUs were selected per stratum with probability proportional to size, using Durbin's method (Durbin, 1967). In some cases, PSUs were subdivided into secondary sampling units (SSUs) which also consisted of contiguous counties. A sample of SSUs was drawn within PSUs in order to decrease travel-related costs.

4.1.3 American Indian Supplemental PSU Sample Design

It is difficult to select an adequate sample of American Indian (AI) births in a general purpose survey. The proportion of AI births in most areas is quite small; furthermore, many AI births are concentrated in a few, sparsely populated areas. Because of these difficulties, an AI supplemental PSU frame was developed, and an AI supplemental PSU sample was selected.

To develop the AI PSU frame the study identified counties or groups of counties that were large enough to provide an efficient interviewer workload (at least 50 AI births) and had a relatively large proportion of American Indian births. The AI PSU frame contained 108 PSUs, most of which consisted of individual counties. The stratification of the frame was geographic by region with finer breakdowns by state within some regions. A supplemental sample of 18 AI PSUs was selected from the AI PSU frame, of which 6 were selected with certainty because they contained more than 2.4 percent of all AI births, which was the overall sampling rate for the AI supplemental sample. The noncertainty PSUs were selected independently of the core sample PSUs, with probability proportional to the number of AI births. One of these PSUs included the Navajo Nation reservation, whose institutional review board did not approve participation in the study. Sampled cases that were contained within the reservation boundaries were treated as nonresponse. This loss of sample size was taken into account when adjusting rates for the overall sample reduction (see section 4.2.2.)

The full AI supplemental sample consists of AI cases selected within the following:

- The 18 PSUs selected from the AI PSU frame and

- The counties *not* covered by the 108 PSUs on the AI PSU frame, but sampled in the core ECLS-B PSU sample.

This provides 100 percent coverage of the AI population.

Thus, AI births within any given county had a single, known chance of being included in the AI supplemental sample. All counties also had a single chance of being included in the core ECLS-B PSU sample. In total, counties covered by the AI PSU frame have *two chances* of participating in the ECLS-B. However, the two samples—core and AI supplemental—are distinct, and any given county had only one chance of being in the AI PSU sample and one chance of being in the core PSU sample.

4.1.4 Consent Processes and Restrictions

Since birth certificates received through the NCHS vital statistics system were used as the sampling frame, the first step in study enrollment was to obtain permission from state registrars. Birth certificates contain information about the child, the mother and the father, including the mother's address. This information was used both to stratify the birth certificates for sampling and to contact mothers of sampled infants. In several states, obtaining permission for access to birth certificates required approval from the state's institutional review board. In addition, some state institutional review boards or registrar offices had requirements that placed restrictions on contacting parents based on birth certificate information. In some cases, these restrictions would have resulted in low response rates or even complete nonparticipation.

The ECLS-B core and AI supplemental PSU samples included births occurring in counties within 46 states and in Washington, D.C. Ten of the 46 states had state institutional review board or registrar office requirements that put restrictions on the standard ECLS-B contact and consent protocol. The types of restrictions were as follows:

- **Passive consent.** In states requiring passive consent, the state registrar sent a letter to the sampled parents informing them about the study. If the parents did *not* respond, then they were assumed to have given consent for Westat to contact them for enrollment into the study.

- **Active consent.** In states requiring active consent, either Westat or the state registrar sent a letter to parents informing them about the study. However, in this case, parents *must* have responded giving consent to contact them. If they did not respond to the letter, they were considered not giving consent.

- **Active consent states with PRAMS.** Some states require active consent to participate in the Pregnancy Risk Assessment Monitoring System (PRAMS). In these states, there was already a program in place where local officials contacted mothers to get consent for participation in PRAMS. Obtaining consent for ECLS-B could have been included in this program.

- **Consent specified on birth certificate.** In one state, parents filled out an item on the birth certificate that allowed persons to contact them. Although this item did not refer specifically to research studies,[7] the state used the response to this question as an indicator of the parent's willingness to be contacted subsequent to the birth. That is, the registrar released information about the parent only if the parent responded affirmatively to the question.

- **No follow-back.** In some cases, state law prohibited the use of birth certificate information to contact parents for "follow-back" research studies like the ECLS-B.

In some cases, larger sample sizes were selected to offset the additional nonresponse that was expected because of enrollment restrictions. However, in states that required active consent or that prohibited follow-back research studies, substitution and alternative frames were used (see section 4.1.5, below). In one state (Illinois), state law allowed for the Bureau of the Census to conduct the interviews through a subcontract with NCHS.

In the 36 states (and Washington, D.C.) that did not impose restrictions, the standard ECLS-B contact and consent protocol were used. See section 5.3.2 for a description of these procedures.

4.1.5 Substitution and Alternative Frames

Substitute PSUs were selected to replace the PSUs originally selected for ECLS-B in a few states. The substitute PSU "replaced" the original PSU in the sample, meaning that a sample of birth certificates that matched the sample planned for the original PSU with respect to sample size and demographic composition was selected in the substitute PSU.

Wherever possible, substitute PSUs were selected from the same sampling stratum as the original PSU, at least for noncertainty PSUs. Substitute PSUs for the main sample were matched on:

- Median household income;
- Percent minority population;

[7] The item on the birth certificate reads, "Parent(s) authorization to receive infant health care information: Yes or No."

- Percent of population in poverty;
- Birth rate;
- Population density; and
- Census region.

Matching was based primarily on a Mahalanobis distance function (Mahalanobis, 1975). In addition, the study looked at the individual rankings of the absolute differences of the values for the potential substitute PSU from the value for the original PSU for each of the above variables. Also, the geographic profile of each potential substitute was compared with that of the original PSU in order to ensure that the selected substitute was a reasonable match for the original PSU not only demographically, but geographically as well.

In the supplemental sample for AI births, substitute PSUs were matched on the number of AI births, percent of mothers with high school education, percent in which both parents were AI, and mean birth weight. In addition, candidate substitute PSUs were reviewed by an in-house Westat expert on American Indians, who provided information about the social structure, culture, and geographic surroundings of American Indians living in the different regions in the United States. Selection of the substitute PSUs was made taking the latter considerations into account. For example, high-population-density California PSUs with no associated tribal lands and a limited tribal structure were not substituted for low population density PSUs located in reservation areas with a strong tribal structure.

Within the substitute PSU (core or AI supplemental), sampling rates that would deliver the same yield as expected within the original PSU by case stratum (e.g., Hispanic, normal birth weight, or twin) were used. A total of seven PSUs were used as substitutes for original ECLS-B PSUs.

In two instances, an alternative sampling frame was used to draw a sample of births occurring within AI PSUs with enrollment restrictions. Specifically, birth records were selected directly from hospital lists of births in counties that defined the original PSUs.

4.1.6 Sampling Birth Certificates Within Primary and Secondary Sampling Units

Individual birth certificates within PSUs or SSUs (hereafter PSUs) were sampled from data files provided by state registrars. These data files were processed through the NCHS state-based vital statistics system. Sampled children subsequently identified by state registrars as having died or having been adopted after the issuance of the birth certificate were excluded from the sample.

NCHS receives birth certificate data from the states on a flow basis, with the number of births received and months included in a given shipment varying by state and throughout the year. Births were thus sampled systematically throughout 2001 and 2002. The within-PSU sampling rates varied by case stratum and depended on the PSU selection probability such that, within each case stratum, each sampled birth had an equal overall probability of selection. After a particular batch was received, births were selected within sampling strata, within PSUs present in the batch, based on appropriate selection intervals, continuing from the point where the last sample selection left off.

The within-PSU sampling rate for a stratum depended on the PSU selection probability and was determined to give each sampled birth in a given case stratum the same overall probability of selection. The within-PSU selection probabilities for each of the sampling strata were determined from the following equation:

Sampling rate = $P_{hik} = P_k / P_{hi}$,

where

P_{hik} = the within-PSU selection probability for case stratum *k* in PSU *i* in PSU stratum *h*;

P_k = the desired overall selection probability for a birth in case stratum *k*; and

P_{hi} = probability of selection for ECLS-B PSU *i* in PSU stratum *h*

Note that the within-PSU sampling rates cannot exceed 1. As can be seen from the above equation, this implies that the PSU selection probability must be no less than the overall selection probability for every sampling stratum. The construction of PSUs was carried out to satisfy this condition.

4.2 Calculation and Use of Sampling Weights

This section describes the ECLS-B 9-month data weighting procedures. Weights are used to adjust for disproportionate sampling, survey nonresponse, and noncoverage of the target population when analyzing complex survey data. The weights are designed to eliminate or reduce biases that would otherwise occur with unweighted analyses. The ECLS-B weights were developed in three steps:

1. Base weights were calculated using the overall selection probabilities;
2. Base weights were adjusted for survey nonresponse; and
3. Raking was used to adjust for undercoverage and to improve precision of survey estimates.

The following sections describe these steps as well as estimation and variance estimation. Section 4.2.1 describes the four sets of weights that were created for use when analyzing the ECLS-B data. Section 4.2.2 describes how base weights were calculated. Section 4.2.3 describes the weight adjustments for nonresponse and undercoverage. Section 4.3 discusses variance estimation and design effects. Further details on these procedures can be obtained from *ECLS-B Methodology Report for the Nine-Month Data Collection, Volume 2: Sampling* (U.S. Department of Education, National Center for Education Statistics, forthcoming [a]).

4.2.1 Four Types of Sampling Weights

The 9-month data collection consisted of three main components:

1. Parent interviews (parent CAPI instrument and parent self-administered questionnaire);
2. Child assessments (the physical measurements, Bayley Short Form–Research Edition [BSF-R], Nursing Child Assessment Teaching Scale [NCATS]); and
3. Father questionnaires (self-administered resident and nonresident father questionnaires).

For the purpose of constructing the ECLS-B 9-month weights, a completed case was defined at the component level, as follows:

Parent-level complete. Cases where sections up to and including the child development (CD) section were complete. The count of items up to and including the CD section is 58, out of 160 minimum for the entire parent component, or about 36 percent of the parent items. Some 10,688 cases meet these criteria.

Child-level complete. Cases where either the physical measurements or BSF-R mental scale or motor scale sections were complete. There are 10,221 cases that meet these criteria. The 10,688 9-month parent completes break down as follows with respect to these child components (table 4-4):

Table 4-4. Parent completes, broken down by completed and missing BSF-R and physical measurements components, 9-month data collection: 2001–02

Parent completes		
Bayley Short Form–Research Edition	Physical measurements	Number
Total		10,688
Complete	Complete	9,948
Complete	Missing	252
Missing	Complete	21
Missing	Missing	467

SOURCE: U.S. Department of Education, National Center for Education Statistics, Early Childhood Longitudinal Study, Birth Cohort (ECLS-B), Nine-Month Data Collection, 2001–02.

Father-level complete. Father-level completes were defined as cases where the resident father questionnaire (RFSAQ) only, the nonresident father questionnaire (NRFSAQ) only, or both component items were complete (i.e., provided one or more item responses in the relevant self-administered questionnaire), as appropriate for the case. Only a very small proportion of cases required both resident father and nonresident father questionnaires. These cases are treated as father-level completes only if both components were completed. There are 6,998 cases that meet these criteria.

These three main components could be completed in various combinations at the case-level (i.e., sampled birth), resulting in seven possible weights. The seven possible weights involve data from the following:

1. The parent interview only;
2. The child assessment only;
3. The father questionnaire only;
4. The parent interview and the child assessment only;
5. The parent interview and the father questionnaire only;
6. The father questionnaire and the child assessment only; or
7. The parent interview, the child assessment, and the father questionnaire.

In principle, a separate set of weights is appropriate for a given set of measures if the set of measures is of analytic interest and if there is an acceptable level of response for the set. A different set of weights should be constructed for each possible type of analysis to maximize the available sample. However, such a proliferation of weights (and the associated replicate weights) is unwieldy and impractical for analysts. Therefore, the sets of weights were restricted to combinations that are of analytic interest and for which the loss in sample size from using a set of more complete records would be sizable. Cross-component frequencies on the 9-month data indicated that the child assessment was rarely completed without the parent interview. Thus, combinations (2) and (6) above were not useful, since analysis of the child assessment data alone (item 2) or child assessment with the father data (item 6) can be conducted using the weights developed for combinations (4) or (7), respectively.

On the other hand, a sufficient number of cases had parent interview and child assessment data but did not have father data, so weights for both combinations (4) and (7) above are useful. Also, 170 cases had parent interview and father component data but did not have child data, so weights for combination (5) were developed in order to utilize these cases. Finally, a sufficient number of cases had parent interview data only so that weights for item (1) would be warranted. An insufficient number of cases had the father data only, so that possibility (3) was eliminated.

Following the above rationale, four sets of weights were constructed. Exhibit 4-2 gives the variable names for each set of weights, and indicates the kind of variables with which they should be used.

Exhibit 4-2. ECLS–B cross-sectional weights in the 9-month data collection: 2001–02

Weight	To be used for analysis of …
W1R0	9-month parent interview data (“P1” prefixed variables), and/or birth certificate data (“BC” prefixed variables) and a limited set of child characteristics (e.g., sex and race/ethnicity)
W1C0	9-month child assessment measures(“C1” prefixed variables) alone or in combination with parent interview and/or birth certificate data
W1F0	Resident (“F1” prefixed variables) and/or nonresident father data (“N1” prefixed variables) alone or in combination with data from the parent interview and/or birth certificate
W1FC0	Resident and/or nonresident father data in combination with child assessment data alone or in combination with parent interview and/or birth certificate data.

SOURCE: U.S. Department of Education, National Center for Education Statistics, Early Childhood Longitudinal Study, Birth Cohort, Nine-Month Data Collection, 2001–02.

Weight W1R0 is used to estimate child-level characteristics associated with data collected through the parent interview and/or birth certificate. Examples include the percentage of children whose mothers are currently employed, the percentage of children who are in a particular type of child care, the percentage of children who are in excellent or very good health, and the percentage of children who were low birth weight.

Weight W1C0 is used to estimate child-level characteristics associated with data collected through the child assessments either alone or in combination with data collected through the parent interview and/or birth certificate. Examples include a comparison of children’s mean height and weight by children’s food security status, a comparison of children’s Bayley mental T-scores by family background characteristics, or the NCATS total parent score by household socioeconomic status (SES).

Weight W1F0 is used to estimate child-level characteristics associated with data reported by fathers either alone or in combination with data collected through the parent interview and/or birth certificate. Analysts who are using data from the resident father questionnaire should include the 51 cases where the father is the respondent to the parent interview (X1RSPREL=3). As noted in section 2.3.1, all questions contained in the resident father questionnaire are also included in special paths for father respondents in the parent interview. Examples of analyses using this weight are the percent of resident fathers who read to their children 3 or more times per week, the percent of resident and/or nonresident fathers who first held their child within an hour after delivery, the percent of resident and/or nonresident fathers who feel that they are very good fathers. This weight should not be used if child assessment data are also being used. For analyses of child assessment data in combination with father data, W1FC0 should be used instead.

The weight W1FC0 is used to estimate child-level characteristics using data reported by fathers in combination with child assessment data. This weight would also be used for analyses involving father data, child assessment data, and data from the parent interview and/or birth certificate data. An example of an analysis using this weight is a comparison of children's Bayley mental T-scores by levels of resident father involvement and family background characteristics such as parent education.

4.2.2 Base Weights

The overall probability of selection for a given birth is computed as the product of each stage's probability of selection. For birth certificates sampled in PSU sampling stratum *h*, PSU *i*, SSU *j*, case stratum *k* at time *t*, the selection probability is

$$p_{kt} = p_{hi}\, p_{hij}\, p_{hijkt}\,.$$

Originally, the sample was designed to yield a constant overall probability of selection for births belonging to a given case stratum. Exceptions to this rule are due to sampling for the American Indian supplement and a reduction in sampling rates; the latter took place in March 2002.[8]

The base weight gives the approximate representation of each sampled birth certificate. The base weight for a given birth record was calculated as the reciprocal of the overall probability of selection:

$$w_{kt} = \frac{1}{p_{kt}}\,.$$

[8] In March 2002, the ECLS-B sample size was reduced in order to control costs.

4.2.3 Weighting Adjustments

The main objective in adjusting base weights is to correct for survey nonresponse and survey undercoverage. All ECLS-B weights were adjusted to correct for survey nonresponse and undercoverage. The data weighting adjustments for the Parent set started with the base weights, whereas the data weighting for the Parent-Child set and Parent-Father set started with the adjusted Parent weight, and the data weighting for the Parent-Father-Child set started with the adjusted Parent-Father weight. Adjusting to population totals through post-stratification or raking can also reduce standard errors. Weighting adjustments can be separated into two general types:

- Sample-based adjustments, such as nonresponse adjustments and
- Population-based adjustments, such as post-stratification and raking.

In sample-based adjustments, data available for both respondents and nonrespondents are used to adjust the respondent sample so that it represents the whole sample (i.e., both respondents and nonrespondents). In population-based adjustments, data from respondents are used to adjust the sample to the population. Sample-based adjustments require that the same data items be available for respondents and nonrespondents. Given the hierarchical nature of the weights being developed for the ECLS-B (see section 4.2.1):

1. Birth certificate data were available for the Parent set of weights;
2. Birth certificate and parent component data were available for the Parent-Child and Parent-Father sets of weights; and
3. Birth certificate, parent, and father component data were available for the Parent-Father-Child sets of weights.

Specifically, the following characteristics were used to define sample-based nonresponse adjustment cells for each set of weights (exhibit 4-3). The characteristics listed were selected based on analyses using segmentation modeling via the Chi-Squared Automatic Interaction Detector (CHAID) (Flores-Cervantes et al., 2002; Kass, 1980; Magidson, 1993). That analysis indicated which characteristics and partitions using those characteristics explain differences in response rates.

Population-based adjustments require that the same data items be available for respondents and the total population. Given the birth certificate data that were available for all sampled cases and for the total population, raking was used to ensure that sums of sampling weights match known population totals. The raking ratio estimation procedure is based on the iterative proportional fitting procedure of Deming and Stephan (1940) and involves ratio adjustments to two or more marginal distributions of the population counts. Raking was proposed by Deming and Stephan as a way to ensure consistency between complete counts and sample data from the 1940 U.S. Census.

The raking procedure is carried out in a sequence of adjustments: The weights are adjusted to one marginal distribution and then to the second marginal distribution, and so on. One sequence of adjustments to the marginal distributions is known as a cycle or iteration. The procedure is repeated until convergence is achieved. The criteria for convergence can be specified either as the maximum number of iterations or an absolute difference (or relative absolute difference) from the known marginal population totals.

Exhibit 4-3. Variables used to define sample-based nonresponse adjustment cells for each set of weights, 9-month data collection: 2001–02

Parent weights

Foreign resident status
Population of county of residence
Child’s race/ethnicity
Child’s birth weight
Five-minute APGAR score
Plurality
Father’s race/ethnicity
Mother’s age
Mother’s education
Live birth order
Number of prenatal visits
Delivery complications
Medical risk factors

Parent-Father weights

Cases where only resident fathers apply:

Foreign resident status
Child’s race/ethnicity
Happiness in marriage as reported in the parent component [1]
Father’s education—composite [1]
Delivery complications
Father’s age—composite [1]
Medical risk factors
Five-minute APGAR score
Population size of MSA
Father’s race/ethnicity [1]
Number of hours per week spouse worked [1]
Population of county of residence
Plurality
Mother’s age
Number of prenatal visits
Live birth order
Mother’s education

Cases where only nonresident fathers apply:

Child’s race/ethnicity
Why the biological father hasn’t seen the child [1]
Population of county of residence
Mother’s age
Region of residence
How far away the biological father lives [1]
Mother’s education

Cases where both resident fathers and nonresident fathers apply:

These cases were assigned to their own adjustment cell.

Parent-Child weights

Foreign resident status
Population of county of residence
Child’s race/ethnicity
Child’s health as reported in the parent component [1]
Child’s hospitalization status as reported in the parent component [1]
Five minute APGAR score
Plurality
Father’s race/ethnicity
Mother’s age
Mother’s education
Live birth order
Number of prenatal visits
Medical risk factors

Parent-Father-Child weights

Cases where only resident fathers apply:

Foreign resident status
Region of residence
Child’s race/ethnicity
Father’s education—composite [1]
Population of county of residence
Delivery complications
Happiness in marriage as reported in the parent component [1]
Medical risk factors
Population size of MSA
Father’s race/ethnicity [1]
Number of hours per week spouse worked [1]
Five minute APGAR score
Plurality
Live birth order
Father’s age—composite [1]
Mother’s age
Number of prenatal visits

Cases where only nonresident fathers apply:

Child’s race/ethnicity
Plurality
Why the biological father hasn’t seen the child [1]
Population of county of residence
Mother’s age
Mother’s education

Cases where both resident fathers and nonresident fathers apply:

These cases were assigned to their own adjustment cell.

[1] These variables came from ECLS-B components; all other variables came from birth certificate data.
SOURCE: U.S. Department of Education, National Center for Education Statistics, Early Childhood Longitudinal Study, Birth Cohort (ECLS-B), Nine-Month Data Collection, 2001–02.

Each of the four sets of sample-based nonresponse adjusted weights was raked to the following 11 dimensions (derived from birth certificate data):

1. Child's sex;
2. Child's race/ethnicity based on information from the birth certificate;
3. Child's plurality (singleton, twin, triplets, and higher order births);
4. Population size of MSA;
5. Region of residence;
6. Presence of medical risk factors;
7. Child's birth weight;
8. Mother's age at time of birth;
9. Total birth order;
10. Mother's education at time of birth; and
11. Number of prenatal visits.

These variables were selected because of their substantive interest as well as their relationship to response propensity, as indicated by logistic regression analyses. Table 4-5 reports selected characteristics of the four sets of weights: the name of the full sample weight, the count of records with a non-missing value of that weight, and the mean, standard deviation, coefficient of variation, minimum, maximum, skewness, kurtosis, and sum of that weight.

Table 4-5. Characteristics of cross-sectional weights, 9-month data collection: 2001–02

Variable name	Number of cases	Mean	Standard deviation	Coefficient of Variation (× 100)	Minimum	Maximum	Skewness	Kurtosis	Sum
W1R0	10,688	373.99	349.31	93.40	4.08	1,854.91	0.745	-0.639	3,997,169
W1C0	10,221	391.07	365.54	93.47	4.72	1,981.78	0.769	-0.559	3,997,169
W1F0	6,988	517.76	509.06	98.32	5.53	3,850.44	1.076	1.031	3,618,138
W1FC0	6,816	530.83	525.25	98.95	5.95	4,050.91	1.164	1.578	3,618,138

SOURCE: U.S. Department of Education, National Center for Education Statistics, Early Childhood Longitudinal Study, Birth Cohort (ECLS-B), Nine-Month Data Collection, 2001–02.

4.3 Variance Estimation

The precision of estimates from a sample survey can be estimated by their sampling variance. Replication and Taylor Series linearization methods have been developed for complex sample designs like the ECLS-B. These methods account for clustering, multiple stages of selection, and differential sampling rates characteristic of more complicated designs. One or both of these methods are supported by software designed specifically for analyzing complex survey data such as AM, Stata, SUDAAN, WesVar, and SAS.[9] The following sections describe the paired jackknife replication (JK2) and the Taylor Series estimation methods as they pertain to the ECLS-B.

4.3.1 Paired Jackknife Replication Method (JK2)

In this method, a survey estimate of interest (e.g., mean BSF-R motor and/or mental score) is calculated from the full sample. Subsamples of the full sample are then selected to calculate subsample estimates of the same parameter. The subsamples are called *replicates*, and the subsample estimates are called *replicate estimates*. The variability of the replicate estimates around the full sample estimate is used to estimate the variance of the full sample estimate. The variance estimator is computed as the sum of the squared deviations of the replicate estimates from the full sample estimate:

$$v\left(\hat{\theta}\right) = \sum_{g=1}^{G}\left(\hat{\theta}_{(g)} - \hat{\theta}\right)^2$$

where

$\hat{\theta}$	is the estimate of the population parameter of interest, θ based on the full sample;
$\hat{\theta}_{(g)}$	is the g^{th} replicate estimate of θ based on the observations included in the g^{th} replicate; and
G	is the number of replicates formed.

For the ECLS-B, in which some first-stage units were selected with certainty and the remaining first-stage units were selected taking two units per stratum, the paired jackknife replication method (JK2) is recommended. The variance estimates of selected survey items presented in section 4.3.3 were produced using the paired jackknife replication method in WesVar. The paired jackknife replication

[9] Information about AM, Stata, SUDAAN, WesVar, and SAS can be obtained at http://www.am.air.org, http://www.stata.com, http://www.rti.org/sudaan, http://www.westat.com/wesvar/, and http://www.sas.com, respectively.

method can also be used in SUDAAN with the appropriate statements and keywords. To achieve this in SUDAAN, the statement JACKWGTS with the keyword ADJJACK can be used as follows:

JACKWGTS varlist / ADJJACK = 1,
where varlist is the list of replicate weights.

A set of 90 replicate weights was created for variance estimation for each of the four ECLS-B sets of sampling weights. Each set of replicate weights was calculated using the same adjustment steps as the full sample weight (see section 4.2.3), but using only the subsample of cases that constituted each replicate. The replicate weights for each set of sampling weights are presented in exhibit 4-4.

Exhibit 4-4. Replicate weights for each set of sampling weights, 9-month data collection: 2001–02

Set	Full sample weight	Replicate Weights
Parent	W1R0	W1R1…W1R90
Parent-Child	W1C0	W1C1…W1C90
Parent-Father	W1F0	W1F1…W1F90
Parent-Father-Child	W1FC0	W1FC1…W1FC90

SOURCE: U.S. Department of Education, National Center for Education Statistics, Early Childhood Longitudinal Study, Birth Cohort (ECLS-B), Nine-Month Data Collection, 2001–02.

4.3.2 Taylor Series Linearization Method

The Taylor Series linearization method produces a linear approximation to the estimate of interest; then the variance of the linear approximation is estimated by standard variance formulas. The stratum and first-stage unit identifiers are required to implement this approach. One set of stratum and first-stage unit identifiers is sufficient for the ECLS-B. The stratum and first-stage unit identifiers are as follows:

Stratum	First-stage unit
W1RSTR	W1RRPSU

The first-stage unit identifier provided for variance estimation (W1RRPSU) is masked to limit the risk of disclosure. That is, units selected within the same PSU share the same values of W1RRPSU; however, the value within W1RRPSU is otherwise meaningless (i.e., does not correspond with FIPS codes, etc.). Exhibit 4-5 provides a quick reference on which weights and variables to use for estimation and variance estimation for the four main kinds of analysis.

Exhibit 4-5. Specifications for computing standard errors in the 9-month data collection: 2001–02

Type of analysis	Full sample weight	Computing standard errors: Replication method (SUDAAN, WESVAR): Replicate weights	Computing standard errors: Replication method (SUDAAN, WESVAR): Jackknife method	Computing standard errors: Taylor Series method (AM, SAS, Stata, SUDAAN): Nesting variables	Approximating sampling errors: DEFT (Median root design effect)
Parent component measures	W1R0	W1R1-W1R90	JK2	W1RSTR, W1RRPSU	1.4492
Parent and child component measures	W1C0	W1C1-W1C90	JK2	W1RSTR, W1RRPSU	1.9256
Parent and father component measures	W1F0	W1F1-W1F90	JK2	W1RSTR, W1RRPSU	1.3436
Parent, father, and child component measures	W1FC0	W1FC1-W1FC90	JK2	W1RSTR, W1RRPSU	1.3476

SOURCE: U.S. Department of Education, National Center for Education Statistics, Early Childhood Longitudinal Study, Birth Cohort (ECLS-B), Nine-Month Data Collection, 2001–02.

4.3.3 Standard Errors and Design Effects for Selected ECLS-B Variables

A large number of variables were collected from parents, children, and fathers in the ECLS-B. The estimate from each variable has its own variance, standard error, and design effect. The design effect compares the statistical efficiency of survey estimates from complex sample designs to what would have been obtained under a simple random sample (SRS) of the same size.

In a stratified clustered design like the ECLS-B, stratification generally leads to a gain in efficiency over simple random sampling, but clustering has the opposite effect because of the positive intracluster correlation of the units in the cluster. Also, the oversampling of rare domains is beneficial for analyses concentrating on those domains, but inefficient for analyses that cut across the oversampled domains due to the differential weighting effect. The basic measure of the relative efficiency of the sample design is called the *design effect* (DEFF), defined as the ratio, for a given statistic, of the variance estimate under the actual sample design to the variance estimate that would be obtained with an SRS of the same sample size:

$$DEFF = \frac{Var_{DESIGN}}{Var_{SRS}}$$

The square root of the design effect, DEFT, is defined as

$$DEFT = \frac{SE_{DESIGN}}{SE_{SRS}}$$

where *SE* is the standard error of the estimate. The design effect indicates the increase or decrease in the variance relative to simple random sampling, and the square root of the design effect indicates the same for the standard error.

The preferred methods of computing variances, standard errors and design effects for ECLS-B variables are the paired jackknife method or the Taylor Series linearization method as described in sections 4.3.1 and 4.3.2, using programs designed specifically for analyzing complex survey data. The standard errors and design effects presented in this section were calculated using the paired jackknife replication method in WesVar.

If a statistical analysis software package such as SPSS (Statistical Program for the Social Sciences) is used, the standard errors should be corrected using DEFT, since these programs calculate standard errors assuming the data were collected with an SRS. The standard error of an estimate under the actual sample design can be calculated as follows:

$$SE_{DESIGN} = \sqrt{DEFF \times Var_{SRS}} = DEFT \times SE_{SRS}$$

Packages such as SAS or SPSS can be used to obtain Var_{SRS} and SE_{SRS}. Alternatively, Var_{SRS} and SE_{SRS} can be computed using the formulas below for means and proportions.

Means

$$Var_{SRS} = \frac{1}{n} \frac{\sum_{1}^{n} w_i \left(x_i - \bar{x}_w\right)^2}{\sum_{1}^{n} w_i} = SE_{SRS}^2$$

where w_i are the sampling weights, n is the number of respondents in the sample, and the sample mean $\bar{x}_w$ is calculated as follows:

$$\bar{x}_w = \frac{\sum_1^n w_i x_i}{\sum_1^n w_i}$$

Proportions

$$Var_{srs} = \frac{p(1-p)}{n} = SE_{SRS}^2$$

where p is the weighted estimate of proportion for the characteristic of interest and n is the number of cases in the sample. In both cases of means and proportions, the standard error assuming simple random sampling should be multiplied by DEFT to get the standard error of the estimate under the actual design.

Tables 4-6 and 4-7 show standard errors and design effects for selected means and proportions based on ECLS-B parent, child, and father component data. For each survey component, the tables present the selected data item, variable name, number of cases, the estimate, the standard error taking into account the actual sample design (Design SE), the standard error under simple random sampling (SRS SE), the design effect (DEFF), and the square root of the design effect (DEFT). The estimate presented for each item in table 4-6 follows particular conventions. If the item is a continuous or interval measure, then the estimate is the mean for all cases (excluding “Don’t knows,” “Not ascertained” etc.) If the item is discrete (i.e., nominal or ordinal), then the estimate is the proportion of cases responding with the first or lowest possible meaningful value in an ordinal sense (excluding “Don’t knows,” “Not ascertained” etc.). For more information on the variables used in this section, refer to chapter 3, which describes the assessment and rating scale scores used in the ECLS-B, and chapter 7, which has a detailed discussion of the other variables.

Note that the design effects for assessment estimates (e.g., X1MTLTSC, X1MTRTSC) are higher than the design effects for some other estimates (e.g., P1ATTN, P1CGDTAP etc.) This can be due to either naturally occurring higher intracluster correlations for assessment estimate items or interviewer effects. In the case of the ECLS-B, where the general relationship between interviewer and cluster is one-to-one, the two are difficult if not impossible to disentangle. See section 5.5.5 for a discussion of assessor effects. Similar observations about the design effects for assessment estimates were made in the Early Childhood Longitudinal Study, Kindergarten Class of 1998–99 (ECLS-K) data (see the *ECLS-K Third Grade Methodology Report*, section 3.1[U.S. Department of Education, National Center for Education Statistics, forthcoming (b)]).

The results presented in tables 4-6 and 4-7 will allow analysts using the appropriate variance estimation software to check their work. Also, the design effects in these tables give analysts a way to approximate the actual sample design's standard errors when they are unable to use the software suggested. One way to use design effects in roughly approximating the actual sample design's standard errors for either overall estimates or subgroup estimates not presented in tables 4-6 and 4-7, (respectively), is to use an average (more specifically, the median) design effect across a number of variables or for a specific subgroup. Table 4-7 presents median design effects by various analytic subgroups and may be used as described above to inflate SRS standard errors.

For example, suppose that analysts are interested in obtaining the estimate of the proportion of children whose parents characterize them as "never demanding attention and company constantly" (P1ATTN=0), and are unable to use the suggested variance estimation software. They would be able to obtain an estimate of 18.99 via a weighted analysis (i.e., using the full sample weight). A SRS standard error can be obtained by using the weighted estimate in the numerator and the unweighted sample size in the denominator; the standard error for this estimate would be 0.377 (the square root of ((18.99 x 80.01)/10,682)).

Applying the design effects (DEFT, specifically) in table 4-7 (Overall row, Parent DEFT column), they would estimate the design's standard error to be 0.5463 (= 1.4491 X 0.377). Note that this is reasonably close to the true design's standard error of 0.550 (from table 4-6). The degree to which this method will produce reasonable approximations of standard errors will vary across estimates. Analysts should review table 4-6 to grasp this variability, accounting for differences between sets of weights and kinds of variables (e.g., demographic versus assessments scores, etc.).

Analysts looking for estimates for various subgroups would look at the other rows (e.g., male, female etc.) in table 4-7. The same approach can be used for means and parameters in regression models, obtaining weighted estimates of the mean or regression parameters and unweighted estimates of the SRS standard errors.

Table 4-6. Standard errors and design effects for the full sample, 9-month data collection (overall): 2001–02

Weight	Survey item	Variable name	Number of cases	Estimate	Design SE	SRS SE	DEFF	DEFT
P	P1 CD120C DEMANDS ATTENTION AND COMPANY	P1ATTN	2,074	18.99	0.550	0.380	2.100	1.449
P	P1 CH005 STATUS OF CHILD HEALTH	P1CHEALT	6,251	61.70	0.661	0.471	1.973	1.405
P	P1 HE102D HOW OFTN TAKE CHILD ON ERRANDS	P1ERRAND	437	1.91	0.174	0.132	1.738	1.318
P	P1 CD120A IS FUSSY OR IRRITABLE	P1FUSSY	3,729	36.90	0.628	0.467	1.808	1.345
P	P1 CD120E NEEDS HELP TO FALL ASLEEP	P1HLPSLP	5,932	56.17	0.786	0.480	2.680	1.637
P	P1 CD120G CRIES FOR FOOD OR TOYS	P1NOWAIT	2,980	27.86	0.624	0.434	2.070	1.439
P	P1 HE102A HOW OFTEN YOU READ TO CHILD	P1READBO	1,383	12.12	0.479	0.316	2.296	1.515
P	P1 HE102C HOW OFTEN YOU ALL SING SONGS	P1SINGSO	237	2.07	0.197	0.138	2.040	1.428
P	P1 CD120F STARTLED BY LOUD SOUNDS	P1STRTL	6,027	58.46	0.742	0.477	2.421	1.556
P	P1 HE102B HOW OFTN YOU TELL CDH STORIES	P1TELLST	2,480	22.29	0.642	0.402	2.545	1.595
P	P1 CD120D WAKES UP 3 OR MORE TIMES	P1WAKES	7,293	70.12	0.816	0.443	3.391	1.841
P	P1 CD120B GOES FROM WHIMPER TO CRYING	P1WHMPR	4,629	45.95	0.714	0.482	2.190	1.480
P	X1 HSHLD FOOD SECURITY - STATUS CTGRY	X1FSSTAT	9,331	87.73	0.554	0.317	3.046	1.745
P	X1 TYPE RES FATHER-BIRTH/ADOPT/STEP/FOST	X1FTHTYP	8,304	78.73	0.483	0.396	1.488	1.220
P	X1 TYPE OF FAMILY	X1HFAMIL	5,558	49.65	0.538	0.484	1.237	1.112
P	X1 CH PARENTS WHO RESIDE IN HOUSEHOLD	X1HPARNT	8,282	78.55	0.482	0.397	1.471	1.213
P	X1 HH FATHER/MALE GUARD-WRK STAT W/IMP	X1IFTHLB	7,916	94.59	0.332	0.246	1.819	1.349
P	X1 HH MOTHER/FEMALE GUARD-WRK STAT W/IMP	X1IMOMLB	6,221	60.28	0.934	0.475	3.865	1.966
P	X1 POVERTY INDICATOR	X1POVRTY	2,603	22.87	0.600	0.406	2.179	1.476
P	X1 PRIM CARE ARRNGMNT WHERE MOST HRS/WK	X1PRIMNW	5,352	49.95	0.794	0.484	2.695	1.642
P	X1 QUINTILE INDICATOR FOR SOCIOECON SCAL	X1SESQ5	2,110	20.05	0.455	0.387	1.379	1.174
P	P1 CD090 AGE WHEN STARTED CRAWLING(MTHS)	P1AGCRWL	9,056	7.17	0.024	0.016	2.360	1.536
P	P1 CD080 AGE SITTING WITHOUT SPPRT(MTHS)	P1AGSIT	10,268	6.08	0.019	0.013	2.042	1.429
P	P1 CD100 AGE PULLED SELF TO STAND(MTHS)	P1AGSTND	8,842	7.99	0.025	0.015	2.733	1.653
P	P1 CD110 AGE WALKING WITH HELP(MONTHS)	P1AGWALK	7,863	8.61	0.029	0.016	3.364	1.834
P	X1 HSHLD FOOD SECURITY-SCALE SCR (RASCH)	X1FSSCAL	2,485	3.66	0.061	0.042	2.135	1.461

See note at end of table.

Table 4-6. Standard errors and design effects for the full sample, 9-month data collection (overall): 2001–02—Continued

Weight	Survey item	Variable name	Number of cases	Estimate	Design SE	SRS SE	DEFF	DEFT
P	X1 TOTAL NUMBER OF HOUSEHOLD MEMBERS	X1HTOTAL	10,688	4.31	0.017	0.014	1.453	1.205
P	X1 NUMBER OF HH MEMBERS LESS THAN 18	X1LESS18	10,688	2.13	0.013	0.012	1.238	1.113
PC	X1 CHILD BIRTH WEIGHT STATUS	X1BTHWGT	7,501	92.51	0.002	0.000	0.000	0.000
PC	X1 CHILD SEX	X1CHSEX	5,221	51.06	0.087	0.494	0.031	0.176
PC	X1 RACE/ETHNICITY - CHILD	X1CHRACE	4,262	53.38	0.541	0.494	1.197	1.094
PC	X1 MULTIPLE BIRTH STATUS INDICATOR	X1MBRTST	8,486	96.82	0.002	0.000	0.000	0.000
PC	X1 CHILD SAMPLED AS TWIN	X1TWSAMP	1,572	2.94	0.029	0.167	0.030	0.173
PC	X1 PM - CHILD'S WEIGHT (KG)	X1CHWGHT	9,993	9.57	0.044	0.017	7.039	2.653
PC	X1 MENTAL PROB1: EXPLORES OBJECTS	X1MTL1	10,195	0.99	0.000	0.000	3.708	1.926
PC	X1 MENTAL PROB2: EXPLORES PURPOSEFULLY	X1MTL2	10,195	0.91	0.003	0.001	7.287	2.699
PC	X1 MENTAL PROB3: BABBLES	X1MTL3	10,195	0.55	0.005	0.002	7.892	2.809
PC	X1 MENTAL PROB4: EARLY PROBLEM SOLVING	X1MTL4	10,195	0.09	0.004	0.002	6.198	2.490
PC	X1 MENTAL PROB5: USES WORDS	X1MTL5	10,195	0.03	0.002	0.001	5.106	2.260
PC	X1 MOTOR PROB1: EYE-HAND COORDINATION	X1MTR1	10,166	0.92	0.001	0.001	3.601	1.898
PC	X1 MOTOR PROB2: SITTING	X1MTR2	10,166	0.95	0.001	0.001	3.096	1.760
PC	X1 MOTOR PROB3: PRE-WALKING	X1MTR3	10,166	0.79	0.004	0.002	4.117	2.029
PC	X1 MOTOR PROB4: INDEPENDENT WALKING	X1MTR4	10,166	0.31	0.007	0.003	5.576	2.361
PC	X1 MOTOR PROB5: BALANCE	X1MTR5	10,166	0.06	0.003	0.001	5.645	2.376
PC	X1 NCATS - TOTAL CHILD SCORE	X1NCATTC	8,606	15.50	0.048	0.029	2.724	1.650
PC	X1 NCATS - TOTAL PARENT SCORE	X1NCATTP	8,608	34.68	0.082	0.049	2.851	1.688
PC	X1 NCATS - TOTAL SCORE	X1NCATTS	8,608	50.18	0.102	0.062	2.698	1.643
PF	F1 Q3A CHANGE CHILDS DIAPER	F1CGDIAP	3,244	48.77	0.849	0.637	1.778	1.333
PF	F1 Q4D STAY HOME W/ILL CHILD	F1STYHM	783	11.71	0.651	0.411	2.506	1.583
PF	F1 Q4C TAKE CHILD TO DOCTOR	F1DCTR	1,333	18.62	0.598	0.497	1.446	1.202
PF	F1 Q2D HOW OFTN TAKE CHILD ON ERRANDS	F1ERRAND	732	9.81	0.474	0.379	1.564	1.251
PF	F1 Q3C FEED/GIVE CHILD BOTTLE	F1FEEDBT	3,133	45.96	0.892	0.637	1.961	1.400

See note at end of table.

Table 4-6. Standard errors and design effects for the full sample, 9-month data collection (overall): 2001–02—Continued

Weight	Survey item	Variable name	Number of cases	Estimate	Design SE	SRS SE	DEFF	DEFT
PF	F1 Q4A WAKE W/CHILD AT NIGHT	F1GETUP	892	14.06	0.777	0.443	3.074	1.753
PF	F1 Q3D PLAY PEEKABOO W/CHILD	F1PEEKC	2,210	34.68	0.798	0.609	1.718	1.311
PF	F1 Q3B PREPARE MEALS/BOTTLES	F1PREPFD	3,018	45.31	0.836	0.635	1.731	1.316
PF	F1 Q2A HOW OFTEN YOU READ TO CHILD	F1READBO	1,774	27.88	0.775	0.572	1.833	1.354
PF	F1 Q2C HOW OFTEN YOU ALL SING SONGS	F1SINGSO	771	11.30	0.514	0.406	1.605	1.267
PF	F1 Q4B SOOTHE UPSET CHILD	F1SOOTHE	1,194	17.91	0.727	0.491	2.195	1.482
PF	F1 Q2B HOW OFTN YOU TELL CH STORIES	F1TELLST	1,874	29.24	0.722	0.584	1.528	1.236
PF	N1 Q12 HOW YOU FEEL ABOUT BEING A FATHER	N1FTHSLF	11	3.08	1.207	0.666	3.284	1.812
PF	N1 Q9 FATHER WANTED CH WHEN BEC PRGNT	N1FWTBBY	363	54.32	2.797	1.936	2.088	1.445
PF	N1 Q13 HOW IS RELATIONSHIP W/CHILD'S MOM	N1RELM	549	79.18	2.326	1.567	2.203	1.484
PF	N1 Q3 HOW OFTEN SPENT 1+ HOURS W/CHILD	N1SPHR1	15	2.42	0.864	0.626	1.906	1.381
PF	N1 Q6 HOW OFTEN TALK WITH CHILD'S MOTHER	N1TALKM	540	79.77	2.162	1.557	1.929	1.389
PF	X1 TYPE RES FATHER-BIRTH/ADOPT/STEP/FOST	X1FTHTYP	6,218	86.15	0.481	0.413	1.355	1.164
PF	X1 HH FATHER/MALE GUARD-WORK STATUS	X1HFEMP	5,177	85.60	0.623	0.450	1.914	1.383
PF	X1 NON-RES FATHER WORK STATUS	X1NRFEMP	362	60.88	2.529	1.956	1.671	1.293
PF	N1 Q2 SEEN CHILD IN LAST 3 MONTHS	N1SEE3MO	576	50.80	1.766	1.393	1.607	1.268
PF	X1 HH FATHER-OCC GSS PRESTIGE SCORE	X1FTHSCR	5,674	42.85	0.180	0.135	1.776	1.333
PCF	F1 Q3A CHANGE CHILDS DIAPER	F1CGDIAP	3,173	48.83	0.857	0.644	1.769	1.330
PCF	F1 Q4D STAY HOME W/ILL CHILD	F1STYHM	766	11.78	0.661	0.418	2.504	1.582
PCF	F1 Q4C TAKE CHILD TO DOCTOR	F1DCTR	1,304	18.75	0.596	0.504	1.397	1.182
PCF	F1 Q2D HOW OFTN TAKE CHILD ON ERRANDS	F1ERRAND	706	9.60	0.514	0.379	1.835	1.355
PCF	F1 Q3C FEED/GIVE CHILD BOTTLE	F1FEEDBT	3,067	46.07	0.899	0.645	1.943	1.394
PCF	F1 Q4A WAKE W/CHILD AT NIGHT	F1GETUP	868	13.98	0.778	0.447	3.028	1.740
PCF	F1 Q3D PLAY PEEKABOO W/CHILD	F1PEEKC	2,164	35.05	0.822	0.618	1.771	1.331
PCF	F1 Q3B PREPARE MEALS/BOTTLES	F1PREPFD	2,953	45.45	0.858	0.643	1.781	1.335
PCF	F1 Q2A HOW OFTEN YOU READ TO CHILD	F1READBO	1,741	28.03	0.794	0.580	1.871	1.368

See note at end of table.

Table 4-6. Standard errors and design effects for the full sample, 9-month data collection (overall): 2001–02—Continued

Weight	Survey item	Variable name	Number of cases	Estimate	Design SE	SRS SE	DEFF	DEFT
PCF	F1 Q2C HOW OFTEN YOU ALL SING SONGS	F1SINGSO	746	11.29	0.540	0.411	1.723	1.313
PCF	F1 Q4B SOOTHE UPSET CHILD	F1SOOTHE	1,163	17.98	0.749	0.497	2.274	1.508
PCF	F1 Q2B HOW OFTN YOU TELL CH STORIES	F1TELLST	1,827	29.38	0.724	0.592	1.497	1.224
PCF	N1 Q12 HOW YOU FEEL ABOUT BEING A FATHER	N1FTHSLF	11	3.18	1.239	0.688	3.246	1.802
PCF	N1 Q9 FATHER WANTED CH WHEN BEC PRGNT	N1FWTBBY	352	54.49	2.635	1.966	1.797	1.341
PCF	N1 Q13 HOW IS RELATIONSHIP W/CHILD'S MOM	N1RELM	533	79.24	2.453	1.591	2.377	1.542
PCF	N1 Q3 HOW OFTEN SPENT 1+ HOURS W/CHILD	N1SPHR1	15	2.47	0.891	0.643	1.920	1.386
PCF	N1 Q6 HOW OFTEN TALK WITH CHILD'S MOTHER	N1TALKM	521	79.08	2.254	1.601	1.983	1.408
PCF	X1 TYPE RES FATHER-BIRTH/ADOPT/STEP/FOST	X1FTHTYP	6,067	86.14	0.472	0.418	1.274	1.129
PCF	X1 HH FATHER/MALE GUARD-WORK STATUS	X1HFEMP	5,059	85.54	0.621	0.456	1.853	1.361
PCF	X1 NON-RES FATHER WORK STATUS	X1NRFEMP	352	60.96	2.575	1.987	1.680	1.296
PCF	N1 Q2 SEEN CHILD IN LAST 3 MONTHS	N1SEE3MO	558	50.61	1.680	1.419	1.402	1.184
PCF	X1 HH FATHER-OCC GSS PRESTIGE SCORE	X1FTHSCR	5,538	42.87	0.178	0.137	1.694	1.302

SOURCE: U.S. Department of Education, National Center for Education Statistics, Early Childhood Longitudinal Study, Birth Cohort (ECLS-B), Nine-Month Data Collection, 2001–02.

Table 4-7. Median design effects in the 9-month data collection, by component measure and subgroup: 2001–02

Domain	Parent		Parent-Child		Parent-Father		Parent-Father-Child	
	DEFF	DEFT	DEFF	DEFT	DEFF	DEFT	DEFF	DEFT
Overall	2.1000	1.4491	3.7080	1.9256	1.8055	1.3436	1.8160	1.3476
Sex								
Male	1.8690	1.3671	2.5780	1.6055	1.8560	1.3624	1.8765	1.3699
Female	1.8950	1.3766	2.7995	1.6725	1.9220	1.3864	1.9380	1.3921
Race/ethnicity								
American Indian or Alaska native, non-Hispanic	1.6690	1.2919	3.2740	1.8094	1.7180	1.3107	1.7245	1.3132
Asian, non-Hispanic	1.6220	1.2736	2.3110	1.5199	1.6230	1.2740	1.5920	1.2617
Native Hawaiian or other Pacific Islander, non-Hispanic	1.8690	1.3671	1.7645	1.3283	1.8985	1.3778	1.8290	1.3524
Hispanic, no race specified	1.4860	1.2190	2.0720	1.4393	1.5330	1.2381	1.5675	1.2520
Hispanic, race specified	1.5630	1.2502	1.6140	1.2703	1.8000	1.3416	1.7880	1.3372
Black, non-Hispanic	2.4425	1.5627	1.5895	1.2607	2.2870	1.5120	2.2780	1.5090
White, non-Hispanic	1.5270	1.2357	1.7420	1.3195	1.4765	1.2125	1.5530	1.2452
More than one race specified	2.4660	1.5704	2.9060	1.7047	2.6050	1.6140	2.8560	1.6900
Plurality								
Singleton	1.8620	1.3646	3.5460	1.8831	1.5385	1.2404	1.5530	1.2462
Twin	1.7900	1.3379	2.1585	1.4692	2.2835	1.5110	2.2340	1.4947
Higher order	1.6760	1.2946	1.4125	1.1883	1.7245	1.3119	1.7055	1.3050

See note at end of table.

Table 4-7. Median design effects in the 9-month data collection, by component measure and subgroup: 2001–02—Continued

Domain	Parent		Parent-Child		Parent-Father		Parent-Father-Child	
	DEFF	DEFT	DEFF	DEFT	DEFF	DEFT	DEFF	DEFT
Birth weight								
Very low birth weight	1.6950	1.3019	3.3670	1.8346	1.5160	1.2312	1.5680	1.2522
Moderately low birth weight	1.4060	1.1857	1.3845	1.1766	1.5595	1.2488	1.6110	1.2692
Normal birth weight	1.2700	1.1269	1.6020	1.2651	1.3390	1.1572	1.3760	1.1730
Region								
Northeast	3.1090	1.7632	3.9090	1.9771	2.0940	1.4470	2.1940	1.4809
Midwest	2.0710	1.4391	4.6190	2.1492	1.8675	1.3663	1.8655	1.3657
South	2.5070	1.5834	4.4820	2.1171	2.0740	1.4401	2.0400	1.4283
West	2.3110	1.5202	3.2810	1.8114	2.1060	1.4512	2.0800	1.4422
Poverty level								
Below poverty threshold	1.8435	1.3577	2.2160	1.4886	1.9695	1.4033	1.9730	1.4046
At or above poverty threshold	2.1100	1.4526	3.4860	1.8671	1.9105	1.3822	1.9240	1.3871

SOURCE: U.S. Department of Education, National Center for Education Statistics, Early Childhood Longitudinal Study, Birth Cohort (ECLS-B), Nine-Month Data Collection, 2001–02.

REFERENCES

Deming, W.E., and Stephan, F.F. (1940). On a Least Squares Adjustment of a Sampled Frequency Table When the Expected Marginal Totals are Known. *Annals of Mathematical Statistics,* 11, 427-444.

Durbin, J. (1967). Design of multi-stage surveys for the estimation of sampling errors. *Applied Statistics*, 16, 152-164.

Federal Interagency Forum on Child and Family Statistics. (1999). *America' Children: Key National Indicators of Well-Being*. Washington, DC: U.S. Government Printing Office.

Flores-Cervantes, I., Brick, J.M., Jones, M.E., DiSogra, C. and Yen, W. (2002). Weighting for Nontelephone Households in the 2001 California Health Interview Survey. *Proceedings of the Section on Survey Research Methods*, *American Statistical Association* (pp. 1002–1007). Alexandria, VA: American Statistical Association. (Available for purchase on CD: http://www.amstat.org/publications.)

Folsom, R.E., Potter, F. J., and Williams, S.R. (1987). Notes on a Composite Site Measure for Self-Weighting Samples in Multiple Domains. In *Proceedings of the American Statistical Association*, Section on Survey Research Methods (pp. 792-96). Alexandria, VA: American Statistical Association. (Available: http://www.amstat.org/sections/srms/proceedings/papers/1987_141.pdf)

Kass, G. (1980). An Exploratory Technique for Investigating Large Quantities of Categorical Data. *Applied Statistics,* 29, 119-127.

Levine, D.B. and Bryant, E.C. (1997). *An Examination of Alternative Approaches to Selecting a Sample of New Births*. (Prepared under contract to the National Center for Education Statistics). Rockville, MD: Westat, Inc.

Magidson, J. (1993). SPSS for Windows: CHAID, Release 6.0. Statistical Innovations Inc.

Mahalanobis, P.C. (1975). On the Generalized Distance in Statistics. In Multivariate Statistical Methods: Among-Groups Covariation; Within-Group Covariation (a collection of reprints), 124-130.

U.S. Department of Education, National Center for Education Statistics. (Forthcoming [a]). *Early Childhood Longitudinal Study, Birth Cohort (ECLS-B), Methodology Report for the Nine-Month Data Collection, Volume 2: Sampling.* Washington, DC: U.S. Department of Education.

U.S. Department of Education, National Center for Education Statistics. (Forthcoming [b]). *Early Childhood Longitudinal Study, Kindergarten Class of 1998–99 (ECLS-K), Third Grade Methodology Report* (NCES 2005–-18). By Karen Tourangeau et al. Elvira Germino Hausken, project officer. Washington, DC: U.S. Department of Education.

U.S. Department of Health and Human Services, National Center for Health Statistics. (1991). *Health Service Areas for the United States*, Series 2, No. 112. Hyattsville, MD: National Center for Health Statistics.

U.S. Department of Health and Human Services, National Center for Health Statistics (1997). *U.S. Vital Statistics System* (DHHS Publication No. (PHS) 97-1003). By Alice M. Hetzel. Hyattsville, MD: National Center for Health Statistics.

This page intentionally left blank.

5. DATA COLLECTION METHODS AND RESPONSE RATES

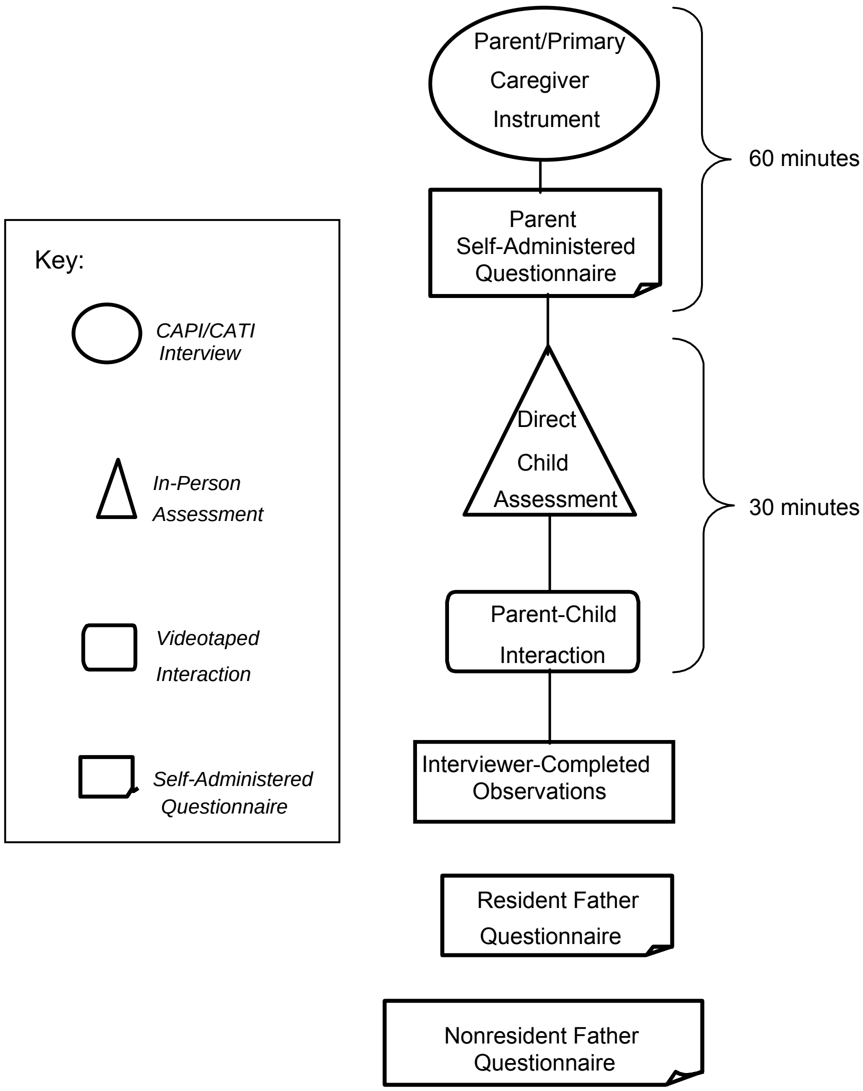
The following sections discuss the procedures used in the 9-month wave of data collection of the Early Childhood Longitudinal Study, Birth Cohort (ECLS-B). Section 5.1 summarizes the methods employed in collecting data from the parent, the child, and the father. Section 5.2 provides an overview of the field staff training procedures that were used to prepare interviewers and field supervisors for the data collection task and describes the measures taken to certify that data collection staff could successfully collect the data. Section 5.3 describes the data collection procedures themselves, beginning with an advance letter to parents and continuing with the initial household contact; it presents information on conducting the parent interview, the child assessments, and the father questionnaires. Section 5.3 also describes the protocol for collecting data from households with limited English proficiency and presents information on the Child Observations and Interviewer Remarks. Section 5.4 presents the response rates for the parent interview, the child assessments, and the father questionnaires. Section 5.5 describes the quality assurance procedures for the data collection task.

5.1 Data Collection Methods

The 9-month wave of data collection began in October of 2001 and continued through December of 2002. Figure 5-1 shows the different methods used to collect the data. Computer-assisted personal interviewing (CAPI) was the principal mode of data collection for the parent interview; self-administered questionnaires were used for gathering information from the spouse or partner of the parent interview respondent (i.e., resident father questionnaire) and for asking sensitive items of the parent respondent. Data were also collected from the child by several means: a series of structured, standardized activities were scored by the field interviewer; a structured interaction with the parent was videotaped for coding; physical measurements (e.g., length and weight) were obtained; and behavior was observed throughout the home visit.

The ECLS-B sample included about 800 twin pairs. Data collection methods were tailored somewhat for twins. Parents of twins were asked to participate in only one twin interview; the sequence of questions was identical to the sequence for parents of singletons and other multiple births. General home and family items were asked only once. Questions specific to the child were asked twice—once for each child (e.g., Child Development section of the parent CAPI instrument). The parent CAPI instrument went through and asked about one of the twins, and then at the end of the sequence of questions that was common to all parents, the instrument looped back to the specific child questions about the other twin. Only one parent self-administered questionnaire was completed. Separate assessments were conducted for the two twins. The father was asked to complete only one questionnaire for twins.

Figure 5-1. Data collection components for wave 1: 2001–02



SOURCE: U.S. Department of Education, National Center for Education Statistics, Early Childhood Longitudinal Study, Birth Cohort (ECLS-B), Nine-Month Data Collection, 2001–02.

5.2 Field Staff and Field Staff Training

For data collection management, the ECLS-B primary sampling units (PSUs) were distributed among 13 regions. Each region was staffed by one field supervisor and between 14 and 19 interviewers, for a total of 256 field staff who were ever active across the length of the field period. Interviewers were responsible for making in-person contact with the sampled cases and completing all data collection activities associated with the cases assigned to them.[1]

In-person training was conducted to prepare field staff for the 9-month data collection tasks. Two different types of training sessions were conducted, one for field supervisors and lead trainers and one for interviewers. Interviewers were trained in two sessions.

The training session for field supervisors and lead trainers was held from August 26 to August 30, 2001, in Rockville, Maryland; 24 individuals were trained. Most of the field supervisors selected for the study had prior experience as field supervisors with Westat, experience on the ECLS-B field test, or experience as field supervisors with another major survey research company. Lead trainers were drawn from the contractor's child development or survey operations staff.

The first interviewer training session was held from October 10 through 17, 2001, in Los Angeles, California; 202 interviewers were trained. A second session was held the week of December 3, 2001, to add more interviewers to the field staff; 41 were trained in the second session in Rockville, Maryland. The focus in hiring interviewers was on those who had worked on the ECLS-B field test or had some other experience working with very young children. This "other" experience included working on studies that involved working with young children (such as work on the ECLS-K study) or work in the field of early child development. The interviewers hired had a mix of experience ranging from prior work on the ECLS-B, education in early childhood development, and/or prior interviewing experience to no background in early childhood development (other than as a parent or grandparent) or in household interviewing.

Before the in-person training, all field staff received 4 hours of home study training. The home study materials covered the study design and procedures for making initial contact, how to conduct a home visit, documenting work, and other study protocol. The training included reading the field

[1] The staff who collected the data are referred to as interviewers, because the activity that required the most time in the home visit was the interview with the parent. However, the reader should keep in mind that the home visit included a number of activities that are not commonly included in an interviewer's tasks (e.g., assessing the child's mental and motor abilities and videotaping a parent-child interaction).

representative manual, viewing a videotape that introduced the home visit, and completing a short-answer exercise on materials covered in the manual and on the video. The home study also included a self-paced CAPI tutorial on using the computer and explaining standard features of conducting a CAPI interview.

5.2.1 Field Supervisor Training

The field supervisor training had several major goals:

1. Develop expertise in ECLS-B interviewing tasks among the field supervisors;
2. Develop supervisor skills in using field management tools; and
3. Build content knowledge and training skills for trainers.

Field supervisor training preceded the interviewer training and lasted for 4 days. The training covered field procedures and protocol as well as using the Interviewer Management System (IMS) and the Field Management System (FMS). The IMS, which resided on the interviewer’s laptop, was used by the interviewers to manage their assigned casework and to collect data about the sampled child and his or her family. The FMS was used throughout the data collection period to transmit data from the interviewer’s laptop computer to the contractor’s home office; to send data (including case assignments) from the home office to the laptop; to manage casework and workload of interviewers; and to monitor production on all data collection activities. During training, supervisors were given hands-on experience in using both systems.

In addition to training on using the IMS and the FMS, topics covered in the field supervisor training session included a review of materials and protocol for making initial contact with the parent; interactive lectures that covered the content and administration of the parent CAPI instrument, the hard-copy parent and father questionnaires, and the Child Observations and Interviewer Remarks; lectures and role plays on conducting the child assessments; and review of administrative materials and procedures, including conducting quality control visits of all interviewers. Supervisors were trained and certified on all activities performed by field interviewers during data collection in the home.

5.2.2 Field Interviewer Training

The major goals of field interviewer training were as follows:

1. Develop skills in gaining cooperation and provide field interviewers with confidence in answering respondent questions about all aspects of the study.

2. Develop skills in conducting the CAPI parent instrument and administering the child assessments.

3. Acquire mastery in applying the study protocol and all administrative procedures associated with the study and field work.

There were 7 days of classroom training, spanning an 8-day period. There were 13 training classrooms each comprising between 18 and 24 field interviewers, two trainers, and at least one support staff. The training approach included interactive lectures; practice in pairs using role-play scripts; practice with children recruited specifically for the training; and videotapes and written material on techniques and strategies for working with the 9-month-old child. The various training activities provided trainees with hands-on experience with all data collection tasks required by the study as well as experience using the IMS and the CAPI programs prior to actual data collection.

Time allocated across training topics is shown in exhibit 5-1:

Exhibit 5-1. Training topics and time allocations in the 9-month data collection: 2001–02

Training topic	Training hours
Total	56.0
Introduction to the ECLS-B	0.5
Making contact	4.5
Parent interview (includes parent CAPI instrument and parent self-administered questionnaire)	10.5
Father questionnaires	6.0
Child assessments (Total)	23.0
Bayley Short Form–Research Edition (BSF-R)	14.5
Nursing Child Assessment Teaching Scale (NCATS)	3.0
Physical measurements	3.5
Child Observations	2.0
Quality control	2.0
Reporting	3.0
Administration/assignments	6.5

SOURCE: U.S. Department of Education, National Center for Education Statistics, Early Childhood Longitudinal Study, Birth Cohort (ECLS-B), Nine-Month Data Collection, 2001–02.

Spanish bilingual staff received one additional day of training, immediately following training of all staff. This training focused on the aspects of data collection that were affected as a result of conducting the home visit in a language other than English. Staff bilingual in Asian languages received similar training by telephone. Aspects of data collection that were discussed and the allocation of training time for each appears in exhibit 5-2.

Exhibit 5-2. Additional training topics and time allocation for bilingual training session in the 9-month data collection: 2001–02

Training topic	Hours allocated
Total	8.0
Welcome/introduction	0.25
Making contact	1.75
Parent interview	3.0
Father questionnaires	1.0
Child assessments (Total)	1.0
Bayley Short Form–Research Edition (BSF-R)	0.50
Nursing Child Assessment Teaching Scale (NCATS)	0.25
Physical measurements	0.25

SOURCE: U.S. Department of Education, National Center for Education Statistics, Early Childhood Longitudinal Study, Birth Cohort (ECLS-B), Nine-Month Data Collection, 2001–02.

5.2.2.1 Training on the Parent CAPI Instrument and Other Nonassessment Instruments

The trainees used laptop computers configured with cases that were part of a training database. The laptops were used to select training cases and complete all of the components associated with data collection, including those components that were a part of the home visit (e.g., parent CAPI instrument) and those that were to be completed after the home visit (e.g., Child Observations and the Interviewer Remarks).

Training on the parent interview utilized an interactive lecture format to introduce the parent CAPI instrument, make specific points about questionnaire content, and provide an opportunity for each interviewer to practice with the group. The trainer worked with a script of answers and played the role of the respondent in a practice case. Interviewers took turns asking the questions and entering the responses. Frequently the trainer used the practice case to make a scripted training point. The first interactive lecture was followed by practice in pairs, with one interviewer playing the role of a respondent following a script and the other interviewer asking the questions and entering the responses into the laptop. Then the interviewers reversed roles. More complex cases were introduced in subsequent interactive lectures, followed by more challenging role-play experiences in pairs. This training approach was also used for the other nonassessment instruments (such as the parent self-administered questionnaire and the resident father questionnaire).

There was no formal certification process for the parent interview *per se*. Each trainee was observed in a number of training contexts for the parent interview throughout the training. The most difficult skills for most interviewers were related to CAPI functions, and almost all interviewers found CAPI easier than the child assessments to learn. From the first day, interviewers who were observed having difficulty achieving mastery of parent interview skills were routed into tutorial sessions, and all were judged capable of administering the parent interview well by the end of classroom training.

5.2.2.2 Training on the Assessment Instruments

A major goal of the interview training session was to train field staff in the proper procedures to conduct the child assessments, including the Bayley Short Form–Research Edition (BSF-R), the Nursing Child Assessment Teaching Scale (NCATS), and the physical measurements. Training on the child assessments began on the first day with a general overview of the three instruments, the purpose of the assessments, and the materials required. The requirements for certification on the assessments were also presented during this introduction. The three different child assessment instruments were introduced in more detail over the next two days as follows.

5.2.2.2.1 Training on the Bayley Short Form–Research Edition

As shown in exhibit 5-1, a considerable amount of training time was devoted to the BSF-R. To achieve consistency in all the classrooms, the training approach relied heavily on a series of ECLS-B videotapes created by the contractor’s child development staff. The BSF-R training began with a detailed introduction that included viewing an introductory videotape of a complete administration conducted by a child development staff member under conditions similar to what field staff would encounter during a home visit. This introduction emphasized that the BSF-R is a standardized assessment and that interviewers needed to adhere closely to the instructions. The required setting for the BSF-R administration (i.e., a quiet, well-lighted, well-ventilated space with the child seated at a table directly facing the interviewer) and the application of the basal and ceiling rules (see chapter 2, section 2.1.1 for a discussion of these rules) rounded out the contents of this introduction.

The foundation for training on the BSF-R was a second training videotape. This second, detailed videotape included the complete instructions for administering and scoring all the mental and motor items of the BSF-R. Supplementary information on the tape included the purpose(s) of an item as well as troubleshooting tips, for example, what to do if the child responded in some unexpected way on an item. This videotape also explained the basal and ceiling rules in detail.

The detailed training videotape had several purposes. The first was to ensure that all trainees received the same comprehensive information about the BSF-R. The second purpose was to streamline and standardize the training of the lead trainers who were not from the child development group; the BSF-R is a complex instrument, and standardizing the instruction in this way ensured that the trainers also received the same information. A third purpose was to support additional opportunities for practice, outside the classroom. All trainees received a copy of the videotape to take home with them so that they could review the administration and scoring of the items, as needed, to improve their administration.

The detailed videotape was presented in two parts over a 2-day period beginning on the second day of training. On the first day, the instructions for administering and scoring the mental scale items, including the core, basal, and ceiling items, were presented over a 2-hour period. This also included instruction about how to determine whether the basal set of items or the ceiling set of items should be administered to the child. Following the videotape presentation of the mental items, all trainees then participated in “directed practice” that involved following the trainer’s instructions for administering and scoring all the mental scale items. This directed practice is similar to role plays in which trainee pairs practice administering an instrument, one trainee taking the role of respondent and the other the role of interviewer. Training on the vocalization and language items of the BSF-R proceeded similarly. A separate (third) videotape of children’s vocalizations was presented. Trainees then practiced saying the different types of vocalizations to learn to recognize the sounds and then practiced scoring the sounds according to the BSF-R criteria.

The motor scale items on the second, detailed videotape were presented the next day over a 1½ hour session (shorter than the mental scale because there are fewer motor scale items in the BSF-R). This was followed by a 1-hour directed practice session in which trainee pairs administered the motor items to a cloth doll following the instructions of the trainer.

Trainees then viewed the first introductory videotape again. During this viewing, trainees completed a BSF-R review form in which they scored the child’s performance on each item and rated how well the tester adhered to the administration instructions for the items. This gave trainees further practice scoring BSF-R items and practice with the administration instructions. This also exposed trainees to the same review form that would be used by the contractor’s staff and field supervisors during quality control visits in the field. This practice session was followed by a precertification quiz (see section 5.2.3 for more details) in which another videotape (the fourth tape used in the training approach) of a complete BSF-R was presented. Using the same BSF-R review form, trainees scored the child’s performance on the items and scored the tester’s adherence to the BSF-R administration instructions for each item. This precertification quiz served as a standardized tool for identifying trainees who needed to improve understanding of administration instructions or scoring before advancing to the live practice session.

Establishing rapport is a critical requirement for obtaining a BSF-R that accurately measures a child’s ability. Therefore, trainees were given tips about how to establish and maintain rapport with an infant, how to monitor the child’s gaze to help them adjust the pace of item administration, and how to adjust the order of items to keep the child interested and move through the BSF-R efficiently. During this instruction, trainees were also taught how to modify any item administration in order to provide children with disabilities an equal opportunity to perform.

As an adjunct to this formal training, trainees were also encouraged to work in pairs or small groups in their rooms at night to practice BSF-R administration and scoring or, alternatively, to attend nightly help labs conducted by the contractor's child development staff members. These nightly help labs were exceptionally well attended. During these help labs, trainers reviewed the BSF-R item by item, and trainees practiced administering and scoring the items as well as administering the basal and ceiling rules.

5.2.2.2.2 Training on the Nursing Child Assessment Teaching Scale

The NCATS was introduced on the fourth day of training in a 2-hour training session. During the first hour, the administration of the NCATS activity was emphasized, including the need to obtain a high quality videotape of the interaction. During this time, trainees viewed a videotape that included examples of a high quality videotape and a low quality videotape. In the second hour, trainees were introduced to the camera, learned how to operate the camera, and then practiced in pairs to administer and videotape the NCATS to each other. Because the videotapes would be coded by trained coders upon receipt at the home office, the focus of this training was on the proper administration of the NCATS task and quality videotaping of the interaction. The forms that would be used to review the quality of the tapes were also discussed so that interviewers would be familiar with the methods that would be used to review their performance in the field as well as reinforce the different elements of a high quality NCATS taping.

5.2.2.2.3 Training on the Physical Measurements

Training on the physical measurements included understanding the purpose of the measurements and becoming familiar with the equipment. The equipment included the SECA scale, the SECA measuring tape (for obtaining the middle upper arm circumference, or MUAC, and, when necessary for very low birth weight babies, the head circumference), and the SECA MeasureMat (a device for measuring child's recumbent length). The trainer demonstrated each measurement, using oversize dolls and other training staff. Hands-on practice was an important part of this training; each trainee practiced weighing a partner and measuring the partner's MUAC and head circumference. (Practice measuring recumbent length was not feasible due to the obvious limitations of the MeasureMat).

5.2.2.3 “Live Practice” of the Child Assessments

On the fifth day of training, all trainees had the opportunity to administer all the child assessments to infants during a live practice session. Each trainee administered the child assessments to an infant.[2] Trainees were instructed to begin the parent CAPI instrument with the volunteer and then break off to begin the child assessments. The child assessments were then administered in the order of the BSF-R, the NCATS, and the physical measurements.

Trainees worked in pairs, with each partner taking turns administering all the child assessments to a child while the other partner videotaped the administration. After the first partner completed an administration of the child assessments, trainees switched roles. A new child was brought in, and the second partner then administered all the child assessments while the first partner videotaped this administration. Thus, trainees gained experience administering the assessments to children in a way that was very similar to what was expected on a home visit. Furthermore, in addition to providing trainees with live practice, the session also served as the means by which trainees were certified on their administrations. The sections that follow describe in greater detail the certification procedures that were undertaken to ensure that field staff administered the child assessments competently and to required standards in the field.

5.2.3 Certification on the Bayley Short Form–Research Edition

As described in chapter 3, the BSF-R is derived from the Bayley Scales of Infant Development, Second Edition (BSID-II). The BSID-II is a comprehensive, standardized measurement that is used to assess children’s developmental status from birth through 42 months of age. The BSF-R, as described in chapter 2, was designed to measure the developmental status of children ranging in age from approximately 8 months to 11 months of age and to be administered by lay interviewers in a home-based setting. As a standardized assessment, the BSID-II, and therefore the BSF-R, requires that the administrator follow clear instructions when administering an item and follow clear criteria when scoring the child’s performance on an item. It also requires that the administrator establish and maintain good rapport with the child in order to elicit the best performance possible from the child. Maintaining good

[2] For the interviewer training in October, two market research firms (Savitz Field and Focus and Ask Southern California) recruited 300 infants. For the interviewer training in December, the contractor’s home office staff recruited 57 infants. For both trainings, each infant’s mother was paid $50.

rapport requires that the interviewer be able to adjust her pace to match the child's ability to receive the instructions and for the interviewer to monitor the child's positive or negative mood to keep the child motivated. In order to monitor the child's performance and mood in this way, interviewers must be in command of the BSF-R item administration procedures.

For these reasons, it was important to ensure that, before collecting data from ECLS-B cases in the field, trainees were able to administer the BSF-R to matching standards for administration and scoring. Standards were developed by the contractor's child development experts, with guidance from external reviewers. A three-level certification component was incorporated into training, beginning with in-class exercises, progressing to a written precertification exam, and culminating in the complete administration of the BSF-R during a "live" practice session with infants. Certification on the BSF-R during the live practice session involved evaluation of the trainees' ability to administer the items according to the standardized instructions, knowledge of each item's scoring criteria, and ability to establish rapport and to interpret the children's responses.

There were several in-class written quizzes beginning when the BSF-R was first introduced. Beginning on the second day of training, when the mental items of the BSF-R were first introduced on the BSF-R training videotape, trainees were quizzed on the scoring of all items, including core, basal, and ceiling items. In addition, trainees were introduced to the vocalization items of the mental scale, practiced saying a wide range of vowel sounds to learn to recognize them, and then completed a quiz on scoring the vocalization items from a videotape. The correct answers for all the quiz items were reviewed immediately after the quiz so that trainees could learn from their mistakes. The same quiz procedure was followed on the third day of training (which was the second day of BSF-R training) when the motor items of the BSF-R were presented on the BSF-R training videotape. The purpose of these quizzes was not to assign grades but to identify individuals who were having trouble scoring the BSF-R items from the videotape. Those identified were then required to attend help labs in the evenings to improve their understanding of the scoring criteria.

On the third day of training, after 2 days of direct BSF-R instruction, modeling, and directed practice role plays, trainees completed a practice exam in preparation for a precertification written exam, which immediately followed the practice exam. The precertification exam included a videotape presentation of a complete BSF-R administration, with core, basal, and ceiling items. Two videotapes showed two people, each administering the core, basal, and ceiling items of the BSF-R, to a number of children. Trainees, as a group, viewed the videotapes item by item and recorded the scores (credit or no credit) they would give for the child's performance on each item. Using a standard BSF-R review form, they also scored the accuracy of the administration of the individual administering the BSF-R on the videotape. The following sample item (exhibit 5-3), which is similar to those on the BSF-R review form, shows that the administration instructions were included so that trainees could compare what the administrator did with the standard instructions, and the scoring instructions were included so trainees could assign credit or no credit for the child.

Exhibit 5-3. Sample item from the BSF-R review form, 9-month data collection: 2001–02

3. Uses Means-End Behavior to Retrieve Object

Administration:

During this item, does **administrator**			
1. If using tray table, turn it lengthwise to child?	❑YES	❑NO	❑NA
2. Suspend object and swing 8-10 from child s face at eye level?	❑YES	❑NO	
3. Place object out of reach; string toward child?	❑YES	❑NO	
4. Make any other errors?____			

Scoring:

During this item, does **child**		
Play with string (doesn t need to grab object)? (Basal Item)......................	❑YES	❑NO
Bang object in play? (Basal Item)..	❑YES	❑NO

Score box: Record C/NC in box.	
Uses Means-End Behavior to Grab Object 3.	

SOURCE: U.S. Department of Education, National Center for Education Statistics, Early Childhood Longitudinal Study, Birth Cohort (ECLS-B), Nine-Month Data Collection, 2001–02.

The practice exam included time for trainees to review their own scoring and receive clarification on any item. The answers for each item were reviewed by the group, and questions were discussed. After completing the practice exam, each trainee completed the precertification exam. Although the videotape was paused between items to give trainees time to complete the item, there was no discussion of the answers, and no questions were permitted. Again the trainees scored the child according to the performance in the video (credit/no credit) and critiqued the administration of each item. In order to pass the precertification exam, each trainee had to receive 90 percent or higher on scoring the items and 85 percent or higher on review of the administration. The higher criterion for scoring accuracy was imposed because scoring errors can result in misapplication of the basal or ceiling rule and therefore the loss of data.

Any trainee who did not pass either of these precertification criteria was required to attend a mandatory help lab to improve her or his understanding of the administration instructions or the scoring rules before being able to advance to the live practice session. This precertification exam identified 27 out of 243 trainees (11 percent) who were required to attend the help lab. In practice, however, almost half of all trainees voluntarily attended all or some help labs. On average, trainees received scores of 91 percent for critique of administration and 87 percent for accuracy of scoring on the precertification exam.

The certification process culminated in the live practice session in which each trainee administered the BSF-R to a child who ranged in age from about 8 to 12 months, the approximate age range trainees would encounter during the ECLS-B data collection. (Although the BSF-R was designed to measure the status of children ranging in ages from approximately 8 to 11 months, most 12-month-olds were more than adequate subjects for certification purposes.) While one trainee administered the BSF-R, the other trainee videotaped the assessment. To ensure that all trainees were obtaining a clearly visible and audible videotape of the BSF-R administration, trainers and technical support staff circulated around the rooms and viewed trainees' camera display screens to make sure that the BSF-R was framed properly.

After the live practice sessions, all BSF-R videotapes were reviewed by the ECLS-B training staff. The qualifications of the training staff to judge trainees' BSF-R administration had been established prior to training. To certify the training staff, four core child development staff members, who had graduate training in child development, early childhood education, or another social science and who had been directly involved in the creation of the BSF-R, were each videotaped while administering the BSF-R to a child. These videotapes were then sent to an expert consultant on the BSID-II. Videotapes were critiqued for administration of items and establishing good rapport with the children. These core child development staff members then trained the other lead trainers in one-on-one tutorials followed by practice with child volunteers. After these trainers had sufficient practice with child volunteers and were confident that they knew the BSF-R administration, they administered a complete BSF-R to a new child and were live-coded by one of the child development staff members using the standard BSF-R review form for certification. All trainers passed this certification exam with a minimum of 90 percent for administration accuracy and 90 percent for scoring accuracy.

After the live practice session, this ECLS-B training staff used the same BSF-R review form that was used during the precertification exam to review each trainee's BSF-R videotape for accuracy of administration, accuracy of scoring the child's responses, and the trainee's ability to maintain rapport. In order to be certified on administration of the BSF-R, each trainee had to earn 90 percent or higher on scoring the child’s responses and 85 percent or higher on accuracy of administration. Trainees who did not pass certification during live practice were required to attend a special help lab where they were given feedback on their performance and guidance for improvement. They were then required to complete a second live practice administration of the BSF-R with another child. There were 21 trainees who were required to do a second live practice, 13 of whom then went on to pass their certification. Five trainees who were not able to achieve certification were released from the study, and three trainees who refused a second practice were also released. Including both training sessions (i.e., October and December), trainees received an average score of 93 percent for BSF-R administration accuracy and an average score of 92 percent for scoring accuracy on the live practice certification. Seven percent of the staff who attended training failed to meet the study standards for the BSF-R and were released.

5.2.4 Interviewer Certification on the Nursing Child Assessment Teaching Scale

The NCATS is a semi-structured parent-child interaction in which the parent is instructed to teach her child a new and challenging task. The parent is given a standard list of age-appropriate activities and told to select one that the child is not yet able to do, such as pretending to drink from a cup. (See sections 2.1.2 and 3.1.2 for more details about the NCATS.) Interviewers were trained to administer the activity the same way to all parents. To ensure standardization of administration, all the NCATS administration instructions, including a script to be read to the parent, were presented in unambiguous detail in the NCATS section of the Child Activity Booklet. The interviewer only needed to follow the step-by-step instructions in the booklet to administer the NCATS accurately. Role plays were conducted during the training to familiarize the interviewers with the NCATS. In groups of three, trainees each took turns administering the NCATS, playing the role of the parent and then playing the role of the child. As trainees participated in the role plays, trainers and all available technical support staff monitored the groups to intervene and correct any problems and to recommend attendance at help labs when necessary.

The NCATS is usually coded *in vivo* by individuals trained on this coding system. For the ECLS-B, however, interviewers were instructed to obtain a videotape of this interaction because live-coding of the interaction within the context of the home visit would be too burdensome for them to do reliably. The NCATS tapes were later coded by trained coders at Westat using the NCATS coding system (see section 5.6). For details about the NCATS and its score, see chapters 2 and 3, respectively.

The main emphases of NCATS certification at training were the following:

- Correct administration of the NCATS;
- Proficient use of the video camera; and
- Production of a high quality videotape.

To ensure that these skills were achieved, trainees received hands-on review of each of these skills during the live practice. During the ECLS-B training, interviewers were trained to introduce the NCATS task to the parent, assist the parent to select an appropriate teaching task, and end the task appropriately. To end the task, interviewers were trained to instruct the parent twice to “tell me when you are done”; record until the parent says something like “We are done”; if five minutes had elapsed and the parent had not yet indicated completion they were to ask “Are you finished yet or do you want to continue?” If the parent chose to continue, the interviewer would tape for 30 more seconds and then say “That’s good. Thank you very much.” After the videotape session was over, the interviewer would rewind the tape and play it back to make sure the camera was recording.

Since the NCATS was not live-coded, high videotape quality was essential for reliable coding. Field staff were trained and certified on videotaping techniques that would produce codeable tapes (e.g., adequate lighting and sound, seeing the parent and child’s facial expressions, hearing vocalizations). Certification procedures for the NCATS therefore focused on obtaining a codeable videotape of the NCATS interaction.

Training on correct procedures for videotaping the NCATS interaction included viewing NCATS training films that showed the interactions between a parent and child during the NCATS teaching task. The films also provided examples of good and bad camera angles, framing, and visual shots for capturing verbal and nonverbal behavior to enable coding the NCATS. During the live practice, each interviewer administered the NCATS task and videotaped a parent and child engaged in the activity. Trainers and technical support staff observed this administration, paying particular attention to the image being captured in the camera’s viewfinder. Hands-on quality control was done to address issues such as inappropriate framing or camera angle during the live practice. Trainers and support staff intervened as appropriate to make corrections and to make sure that the trainee was videotaping appropriately. Recommendations for mandatory help labs on the NCATS (both administration and videotape skills) were made for those trainees showing problems that would compromise NCATS coding.

In addition to hands-on review of their videotaping skills during live practice, interviewers were precertified on the basis of the quality of their videotapes of their partner's BSF-R administration. (As noted earlier, each trainee videotaped his/her partner administering the BSF-R during live practice. See section 5.2.3.) A final certification of NCATS videotape quality during training was not possible because trainers and technical support staff were rigorously monitoring BSF-R videotape quality to ensure that all BSF-R administrations could be clearly evaluated. In addition, NCATS videotape certification would require the partners to switch the camera, insert a new videotape, and remember to put on a new label with the NCATS filmer's name. This seemed to introduce too much potential for error in the context of the scale of the ECLS-B training. Therefore, it was decided to concentrate on final certification after the first NCATS tapes were received from data collection in the field.

Final certification of videotaping skills took place when NCATS tapes were receipted at the home office and were quality checked by trained NCATS coders using the NCATS Videotape Quality Control Checklist. This checklist consists of 11 aspects of quality that are critical to the success of the NCATS coding, such as quality of lighting, sound level, and camera position and angle. The first three NCATS videotapes that each interviewer sent in were coded and reviewed for quality immediately upon receipt to assess a final certification and to make sure that any potential errors would not be duplicated on incoming tapes. If, after three videotapes, an interviewer did not achieve 85 percent on the checklist scores, a discussion was held with the contractor's project management and the individual's field supervisor to determine a further course of action, such as additional home study or additional training on the NCATS in the contractor's office. No trainees were released because of NCATS administration problems. Of the 9,439 videotapes processed in the coding activity, 731 (8 percent) were uncodeable. The uncodeability rate was higher during the early months of the coding effort.

Immediate feedback on an ongoing basis was given to field staff to correct videotaping techniques, as required. Problems reported on the Videotape Quality Control Checklist covered a number of items, some more critical to codeability of the tape than others. For example, the Quality Control score of field interviewers was marked down if he or she failed to provide an appropriate date stamp or failed to note the language used during the NCATS administration, though neither had impact on the codeability of the tape. However, field interviewers were also given feedback on items critical for coding the tapes such as improper camera position (e.g., shaky or not properly framing the child and parent) or inadequate volume level. Regardless of the nature of the problem, field interviewers received complete and thorough feedback. Over the course of the field period, feedback to correct videotaping techniques was given to about 20 percent of the field interviewers. In a typical month (selected from the middle of the data collection so as to be representative), videotape quality control scores averaged 85 percent across all interviewers (range 35 to 100 percent).

The videotape quality review also included the appropriateness of the chosen NCATS activity, capturing the introduction of the materials by the mother to the child, ending the task appropriately, and not interrupting the interaction. Feedback about administration errors was given to field supervisors on a weekly basis, who then followed up with the interviewers.

5.2.5 Certification on the Physical Measurements

Training on the physical measurements involved hands-on practice during which trainees worked in triads to role play obtaining physical measurements from the mother and child. The triad included one interviewer obtaining measurements, one interviewer being measured, and a third interviewer observing the interaction. During that role play and others, trainers and technical support staff monitored trainees' administrations of the physical measurements, correcting them as necessary. See chapters 2 and 3 of this manual for further information about the procedures and characteristics of the physical measurements.

The physical measurements were the final activity during the live practice session and included the child's weight, the mother's weight, the child's length, the child's middle upper arm circumference (MUAC—this measurement includes the upper arm length, the midpoint of the upper arm length, and the circumference around this midpoint), and the child's head circumference (for very low birth weight babies only). As standard procedure, all measurements were obtained twice. Certification was defined as a less than 5 percent difference between the two sets of measurements. If any measurement had more than a 5 percent difference between the two recordings, the trainee was required to demonstrate to the trainer that she or he knew how to obtain the measurement. The purpose of this certification procedure, therefore, was not to quantify how many trainees were "certified" but rather to catch those with large discrepancies while they were still at training, identify the problem, retrain them on the problem measurement, and require them to demonstrate that they knew how to administer and record the measurement properly. All trainees were certified on the physical measurements before leaving training. No trainees were released because of errors in the physical measurements procedures.

5.3 Nine-Month Data Collection Procedures

This section details the procedures employed to administer the instruments of the 9-month data collection. Procedures for conducting the data collection included advance mailing to the parents of sampled children; making initial contact; conducting the parent interview; administering the child assessments; presenting and prompting for the resident and nonresident father questionnaires; recording Child Observations; and completing field Interviewer Remarks. The first two procedures established contact with the family. The parent, child, and father instruments were a part of an in-home visit. The final two procedures were completed by field interviewers after the home visit.

5.3.1 Advance Mailings

Beginning in mid-September 2001, advance packages were mailed to the parent respondent of each of the sampled children about 3 weeks before the child's 9-month birth date. The advance package included a letter signed by the contractor's project director. The letter described the study and the data collection at 9 months, explained how the child was chosen, and detailed the confidential and voluntary nature of the study. It informed the children's parents that a representative from the contractor would be contacting the family soon.[3] The advance package also included a color brochure with additional information about the study, why it was important, and how it would be conducted. A small incentive, a bib imprinted with the study logo for the infant and a small, colorful board book, was also included in the advance package to encourage participation in the study.

Advance letters and brochures were translated into Spanish and six Asian languages (Hmong, Vietnamese, Japanese, Khmer, Korean, traditional Chinese, and simplified Chinese). All respondents were sent a letter and brochure in English. If the birth certificate indicated that the mother was Hispanic and that she was born outside the U.S., a letter and brochure in Spanish were sent as well. If the birth certificate indicated that the mother was Asian and that she was born outside the U.S., an insert to request a letter and brochure in one of the six Asian languages was sent along with the English materials. (The insert itself included statements in all six languages, to enable the parent to understand the statement and make the request for additional information by checking a box next to the statement.)

---

[3] Special letters were sent to families in three states. Those letters were similar in content to those sent to families in the other states, but featured slightly different wording based on requirements by either the state registrar's office or the state institutional review board.

Advance letters that were returned either with a forwarding address or as undeliverable were entered into a tracking database. Advance letters were re-mailed to the updated addresses, and returned letters having no forwarding address underwent intensive tracing efforts by the home office. These tracing efforts were implemented by the contractor's Telephone Research Center staff who had extensive background in tracing activities. The tracing efforts included Internet searches, contact with direct assistance, use of criss-cross directories, and contact with local institutions that might be sources of address information, such as schools and local post offices. Requests for addresses and telephone numbers for 14,214 cases were sent to a database search firm, and 6,803 (48 percent) were confirmed; 15,263 post office confirmation letters were sent, and 10,783 (71 percent) were confirmed; 3,371 cases were submitted for Internet tracing, and 159 (5 percent) were confirmed; 1,130 were submitted for telephone tracing, and 379 (34 percent) were confirmed. Altogether, 15,509 cases were traced in this pre-field effort, and 87 percent were confirmed.

5.3.2 Initial Contact

Initial contact with the household was usually made in person at the most current address provided. However, if a case was more than 50 miles from his or her home, remote from all other work, and a telephone number was available, the field interviewer called the household first and determined whether anyone would be present on the day and time that the field interviewer planned to travel there. If the field interviewer had difficulty locating the sampled child's family at the address provided, the field interviewer implemented field tracing procedures. Such tracing procedures included going to the local post office and talking with the landlord, neighbors, and current residents.

When the sampled child's family was contacted, first it was determined whether the child was living in the home.[4] Next, the parent respondent was identified. In most instances, the person named as the birth mother was identified as the parent respondent. If the mother did not live in the home, another parent or guardian responsible for the care of the sampled child was identified as the parent respondent. In

[4] The ECLS-B home visit was conducted in the residence of the sampled child. If it was determined that the child was not living at the contacted address, contact information was obtained to locate the child.

99 percent of the cases the parent respondent was the child’s biological mother. Respondents for the parent interview were selected according to the following order of preference:

1. The child’s biological mother.
2. The child’s biological father.
3. Another parent or guardian.
4. Another adult household member.[5]

After identifying the parent respondent, the study was explained. For most of the participating cases, the introduction to the study was all that was needed to gain the cooperation of the parent respondent. However, there were times when respondents were hesitant to cooperate. Extra reassurance was given to those parents by using various other outreach materials such as study brochures. Field interviewers were trained to avert refusals in as many of these cases as possible by clearly explaining the study and answering specific questions from the parent respondent.

In addition to gaining initial cooperation from the parent respondent, signed informed consent was obtained from the respondent before the parent interview began. The consent form was available in English and Spanish. It explained the purpose of the ECLS-B to parent respondents, that their participation was voluntary, and that all information they provided would be kept confidential.[6] Before the interview began, the parent consent form was presented and explained to the parent respondent, and the respondent was asked to read and sign the form. If a signed parent consent form was not obtained, the home visit was not conducted and the case disposition was set to “refused.”

During initial contact, if the field interviewer determined that the respondent did not speak English very well, she or he attempted to determine what language the respondent did speak well. If the respondent spoke Spanish, the interviewer arranged for the home visit to be conducted in that language by a qualified bilingual ECLS-B interviewer. If the respondent spoke some other language (e.g., Chinese or Vietnamese), the interviewer attempted to find someone else in the household who spoke the respondent’s

---

[5] The term “adult” was not defined for interviewers.

[6] Confidentiality was defined for respondents. The following example is from the Washington State consent form: “Only the study team will know your answers to the interview or see information about your baby. All study files will be kept locked. We will assign a code number to you and your family members to protect your privacy. We plan to allow researchers to use a special version of the study information without any names. These researchers must promise to follow very strict confidentiality procedures to protect your privacy. We will not use any names of parents or children in our study reports. There are three exceptions to our promise of confidentiality. You need to be told that we are required by law to report any child abuse or neglect we may see or hear about to DSHS Child Protective Services. We are also required to make a report and to get help if we are concerned that you may hurt yourself or someone else. As required by law, the Attorney General of the United States could get information collected in this study under court order to use to investigate and prosecute acts of terrorism.”

language in addition to English. This person was asked to introduce the field interviewer to the respondent and explain that he or she would like to make an appointment to meet with the respondent and introduce the study through the use of an interpreter. (A small number of interviews were conducted in Mandarin/Cantonese, Vietnamese, Khmer, and Korean by local interviewers who spoke those languages, but these interviewers could not travel to cover most of the Asian language-speaking families in the study. In some distant cases a decision was made for these interviewers to conduct the parent interview by telephone; most of the time when this occurred, no child assessments were conducted.)

5.3.3 Conducting the Home Visit

Exhibit 5-4 shows the preferred order of administration for all of the home visit components. This order was selected to ensure standardization and to allow the child to become familiar with the field interviewer. Interviewers were trained to begin with the parent CAPI instrument. Watching friendly interaction between the field interviewer and the parent respondent enhanced the child's level of ease and fostered the best possible performance by the child during the administration of the child assessments. Interviewers were encouraged to interrupt the parent interview and administer the Child Activity Booklet, once they observed that the child was comfortable in their presence. Optimal conditions varied by case, and sometimes required departures from the preferred protocol for administration order.

Exhibit 5-4. Preferred order of administration of the home visit components, 9-month data collection: 2001–02

Parent questionnaires	Child assessments
	These instruments are all packaged into the Child Activity Booklet.
1. Parent CAPI instrument	2. Bayley Short Form–Research Edition (BSF-R)
5. Resident father questionnaire (if applicable)	3. Nursing Child Assessment Teaching Scale (NCATS)
6. Parent self-administered questionnaire	4. Physical measurements
7. Parent respondent contacting and locating form	
8. Father contacting and locating information (if applicable)	
9. Nonresident father questionnaire (if applicable)	

SOURCE: U.S. Department of Education, National Center for Education Statistics, Early Childhood Longitudinal Study, Birth Cohort (ECLS-B), Nine-Month Data Collection, 2001–02.

5.3.4 Conducting the Parent Interview

The parent interview included two instruments: the parent CAPI instrument and the parent self-administered questionnaire. The parent interview was conducted in person by trained field interviewers using computer-assisted personal interviewing techniques. The parent CAPI instrument consisted of 18 sections covering 22 topics (see chapter 2, exhibit 2-3) that obtained information from the parent respondent about aspects of the child's health, growth, and development and environmental factors that affect children's growth and development (see section 2.2). The parent CAPI instrument took on average about 60 minutes to administer.

Most of the interviews were conducted in English, but provisions were made to interview parents who preferred Spanish or some other language. Thirty bilingual field interviewers were trained and certified to conduct the parent interview in English or Spanish. When the parent respondent preferred Spanish, but the field interviewer was not certified to conduct interviews in Spanish, the case was turned over to the field supervisor who reviewed the case and either assigned it to a certified Spanish-speaking field interviewer or (if none was available) assisted the field interviewer in locating an interpreter to conduct the interview. Most interviews in other languages were conducted with the aid of interpreters (although, as noted above, some were conducted in an Asian language in person or by telephone by ECLS-B interviewers who received special training via telephone). About 9 percent of the parent CAPI instruments were conducted in a language other than English; about 77 percent of the non-English interviews were conducted in Spanish. Less than 0.1 percent of the parent interviews could not be conducted because of language problems (e.g., respondent and family spoke Bengali and no interpreter was found; respondent spoke an unknown Asian language and no interpreter could be found).

In addition to the parent CAPI instrument, parent respondents were given the parent self-administered questionnaire to complete during the home visit. The paper and pencil instrument was presented during the parent CAPI instrument for the respondent to complete and return in a sealed envelope. The parent self-administered questionnaire was short—23 questions on topics some people might prefer to answer privately (see section 2.2.2). The parent self-administered questionnaire was translated into Spanish. If the parent CAPI instrument was conducted in another language (i.e., neither English nor Spanish), then the parent self-administered questionnaire was not given to the parent respondent. Ninety-five percent of the complete parent self-administered questionnaires were in English, and 5 percent were in Spanish. A total of 119 cases were not given a parent self-administered questionnaire because of language issues.

The closing section of the parent interview combined a number of tasks. It began by thanking the parent respondent for his or her participation in the study. Next, the parent respondent was asked to provide updated information about the family's address and telephone number and information about any new address, if applicable, that could be used to locate and contact the family for the next home visit. In addition, respondents were asked to give the name, address, and telephone numbers of persons who would always know their whereabouts. This information was collected using a self-administered form. The section concluded with paying the parent $50 for her or his participation in the study ($100 for twins) and presenting a gift book for the child (a book for each child in twin cases). This occurred after the child assessments were completed.

In order to build response rates for the home visit, parent refusal conversion letters were developed. Letters were sent to all parents who were initially coded as "refusal," except those that were judged to be extremely firm in their decision not to participate, and parents in Washington State.[7] Different versions of the refusal conversion letters were prepared including a general refusal letter and letters more specific to the type of refusal, such as concerns about confidentiality, the government's role, and time constraints. A refusal conversion letter addressed to apartment managers and landlords was also developed to help gain access to respondents in apartment buildings. Spanish versions of these letters were also created to help in refusal conversions. These letters were mailed directly to respondents by field supervisors or field interviewers (who often had a mailing address closer to the respondent's address), usually within several weeks of the respondent's initial refusal.

5.3.5 Conducting the Child Assessments

The child assessments were administered in the home visit along with the parent interview.[8] The ECLS-B field staff arranged to work with the infants at their best time. Home visits were scheduled at times when the child was awake and alert.

---

[7] Washington State's institutional review board requested this protocol change.

[8] Although for the majority of cases the assessments were conducted on the same day as the parent interview, about one percent of the households requested that the interviewers return at a later time to conduct or complete the assessments. Usually the interviewers scheduled the second visit in the same week, but occasionally a longer time period elapsed.

5.3.5.1 Overview of the Child Assessments

In the three-part assessment, the child's developmental status was measured using the BSF-R; a videotape of the NCATS (caregiver-child interaction) was created, and the child's physical measurements were obtained. To facilitate interviewer-child rapport, the recommended order of activities was to begin with the BSF-R, follow with the NCATS, and conclude with the physical measurements.

Information was gathered using hard-copy materials. Interviewers recorded information in the Child Activity Booklet. The Child Activity Booklet also included administration and scoring instructions and record forms. A Spanish version of the Child Activity Booklet was also developed. If the family spoke a language other than English or Spanish, interviewers used an interpreter recruited either through the family or through nearby community agencies or organizations.

To conduct the child assessments in a home setting, interviewers worked with the parent to find a well-lit, quiet setting, away from sources of noise, such as a television or radio, and away from any other distractions, such as the child's toys, family pets, and so forth. The presence of other family members was discouraged whenever possible.

Interviewers conducted the BSF-R with the child in a high chair, tray removed, and pulled up to a kitchen or dining room table whenever possible. If the household did not have available table space, the BSF-R was conducted at a small folding table provided for this reason, with the child sitting in the parent respondent's lap. Interviewers were encouraged to sit facing the child in all circumstances in order to facilitate rapport and for accurate scoring.

For the NCATS, interviewers videotaped the interaction with the parent and child seated on the floor, or at the same table used for the BSF-R. If other siblings, family members, or friends were present, interviewers requested a quiet space with no interruptions. Typically, interviewers brought crayons and coloring books to keep siblings busy and away from the testing area.

The physical measurements were obtained at the end of the home visit because the procedures for obtaining the middle upper arm circumference are intrusive and infants become distressed and try to pull away. In addition, obtaining the child's length with the child recumbent on the MeasureMat also leads children to become distressed and fussy. By completing these at the end of the home visit, the other two child assessments could proceed smoothly.

5.3.5.2 Standardization in Conducting the Child Assessments

Data collection procedures for each of the three components of the child assessment are described below. For more information on the instruments or their scores, see chapters 2 and 3, respectively.

The BSF-R. The BSF-R is a standardized measure of the mental and motor developmental functioning of young children. (For further details about the contents of the BSF-R and how it was developed, see section 3.1.1.) Interviewers worked one-on-one with each child, using standardized toys, verbal prompts, and modeling to administer a variety of tasks. The necessary standardization of administration and scoring was pursued in three ways. First, a training videotape was produced that included item-by-item instructions for administering and scoring the items, which was given to all field interviewers to take home so that they could review the information as needed. Second, item-by-item administration instructions were included in the Child Activity Booklet, along with the scoring instructions, closely based on the BSID-II manual. Additional instructions and clarifications were added to assist interviewers. Third, a videotaped BSF-R administration was sent to interviewers quarterly. The interviewers used a review form to critique the administration and to assign item scores for the child's performance. Interviewers were required to score 85 percent or higher for their critique and 90 percent or higher for scoring. For further information about quality control procedures for the BSF-R in the field, please see section 5.5.4.

The NCATS. The NCATS is a measure for describing characteristics of mother-child interaction, such as how they communicate, their responsiveness to each other, and their adaptability to each other. For the NCATS, the interviewer videotaped a parent-child interaction in which the parent taught the child how to do a new and previously unknown activity. To ensure standardization of administration, step-by-step instructions were included in the NCATS section of the Child Activity Booklet. Furthermore, the interviewer helped the parent choose an activity from a standard list of age-appropriate activities based on a standard set of toys or materials. Once the parent respondent selected the activity, the interviewer then presented the parent with the necessary materials and then videotaped the interaction. Per the NCATS design, interactions typically lasted no longer than 5 minutes and 30 seconds. If an interaction lasted longer than that, then only the first 5 minutes 30 seconds were coded. An interaction was defined as uncodeable if it lasted less than 45 seconds.

The Physical Measurements. The physical measurements component included the child's weight, length, and middle-upper arm circumference (MUAC) (and head circumference for very low birth weight babies). The administration of the physical measurements generally followed the administration of the NCATS, but could be done at any time during the home visit, at the interviewer's discretion. Interviewers used a variety of equipment to conduct the physical measurements. A scale was used to measure the child's weight; first the parent was asked to step on the scale while holding the child; the parent stepped on the scale a second time without the child, and the difference between the two weights was taken as the child's weight. If the parent was unable to perform this task, the interviewer or another adult could perform it. Note that not all children were measured by the same method. Some children (less than 1 percent) were not weighed with an adult, but rather alone on the scale. (In addition, as noted in chapter 3, in one case the scale malfunctioned, and the child weight was recorded by parent report of a well-baby check at a medical facility the day before. The values for this case were retained as reported.) Children's length was assessed with a measure mat; the interviewer laid the child down on the mat with the head at one end; the parent was asked to hold the child straight, and length was measured with the soles of the feet at right angles to the mat. Children's MUAC and head circumference was measured with a tape measure formed into a loop. For the MUAC, dots were placed on the right elbow and the shoulder joint; the distance was measured between the two dots, and a point exactly halfway between them was marked. The tape measure was placed snugly around the arm at the mark. Head circumference was measured around the crown of the head at its point of greatest circumference. As with the BSF-R and NCATS, to ensure standardized administration, step-by-step instructions for obtaining the physical measurements were included in the Child Activity Booklet. Furthermore, all of the measurements were taken twice in keeping with standard procedures used in major health studies to ensure the accuracy. All physical measurements were recorded in a form in the Child Activity Booklet.

5.3.6 Data Collection for Resident and Nonresident Fathers

Resident and nonresident fathers were identified in the closing section of the parent CAPI instrument, based on data that were collected about the members of the household. During the parent interview, the person living in the household who was identified as the partner or spouse of the parent respondent was defined as the "resident" father.[9] A "nonresident" father was the sampled child's biological father who did not live in the house with the child.

The resident father information was collected through a hard-copy self-administered questionnaire. If a household member was identified as the parent respondent's partner or spouse, the field interviewer prepared the questionnaire package for the resident father. The questionnaire was titled "Questions for Fathers and Other Important People" (referred to often as the resident father questionnaire). A Spanish version of this questionnaire was also available. Besides the questionnaire, the package contained a cover letter for the father, a self-addressed postage-paid envelope, and a gift pen. (No monetary incentive was offered to the resident father. The household incentive was intended to compensate the family for participation in the entire set of activities for the 9-month wave of the study, including the parent interview, the child assessments, and the resident father questionnaire.) If the father was present during the home visit, he was asked to complete the questionnaire before the field staff interviewer left the home. Otherwise, the resident father questionnaire was left in the home for the parent respondent's spouse/partner to complete and send back to the contractor's home office. In 99 percent of the cases, the resident father was the sampled child's biological father.

In cases where the child's biological father did not reside with him or her, information was gathered from the biological mother about the biological father. A nonresident biological father may have been asked to complete a nonresident father questionnaire. This was dependent on whether the biological mother was the respondent to the parent interview, the father's frequency of contact with the child or the birth mother, and whether the mother granted permission to contact the father. The contact level to be eligible for the questionnaire was one of the following: (1) the father had seen the child at least once in the last month; (2) the father had seen the child at least 7 days in the last 3 months; or (3) the father had been in touch with the child's birth mother at least once a month in the 3 months preceding the parent interview. If the biological mother was the respondent to the parent interview and the nonresident father had sufficient contact, the parent CAPI instrument was programmed to prompt the field interviewer to introduce the nonresident father questionnaire. The field interviewer then asked the child's mother for permission to contact the father and for information to contact and locate the nonresident father. If the mother would be seeing the nonresident father in the next 7 days, then the nonresident father package, including the nonresident father questionnaire, cover letter for the father, a check for $20, and a self-

[9] If the biological father resided in the home but was not identified as the spouse/partner of the parent respondent, he was not given the resident father questionnaire. Similarly, if the biological father was identified as the parent respondent, then he was not given a resident father questionnaire; rather, the resident father items were part of the CAPI parent instrument. If he had a spouse/partner living in the household, that person was given a resident father questionnaire.

addressed postage-paid envelope, was left with her for him to complete. If the mother would not see the nonresident father within the week following the home visit, the field interviewer contacted the nonresident father by telephone with no advance letter.[10] A Spanish version of the nonresident father questionnaire was also available. In 12 percent of the cases with a completed parent interviewer, a nonresident father was identified. Of the completed nonresident father cases, 94 percent of the questionnaires were received by mail without telephone followup. See section 5.4.3 for more information on nonresident and resident father questionnaire response rates.

If the contractor's home office did not receive the resident or nonresident father questionnaires within 2 weeks of the completion of the home visit, interviewers were instructed to follow up by telephone. The field interviewers used a telephone script to introduce themselves and the study to the father. The script also included consent information that had to be read to the father. Either the interview was completed over the phone or another questionnaire was sent to the father to complete and return in a postage-paid envelope. Field interviewers completed 3.25 percent by telephone.

Response rate problems became apparent during the course of father data collection. Although the same followup procedures were in place for the national study that were used successfully in an earlier field test, some interviewers on the national study did not follow up with fathers who had failed to return their questionnaires in a timely fashion. Furthermore, although the majority of interviewers implemented the followup protocol correctly, the initial response rates for fathers (especially nonresident fathers) did not meet expectations established by field test results. These problems were not recognized during the first half of the data collection period because they were overshadowed by production and cost problems with the home visits.

Once the father response rate problems were recognized and diagnosed, a number of steps were taken to remedy the situation. First, about 500 cases that had not received appropriate telephone followup were reassigned to senior interviewers in September 2002. These interviewers were instructed to devote about 20 hours per week to telephone calls to these fathers, in an intensive effort to increase the response rate. Second, for the remainder of the data collection period, all field interviewers were routinely sent a report to alert them that the field management system on their laptop computers showed father followup activity was required, to focus interviewer attention on the followup calls. Third, if these activities did not produce the desired results, in December 2002 another questionnaire was mailed to

[10] Given the sensitivity of paternity in these cases, a decision was made to send no letters in advance of the telephone call. The Institutional Review Board at the National Center for Health Statistics was concerned that mail addressed to the father might be opened by another member of his household.

about 1,000 fathers who had refused to answer the questions by telephone, but had agreed in the followup call to complete the questionnaire if it were remailed to them. No questionnaire had been received from these fathers in response to the remailing. For many, this was the third questionnaire mailing, because (as noted above) they had reported earlier in the followup telephone calls that they had not received the first questionnaire, and a second one had been mailed. This final questionnaire mailing increased resident father response rates by around 1 percent; nonresident father rates were not affected.

5.3.7 Data Collection Protocol by Component for Spanish-Speakers and Speakers of Other Languages

Exhibit 5-5 summarizes the approach for collecting ECLS-B 9-month data in languages other than English. All data collection instruments were translated into Spanish. For other languages, interpreters were used for most, but not all of the instruments. Three percent of the home visits employed interpreters.

Exhibit 5-5. Non-English administration of ECLS-B 9-month instruments: 2001–02

Instrument	Spanish version	Procedure in other non-English language households
Child Activity Booklet	Yes	English version used, interpreter translated instruction/question
Parent CAPI instrument	Yes	English version used, interpreter translated instruction/question
Parent self-administered questionnaire	Yes	Not administered
Resident father Questionnaire	Yes	English version used, interpreter translated instruction/question if father at home
Nonresident father questionnaire	Yes	Not administered

SOURCE: U.S. Department of Education, National Center for Education Statistics, Early Childhood Longitudinal Study, Birth Cohort (ECLS-B), Nine-Month Data Collection, 2001–02.

As part of the initial contact, the field interviewer determined whether the parent interview would best be conducted in English, Spanish, or another language. Interviewers were trained to ask the respondents in the Spanish-speaking and Asian language-speaking households which language the respondent preferred before the interview began.

For areas where Hispanic births are concentrated, Spanish-speaking interviewers were recruited who resided in those areas. Twelve percent of the ECLS-B field interviewers were bilingual in Spanish; they conducted 4 percent of the ECLS-B home visits in Spanish. In the event that a non-Spanish-speaking interviewer encountered a family who preferred to (or could only) speak Spanish, the situation was reported to the field supervisor, who transferred the case to a Spanish-speaking field interviewer. In areas where there was no Spanish-speaking interviewer, the supervisor determined whether it was economically feasible for a Spanish-speaking interviewer to travel to the area to complete the case. The guidelines for "economically feasible" were defined for the supervisors.[11] If it was not feasible to send a Spanish-speaking interviewer to the parent's home, the supervisor attempted to identify a local interpreter who could work with the interviewer or authorized use of an adult member of the household as an interpreter.

For other languages, interviewers were instructed to use a member of the local community or a household member as an interpreter.[12] Where language minority populations were concentrated in an ECLS-B primary sampling unit (e.g., Los Angeles), local groups (churches, charities, and immigrant action organizations) and universities that have experience with the population were canvassed to identify qualified translators with ties to the community. In smaller communities, where the interviewer might encounter only one case where the parent of the sampled child spoke a particular language (e.g., Mandarin) and had limited English proficiency, the interviewer was authorized to use a household member as an interpreter.

Community interpreters were required to sign a notarized affidavit of confidentiality. Household interpreters had to be 15 or older. A job aid with instructions was presented to interpreters, encouraging them to try their best to convey the exact meaning of the English words in each question, to remain neutral, and to maintain the family's confidentiality. Interviewers were instructed to review these instructions with the interpreters before the interview and to make sure the interpreter was familiar with the informed consent language and comfortable summarizing it for the respondent. During the second half

---

[11] In general, the guidelines for determining whether a case was economically feasible were based on the estimated total cost of traveling a Spanish-speaking interviewer into the area to conduct the home visit. If the cost was estimated to be greater than $1000, it was considered infeasible.

[12] A small number of ECLS-B bilingual interviewers were available who spoke Mandarin, Cantonese, Vietnamese, Khmer, and Korean, but (as noted in section 5.3.2) these interviewers were not able to travel to areas where most families who spoke these languages lived.

of the data collection period, additional procedures were developed to certify community interpreters' skills by telephone from the home office, using a protocol similar to that used for certifying bilingual ability for interviewers.

Interviewers were advised to read the questions aloud in English, to frequently ask the respondent "Did you understand the question?", to enter all responses in the laptop, and to invite the interpreter and the respondent to look at the computer screen to follow along with the interview.

The Child Activity Booklet (both English and Spanish versions) contained a number of instructions for the interviewer to say specific phrases to the child during the BSF-R, as part of the process of engaging the child in a set of standardized activities. Interpreters were asked to say these phrases to the child in the child's language. The booklet also included a short set of instructions for the parent in preparation for videotaping the NCATS**.** Interpreters were asked to convey these instructions in the parent's language.

The NCATS videotapes were coded in a central location, typically with three viewings of the each tape. The coding staff included individuals proficient in English and Spanish. For NCATS videotapes that featured a different language, an interpreter was identified. The interpreter sat with the coder during the second viewing of the tape and conveyed the English meaning of any words that were exchanged between the parent and child. Fourteen percent of the videotapes included dialogue in languages other than English. A total of 8.5 percent of these were in Spanish, and 5.8 percent were in an Asian language. A total of 37 tapes (0.39 percent) were not coded because of a foreign language for which an interpreter could not be obtained.

The parent self-administered questionnaire and the nonresident father questionnaire were not administered if the respondent did not speak English or Spanish. Given the prevalence of nonresident fathers in the population, the ECLS-B contact frequency requirements for them, and the need to obtain the mothers' cooperation to approach the fathers, the total number of ECLS-B eligible nonresident fathers in the sample was expected to be less than 25 percent of the number of resident fathers. Only a very small subset of the nonresident fathers who did not speak Spanish was expected to have limited proficiency in English. Of those with limited proficiency, it was not possible to predict with confidence whether one or more than one would speak a particular language. Therefore, developing translations of the nonresident father questionnaire in other languages or arranging for an interpreter in the telephone conversation was not cost effective.

5.3.8 Child Observations and Interviewer Remarks

After the home visit, the Child Observations and Interviewer Remarks components had to be completed in order to finish the case. The Child Observations component included questions about the child's behavior during the child assessments and questions about the home environment. It took about 10 minutes for the interviewers to complete. The Interviewer Remarks component was a self-administered questionnaire consisting of case-specific information. The Child Observations and Interviewer Remarks were computer-assisted instruments (CAI). Each was done by the field staff interviewers using laptops to answer questions and to record responses. The components were completed as soon as possible after the interviewers left the homes.

The Child Observations consisted of two parts—selected questions from the Behavior Rating Scale (BRS; see sections 2.4 and 3.1.3) and items from the Home Observation for Measurement of the Environment[13] (HOME; see sections 2.4 and 3.3.2). The first part, a set of seven questions from the BRS,[14] a subsection of the Bayley Scales of Infant Development, Second Edition, contained questions about the child's reactions during the BSF-R assessment. None of these items was answered if the BSF-R was not completed in the home.

The second part of the Child Observations component was a set of eight questions from the HOME, an inventory of items that assessed the quality of various characteristics of children's home environments. The eight HOME items were always completed as long as the baby was present and interacted with the parent respondent during the home visit. Most of them could be scored "Yes" from one observation during the home visit, whereas others had to occur more frequently.

The Interviewer Remarks component included questions about the administration of the parent interview and other aspects of data collection. This information can be used to evaluate the home visit. For example, some questions referred to the order in which home visit tasks were completed, such as whether the field staff interviewer began the home visit with the parent CAPI instrument or the child assessments. Other questions referred to whether certain aspects of the home visit were of concern to the parent respondent. Exhibit 5-6 presents a list of interviewer remarks that are included on the ECLS-B 9-month restricted-use data file as well as the 9-month public-use data file.

---

[13] Note that the items selected do not necessarily represent the HOME scale or a short form of it.

[14] Note that the items selected from the BRS are intended to be single-item indicators of child behavior during an assessment. They do not represent a short form of the BRS.

Exhibit 5-6. Interviewer Remarks items, 9-month data collection: 2001–02

Variable name	Variable description[1]
R1CMPPRN	R1 IR010 Parent interview completed before child assessments performed
R1PRNSAQ	R1 IR030 PSAQ caused anger/confusion/discomfort for respondent
R1CABUNU	R1 IR050 Any unusual events occurred during Child Activity Booklet administration
R1PSQSLF	R1 IR070 PSAQ completed alone
R1NRFCNT	R1 IR090 Issue occurred when asking to contact nonresident father
R1RFQSLF	R1 IR097 RFSAQ completed alone
R1CAPICN	R1 IR100 Unusual/concern about CAPI administration
R1PRNLNG	R1 IR120 What language used for parent interview
R1PRNTRN	R1 IR125 Was an interpreter used during parent interview
R1PRNOBS	R1 IR127 An observer was present during the home visit
R1PRNOTH	R1 IR130 Were others present during the home visit
R1PRNOT1	R1 IR131 Who was present during the home visit
R1OTHEVS	R1 IR132 Were others present for entire visit
R1OTPPSQ	R1 IR133 Were others present during-PSAQ
R1OTPBSF	R1 IR133 Were others present during-Bayley
R1OTPNCT	R1 IR133 Were others present during-NCATS
R1OTPPM	R1 IR133 Were others present during-physical measurements
R1OTPOTH	R1 IR133 Were others present during-other, please specify
R1ONEDAY	R1 IR140 Parent interview and child activities occurred in one day
R1VSTINT	R1 IR160 Interruptions occurred during home visit
R1SCHPRB	R1 IR192 Any problems scheduling the home visit
R1XFRCAS	R1 IR220 Was this case transferred between interviewer staff
R1TOTCNT	R1 IR222 # Total contacts for case
R1INPCNT	R1 IR223 # In-person contacts for case
R1TELCNT	R1 IR224 # Phone contacts for case

[1] R1 refers to the first data collection wave, when the ECLS-B children were about 9 months old.

SOURCE: U.S. Department of Education, National Center for Education Statistics, Early Childhood Longitudinal Study, Birth Cohort (ECLS-B), Nine-Month Data Collection, 2001–02.

5.4 Nine-Month Response Rates

The proportion of the sample responding to a survey can be an important indicator of survey quality. The National Center for Education Statistics has adopted standards for achieving acceptable response rates in surveys, and for computing and reporting response rates, "… encouraging high rates of response across all strata … since high response rates help ensure that results are representative of the target population." (U.S. Department of Education, National Center for Education Statistics, 2002). Standards for calculating and documenting response rates have also been published by the American Association for Public Opinion Research (2000) and the Council of American Survey Research

Organizations (1982), to allow users to evaluate potential for bias and to support comparison of results with other surveys. The American Association for Public Opinion Research and Council of American Survey Research Organizations standards are generally consistent with the NCES standards.

There are a number of approaches for calculating response rates, and surveys with complex designs such as the ECLS-B may require the use of more than one approach. For example, in a longitudinal survey, one might report the response rate for each wave of data collection using the initial eligible sample as the denominator; alternatively, the response rate for each wave might be calculated using the number of sampled cases that were assigned for data collection in the wave. In order to monitor data collection activities, unweighted response rates may be used, but response rates for evaluating survey estimates should be weighted. Response rates may be calculated based only on cases with all items or components complete, or calculations may treat partially complete cases as complete.

The approach chosen for the 9-month data collection wave of the ECLS-B is based primarily on the parent CAPI instrument. The response rate is calculated as the number of completed parent interviews divided by the number of eligible sample children, following the NCES standards. All response rates are at the child level. All sampled children were eligible except those children who died before the home visit occurred, children born to mothers younger than 15 years, children who were adopted before the age of 9 months, and those who were removed from the sample as part of a cost reduction process in February 2002. (See section 5.4.2.)

The parent CAPI instrument was chosen as the primary vehicle for determining the overall response rate for the ECLS-B 9-month data collection because there were very few instances (0.3 percent of the total sample) in which other components of the study (i.e., direct child assessments or father questionnaires) were complete for the sampled child, but the parent interview was not complete. Most completed parent CAPI instruments (96 percent) were accompanied by direct child assessment data. The parent CAPI instrument was the largest component of the home visit, representing more than half of the time required in the household. In addition, the data from this instrument alone were regarded as having a high level of interest for a broad range of analysts.

5.4.1 Steps to Improve Response Rates

Through the course of development activities for the ECLS-B, and throughout the data collection period for the 9-month wave, a number of steps were taken to improve response rates. This section discusses four major aspects of the response rate improvement efforts in the data collection task: the use of incentives, the refusal conversion process, procedures for locating the sampled cases, and procedures for other nonresponse cases.

5.4.1.1 Incentives

Offering incentives to potential respondents can have a positive effect on response rate. A variety of monetary and nonmonetary incentives were employed in the 9-month wave of the ECLS-B. As part of the advance package, parents of all sampled children were mailed a small book, an ECLS-B magnet, and a child's bib bearing the ECLS-B logo. In appreciation for the household's participation in the ECLS-B, the parent respondent was offered a check for $50. Households with twins were offered $100. The check was presented at the end of the home visit and was targeted toward the entire household, including the parent, the child or children, and the resident father (if identified in the household). In addition, a book was offered to the parent to give to the child at the end of the home visit; households with twins were offered two different books. Fathers were given an ECLS-B ballpoint pen to use in completing the questionnaire. Nonresident fathers were sent checks for $20 upon receipt of the completed nonresident father questionnaire. The effectiveness of these incentives can be evaluated in part by a comparison of response rates between the 2000 field test and the national study. The parent interview response rate improved somewhat with the increased incentive (from 73.7 to 76.8 percent, both unweighted),[15] but the response rate for both groups of fathers decreased. The resident father rate (75.1 percent, unweighted, a decrease of more than 9 points from the field test)[16] may have been affected negatively by the elimination of a targeted prepaid monetary incentive. However, the nonresident fathers rate (50.9 percent, unweighted, 16 points less than the field test) decreased despite an increase in the monetary incentive. This is likely to have been the result of less than universal, timely prompting for nonresponse, as discussed in section 5.3.6.

---

[15] Weights were not developed for the field test, so the appropriate comparison is between unweighted response rates.
[16] The field test response rate for fathers was higher than the field test response rate for parents. Father response rates were conditional on the completion of a parent interview.

5.4.1.2 Refusal Conversion

On most voluntary surveys, the largest component of nonresponse consists of sample members who refuse to participate, and the ECLS-B 9-month data collection is no exception. To minimize refusals, ECLS-B interviewers were trained to gracefully exit interactions with potential respondents that seemed likely to lead to a refusal to participate. These cases were given dispositions of "initial refusal." A letter tailored to the potential respondent's concerns was mailed. An interviewer was assigned to visit or call the respondent again, to determine whether the respondent's concerns had been addressed by the letter. Whenever possible, a different interviewer was assigned to follow up on the initial refusal. This process is known as "refusal conversion." It was not attempted when the potential respondent indicated clearly in the initial interaction that he or she was adamantly refusing to participate in the study; all such cases were classified as "final refusal" and were not subjected to the refusal conversion process. A total of 1,168 cases were classified as initial refusal in the 9-month round; 22 percent were completed after refusal conversion was attempted. At the end of data collection, 14.0 percent of the sampled cases (unweighted) were classified as "final refusal."

5.4.1.3 Locating Activities

A unique feature of the ECLS-B study design that has an impact on nonresponse is its dependence on the nation's birth records as a sample frame. The birth record is almost always completed in the hospital when the birth occurs. The hospital transmits the record to the state's vital statistics department, which in turn passes it to the National Center for Health Statistics (NCHS). Early in the development of the study design, it was determined that the flow of birth records to NCHS was not uniformly rapid enough to support data collection in the child's home earlier than 9 months after the birth occurred, or to support pre-field locating activities earlier than 7 months after the birth. Thus, locating information was 9 months old by the time the home visit was scheduled.[17]

"Unlocatable" was a large component of nonresponse, second only to refusals. Of the completed parent CAPI instruments in the 9-month data collection wave, 34 percent were conducted at a different address than indicated on the birth record, indicating the high degree of mobility of families in

[17] In order to get the birth record information within seven months of the birth, special arrangements were made with the National Center for Health Statistics, Centers for Disease Control and Prevention in the U.S. Department of Health and Human Services, which administers the National Vital Statistics System, and with each state's vital statistics department. Otherwise, the address information would have been much older.

early stages of formation. To minimize "unlocatable" as a component of nonresponse, a number of locating steps were undertaken. Before assignment to a field interviewer, address correction requests were mailed to post offices, and telephone number and address searches were conducted using large electronic databases of American households. Locating specialists were assigned to call directory assistance and explore Internet sources to find cases with unconfirmed addresses. All cases without confirmed addresses were assigned to field interviewers, and if the interviewer could not locate the household at the address provided, a number of field locating steps were implemented, including talking to residents and neighbors of the address provided, talking to the landlord, and visiting the local post office and library. In the end, 6.1 percent of the sample (unweighted) could not be found.

5.4.1.4 Procedures for Other Nonresponse Cases

The "other nonresponse" category (3.2 percent of the eligible sample) includes cases that posed a language barrier, as well as a residual "other" category. Language barriers were a subject of considerable concern to the ECLS-B. For the population of children born in the U.S. in 2001, 23 percent had mothers born outside the U.S. Many of these mothers did not speak English very well. This problem was addressed in the ECLS-B 9-month data collection by determining the language spoken by the parent respondent and transferring the case to an interviewer who spoke the language or identifying an interpreter who spoke the language. A set of procedures was developed to ensure that standard practices were followed in collecting data from households with limited English proficiency and in using interpreters. In the end less than 0.1 percent of cases could not be completed because of language issues.

The residual "other" category includes cases that were judged too far to interview, based on an estimate of travel cost. Some of these were households that moved far away from the PSU in which the child was born, and far from any other ECLS-B PSU. Others were households that, even when the child was born, lived far way from the place of birth. The sample included some cases residing outside the U.S. The cost of traveling to each of these were evaluated, and those deemed too costly were considered too far to interview and treated as nonresponse. Other types of cases included in the residual "other" group were cases that could never be contacted despite many attempts, and cases where the parent was not capable of participating in the interview, due to illness or disability. About 3.1 percent of the sample (unweighted) was too far to interview, never contacted, or unable to participate.

5.4.2 Overall Response Rate

As discussed in chapter 4, sections 4.1.4 and 4.1.5, the ECLS-B 9-month sample included cases selected as substitutes for cases in PSUs with severe state registrar or institutional review board restrictions. The response rates presented in this section treat all substitute cases as original cases; that is, the substitute cases contribute to both the numerator and denominator of the unweighted and weighted response rates. See chapter 5 of the *Early Childhood Longitudinal Study, Birth Cohort (ECLS-B) Design and Operations Report for the Nine-Month Data Collection* (U.S. Department of Education, National Center for Education Statistics, forthcoming [a]) for more details on substitution and response rates before and after substitution.

The response rate for the 9-month data collection was 74.1 percent (table 5-1), based on weighted data (i.e., using the base weights). The response rate is the number of completed parent interviews divided by the total eligible sample. To be considered complete, the parent interview had to have successfully passed through the child development section of the CAPI program (that is, the first three sections—Introduction (IN), Family Structure (FS), and Child Development (CD)—see exhibit 2-3). Ineligible cases were excluded from the denominator in the response rate calculation. Ineligibles consisted of 4 groups: (1) deceased cases;[18] (2) cases where the child was adopted before 9 months of age; (3) cases where the child's mother was younger than 15; and (4) cases that were removed from the sample as part of a sample reduction in February 2002, in an effort to reduce costs. All cases completed before the sample reduction were considered eligible, but cases selected for sample reduction and completed after the reduction were considered ineligible.

The unweighted and weighted response rates in table 5-1 can differ for two reasons:

1. Differences in response rates across the 36 birth certificate sampling strata, which have different weights.
2. Differences in response rates between cases before and after sample reduction.

The sample was reduced during the course of data collection, before sufficient effort had been expended to achieve the highest possible response rate for cases fielded with early birth months, and

[18] The biggest effect of this group was on very low birth weight cases, where 26.4 percent of the sample was deceased by the time of the home visit. (This percentage includes a small but unknown number of adoptions, because some states did not distinguish between deaths and adoptions in reporting to NCHS on selected cases that should be considered ineligible. It does not include some deaths that were identified by states and removed from the list of births prior to sampling.) This experience was consistent with expectations; the sample design assumed a 26.0 percent mortality rate for very low birth weight children in the first 9 months of life, based on an analysis of birth and mortality files from prior years, and this assumption was incorporated in the sample selection rates.

before cases with later birth months had been fielded. The weighted response rates (the right column in table 5-1) use the base weights, which incorporate the adjustment for cases completed before the sample reduction.

Nonresponse cases were classified in three categories. “Unlocatable” signifies that the case could not be found at the parent’s residential address on the birth record, and attempts to determine a current address were unsuccessful. “Refusal” indicates that the case was found but the parent did not agree to participate in the study. “Other” represents all other types of nonresponse, including final language barrier, too far to interview, out of the country, and other reasons the parent interview could not be completed. The (weighted) unlocatable rate was 6.9 percent, the (weighted) refusal rate was 16.0 percent, and the (weighted) other nonresponse rate was 2.9 percent.

Table 5-1. Final case status of the 9-month data collection, based on parent interview status: 2001–02

Status	Number	Unweighted percent	Weighted percent
Total	13,921	100.0	100.0
Complete	10,688	76.8	74.1
Unlocatable	848	6.1	6.9
Refusal	1,943	14.0	16.0
Other	442	3.2	2.9

NOTE: All unweighted and weighted response rates based on completion of the parent CAPI instrument. The base weight was used (W1BASEWT).
SOURCE: U.S. Department of Education, National Center for Education Statistics, Early Childhood Longitudinal Study, Birth Cohort (ECLS-B), Nine-Month Data Collection, 2001–02.

5.4.3 Component Response Rates

The overall response rate was based on the completion of the parent interview. However, it was possible to complete the parent CAPI instrument, but not one or more of the other study components. The response rates for the child assessment component included the BSF-R, NCATS, the physical measurement data, and the Child Observations. Completed child assessments were defined as cases with either (1) scores for the BSF-R mental or motor scales or (2) complete physical measurements (both length and weight). The response rate was 96.9 percent. It was conditioned on a completed parent interview. In other words, only completed parent interviews were included in the denominator; the numerator is the number of completed child assessments amongst completed parent interviews. This approach was chosen because most child assessments were preceded by the parent CAPI instrument. Relatively few child assessments were not associated with a completed parent interview, and the value of the child assessment data is greatly enhanced by the parent data. An alternative approach to calculating component response rates would use all eligible cases (regardless of parent interview completion status) as the denominator. With that approach, the child assessment component unweighted response rate would be about 73.4 percent, and the weighted response rate would be about 70.8 percent.

The response rate for the resident father questionnaire is also conditioned on the completion of the parent CAPI instrument.[19] This approach is appropriate because the resident father could only be identified through the parent CAPI instrument. This response rate, based on the identification of a spouse or partner of the parent respondent who was living in the same household, was 76.1 percent (weighted). It is the number of completed resident father questionnaires received divided by the number of eligible resident fathers. The response rate for the nonresident father questionnaire (conditioned on the completion of the parent CAPI instrument and based on (1) the identification of the biological father by the mother in the parent interview, (2) her consent to his participation in the study, and (3) criteria for frequency and recency of contacts with either mother or child) was 50.0 percent (weighted).[20] It is the number of completed nonresident father questionnaires received divided by the number of eligible nonresident fathers.

Tables 5-2 and 5-3 show the response rate for all of the ECLS-B components in the 9-month wave. A nonresponse bias analysis was conducted; see *Early Childhood Longitudinal Study, Birth Cohort (ECLS-B), Methodology Report for the Nine-Month Data Collection. Volume 2: Sampling* (U.S. Department of Education, National Center for Education Statistics, forthcoming [b]). The evaluation consisted of the following elements:

- Evaluation of response rates;
- Comparison of frame data between respondents and nonrespondents;
- Comparison of survey data between respondents and "proxy" nonrespondents;
- Sensitivity analysis of potential for nonresponse;
- Comparison of ECLS-B data with other surveys; and
- Analysis of factors that influence likelihood of survey response; and

[19] Note that the overall response rate for resident fathers was 56.4 percent (76.1 percent times the parent response rate of 74.1 percent).
[20] Note that the overall response rate for nonresident fathers was 37.1 percent (50.0 percent times the parent response rate of 74.1 percent).

- Evaluation of the impact of substitution on nonresponse bias.

The analysis benefited from the usually rich information available on the frame. The birth record contains a number of important variables on the mother and the child that support many comparisons between respondents and nonrespondents. For fathers, many additional data items were collected in the parent interview.

In table 5-3, the column for all fathers represents children for whom a parent interview and the appropriate father questionnaire or questionnaires were completed. Table 5-4 provides the number of cases that passed through each stage of father identification and completion. For resident fathers, the number of “children identified with fathers” (8,344) was determined in the parent interview; these fathers were identified as spouses or partners currently living in the child’s household with the parent. For nonresident fathers, the number of “children identified with fathers” (2,238) was determined in the parent interview as the biological fathers identified by the parent who were not currently living in the child’s household. (Note that for 120 completed parent interviews, no father was identified. These children were living in households where the parent had no spouse or partner and did not identify the biological father.) Table 5-4 shows the effects of the contact and consent criteria that were applied in the parent interview on nonresident father eligibility. Note that the father and nonresident father types are not mutually exclusive. Both resident and nonresident fathers were identified for 27 children. Thus the number of children identified without fathers plus the number of resident and nonresident fathers in tables 5-4 and 5-5 does not sum to the total number of children. Table 5-5 provides the final case status of the father questionnaires. For each father type, the final outcomes (completed, unlocatable, refused, or other nonresponse) are shown for all eligible cases.

Table 5-6 shows the child-level response rate (conditioned on the completed parent interviews) by the different sample domains and other characteristics of the child. The three sample domains (race/ethnicity, birth weight, and plurality) overlap so that all sample children can be classified in each domain. The percent in the “Complete” column is calculated in the same way as the weighted response rate (with base weights) in table 5-1.

Table 5-2. Unit response rates for the 9-month data collection, by parent CAPI instrument and child assessments: 2001–02

Child characteristics	Parent CAPI instrument			Child assessments		
		Response rates[1]			Response rates[1]	
	Number	Unweighted percent	Weighted percent	Number	Unweighted percent	Weighted percent
Total	10,688	76.8	74.1	10,221	95.6	95.6
Race/ethnicity						
American Indian or Alaska Native	873	79.2	79.3	835	95.6	95.6
Chinese	466	63.7	63.2	418	89.7	90.8
Other Asian/Pacific Islander	1,345	73.4	70.1	1,287	95.7	95.5
Hispanic	1,581	75.9	72.3	1,501	94.9	94.6
Black	1,771	82.0	79.7	1,715	96.8	96.8
White	4,652	77.4	73.7	4,465	96.0	95.6
Birth weight (in grams)						
Less than 1,500	1,151	78.1	76.6	1,093	95.0	94.6
≥ 1,500 and < 2,500	1,647	80.1	77.3	1,585	96.2	96.0
2,500 or more	7,890	75.9	73.9	7,543	95.6	95.6
Plurality						
Twin	1,658	82.0	80.7	1,585	95.6	95.5
Non-twin (single birth or other multiple births)	9,030	75.9	73.9	8,636	95.6	95.6
Household poverty level[2]						
Less than 100 percent of poverty	2,603	†[3]	†[3]	2,491	95.7	95.5
100–149 percent of poverty	1,633	†[3]	†[3]	1,557	95.3	95.3
150–199 percent of poverty	1,735	†[3]	†[3]	1,659	95.6	95.7
200 percent or more of poverty	4,717	†[3]	†[3]	4,514	95.7	95.6

See notes at end of table.

Table 5-2. Unit response rates for the 9-month data collection, by parent CAPI instrument and child assessments: 2001–02—Continued

Child characteristics	Parent CAPI instrument			Child assessments		
		Response rates[1]			Response rates[1]	
	Number	Unweighted percent	Weighted percent	Number	Unweighted percent	Weighted percent
Household SES quintile						
1 – lowest	2,110	†[2]	†[2]	2,007	95.1	94.9
2	2,141	†[2]	†[2]	2,055	96.0	96.0
3	2,086	†[2]	†[2]	2,000	95.9	95.8
4	1,986	†[2]	†[2]	1,896	95.5	95.8
5 – highest	2,365	†[2]	†[2]	2,263	95.7	95.3
Region of residence						
Northeast	1,646	72.5	68.4	1,538	93.4	93.0
Midwest	2,460	79.4	76.5	2,399	97.5	97.7
South	3,705	77.6	75.0	3,577	96.5	96.4
West	2,871	76.5	75.1	2,701	94.1	94.0
Outside the 50 states and the District of Columbia	6	25.0	20.9	6	100.0	100.0
Mother’s highest grade						
Less than 4 years of high school	2,115	77.2	73.8	2,027	95.8	95.2
Four years of high school	3,241	76.9	73.6	3,106	95.8	96.0
Some college	2,214	77.3	74.8	2,113	95.4	95.3
Four or more years of college	2,868	75.8	74.4	2,742	95.6	95.9
Not classifiable	250	78.1	74.9	233	93.2	90.4

† Not applicable.

[1] All response rates are computed at the child level. The parent CAPI instrument component rates are based on the entire eligible sample. The child assessment rates are conditioned on completing the parent CAPI instrument.

[2] Poverty level was determined using data on household income and household size obtained during the parent interview. This information was compared to the U.S. Bureau of the Census’ weighted poverty thresholds for 2001 for households with children. For example, a household was defined as being at 100 percent of poverty if their household income was below the poverty threshold for a family of that size.

[3] Household poverty level and SES quintile rate are not available on the birth record, so it is not possible to calculate a parent CAPI instrument response rate for households by income level.

NOTE: The base weight (W1BASEWT) was used. Source of information for household poverty level and SES quintile was the parent questionnaire. Source of information for all other child characteristics was the birth record. Race/ethnicity here represents the race/ethnicity of the child, as derived from the mother’s race/ethnicity on the birth record (and the father’s race/ethnicity on the birth record, for Chinese and American Indian or Alaska Native births).

SOURCE: U.S. Department of Education, National Center for Education Statistics, Early Childhood Longitudinal Study, Birth Cohort (ECLS-B), Nine-Month Data Collection, 2001–02.

Table 5-3. Unit response rates for the 9-month data collection, by resident father questionnaire, nonresident father questionnaire, and all fathers: 2001–02

Child characteristics	Resident Father Questionnaire			Nonresident Father Questionnaire			All Fathers[3]		
		Response rates[1]			Response rates[1]			Response rates	
	Number	Unweighted percent	Weighted percent	Number	Unweighted percent	Weighted percent	Number	Unweighted percent	Weighted percent
Total	6,270	75.1	76.1	679	51.2	50.0	6,937	71.8	72.6
Race/ethnicity									
American Indian or Alaska Native	470	72.6	71.7	82	61.2	62.1	548	70.5	69.8
Chinese	347	75.9	75.8	‡	‡	‡	348	75.8	75.7
Other Asian/Pacific Islander	833	70.5	70.2	43	49.4	49.0	876	69.2	68.9
Hispanic	848	67.3	66.1	68	44.4	44.1	913	64.9	63.9
Black	461	62.6	62.0	351	51.9	51.5	811	57.6	56.9
White	3,311	81.5	81.5	134	47.3	48.7	3,441	79.3	79.3
Birth weight (in grams)									
Less than 1,500	615	74.3	73.3	88	50.3	50.0	701	70.1	69.2
≥ 1,500 and < 2,500	922	75.1	73.6	107	45.7	48.2	1,026	70.6	69.2
2,500 or more	4,733	75.3	76.3	484	52.2	49.8	5,210	72.3	72.8
Plurality									
Twin	1,052	78.2	77.4	77	43.5	43.8	1,129	74.3	73.7
Non-twin (single birth or other multiple births)	5,218	74.6	76.0	602	51.9	49.8	5,808	71.4	72.5
Household poverty level[2]									
Less than 100 percent of poverty	915	67.0	67.3	378	52.4	50.5	1,289	62.0	61.7
100–149 percent of poverty	775	67.2	66.5	134	51.0	49.1	906	64.2	63.4
150–199 percent of poverty	1,050	74.6	75.7	95	49.2	49.5	1,143	71.7	72.7
200 percent or more of poverty	3,530	79.9	81.3	72	45.3	47.4	3,599	78.8	80.0

See notes at the end of table.

Table 5-3. Unit response rates for the 9-month data collection, by resident father questionnaire, nonresident father questionnaire, and all fathers: 2001–02—Continued

	Resident Father Questionnaire			Nonresident Father Questionnaire			All Fathers[3]		
		Response rates[1]			Response rates[1]			Response rates	
Child characteristics	Number	Unweighted percent	Weighted percent	Number	Unweighted percent	Weighted percent	Number	Unweighted percent	Weighted percent
Household SES quintile									
1 – lowest	692	61.7	62.2	294	53.1	48.3	984	58.9	57.9
2	1,054	68.9	69.5	174	47.8	50.5	1,224	65.0	66.1
3	1,201	73.1	73.0	140	52.0	52.2	1,337	70.2	70.3
4	1,432	80.5	82.8	56	50.0	50.9	1,487	78.7	81.0
5 – highest	1,891	83.3	85.8	15	39.5	42.5	1,905	82.3	84.8
Region									
Northeast	1,001	76.2	74.5	78	44.3	45.1	1,078	72.5	71.1
Midwest	1,571	81.3	83.2	148	48.4	47.0	1,709	76.8	78.6
South	2,058	75.9	77.0	345	55.6	54.4	2,403	72.2	73.1
West	1,634	68.7	69.0	108	46.0	42.9	1,741	66.7	66.7
Outside the 50 states and the District of Columbia	6	100.0	100.0	‡	‡	‡	6	100.0	100.0
Mother's highest grade									
Less than 4 years of high school	878	67.0	67.3	226	50.2	49.6	1,101	62.6	63.0
Four years of high school	1,619	70.6	71.7	295	52.4	50.9	1,909	67.1	67.8
Some college	1,432	78.4	78.5	105	47.5	47.2	1,536	75.2	75.5
Four or more years of college	2,232	81.5	83.9	31	59.2	50.4	2,261	80.8	83.1
Not classifiable	109	62.6	62.0	22	59.5	51.9	130	61.9	59.6

See notes at the end of table.

Table 5-3. Unit response rates for the 9-month data collection, by resident father questionnaire, nonresident father questionnaire, and all fathers: 2001–02—Continued

	Resident Father Questionnaire			Nonresident Father Questionnaire			All Fathers[3]		
		Response rates[1]			Response rates[1]			Response rates	
Child characteristics	Number	Unweighted percent	Weighted percent	Number	Unweighted percent	Weighted percent	Number	Unweighted percent	Weighted percent
Parents in household									
Biological mother, biological father	6,162	75.3	76.3	†	†	†	6,162	75.3	76.3
Biological mother, other father	63	61.2	61.4	9	50.0	52.2	60	58.3	60.6
Biological father, other mother	5	100.0	100.0	†	†	†	5	100.0	100.0
Biological mother only	†	†	†	670	51.1	49.8	670	51.1	49.8
Two adoptive parents	13	100.0	100.0	†	†	†	13	100.0	100.0
Related guardian(s)	22	78.6	73.7	†	†	†	22	78.6	73.7
Unrelated guardian(s)	5	55.6	34.5	†	†	†	5	55.6	34.5

† Not applicable.

‡ Reporting standards not met.

[1] All response rates are computed at the child level. The resident father questionnaire and the nonresident father questionnaire are conditioned on completing the parent CAPI instrument. Response rates for the resident father and nonresident father questionnaires do not add to the response rates for all fathers, because a case could have both questionnaires completed. The weights were developed for all fathers; cases eligible for both questionnaires have a positive weight only if both were completed for the case.

[2] Poverty level was determined using data on household income and household size obtained during the parent interview. This information was compared to the U.S. Bureau of the Census' weighted poverty thresholds for 2001 for households with children. For example, a household was defined as being at 100 percent of poverty if their household income was below the poverty threshold for a family of that size.

[3] "All fathers" includes both resident and nonresident fathers. A child could have a resident father, a nonresident father, both, or neither. Children with both father types are counted as having unit response if both father questionnaires were complete.

NOTE: The base weight (W1BASEWT) was used. Source of information for household poverty level and SES quintile was the parent questionnaire. Source of information for all other child characteristics was the birth record. Race/ethnicity here represents the race/ethnicity of the child, as derived from the mother's race/ethnicity on the birth record (and the father's race/ethnicity on the birth record, for Chinese and American Indian or Alaska Native births).

SOURCE: U.S. Department of Education, National Center for Education Statistics (ECLS-B), Nine-Month Data Collection, 2001–02.

Table 5-4. Stages in identifying and completing father questionnaires, by father type, 9-month data collection: 2001–02

Stage	Number	Unweighted percent[1]	Weighted percent[1]
Total	10,688	100.0	100.0
Children with no father identified	120	1.1	†
Resident father questionnaire			
Children identified with fathers	8,344	100.0	100.0
Completed interviews	6,270	75.1	76.1
Nonresident father questionnaire			
Children identified with fathers	2,238	100.0	100.0
Children with fathers who met the contact frequency eligibility criteria	1,778	79.4	78.4
Children whose mother provided information to contact father	1,326	74.6	74.5
Completed interviews	679	51.2	50.0

† Not applicable.
[1] All response rates are computed at the child level. The resident father questionnaire and the nonresident father questionnaire are conditioned on completing the parent CAPI instrument.
NOTE: The base weight (W1BASEWT) was used. Source of information for stage was the parent and father questionnaires. Race/ethnicity here represents the race/ethnicity of the child, as derived from the mother's race/ethnicity on the birth record (and the father's race/ethnicity on the birth record, for Chinese and American Indian or Alaska Native births).
SOURCE: U.S. Department of Education, National Center for Education Statistics, Early Childhood Longitudinal Study, Birth Cohort, (ECLS-B), Nine-Month Data Collection, 2001–02.

Table 5-5. Final case status of the father questionnaires, by father type, 9-month data collection: 2001-02

Father type and status	Number	Unweighted percent[1]	Weighted percent[1]
Resident father questionnaire			
Total	8,344	†	†
Complete	6,270	75.1	76.1
Unlocatable	22	0.3	0.3
Refusal	161	1.9	1.8
Other	1,891	22.7	21.8
Nonresident father questionnaire			
Total	1,326	†	†
Complete	679	51.2	50.0
Unlocatable	42	3.0	2.7
Refusal	88	6.6	6.3
Other	517	39.0	41.0

† Not applicable.

[1] All response rates are computed at the child level. The resident father questionnaire and nonresident father questionnaire are conditioned on completing the parent CAPI instrument.

NOTE: The base weight (W1BASEWT) was used. Source of information for father status was the field management and receipt control system.

SOURCE: U.S. Department of Education, National Center for Education Statistics, Early Childhood Longitudinal Study, Birth Cohort, (ECLS-B), Nine-Month Data Collection, 2001–02.

Table 5-6. Final case status of the 9-month data collection, based on parent interview status, by child characteristics: 2001–02

	Complete[1]		Unlocatable		Refusal		Other		Total	
Child Characteristics	Number	Percent	Number	Percent	Number	Percent	Number	Percent	Number	Percent
Total	10,688	76.8	848	6.1	1,943	14.0	442	3.2	13,921	100.0
Child's race/ethnicity										
American Indian/ Alaska Native	873	79.2	94	8.5	111	10.1	24	2.2	1,102	7.9
Chinese	466	63.7	43	5.9	150	20.5	73	10.0	732	5.3
Other Asian/Pacific Islander	1,345	73.4	94	5.1	301	16.4	92	5.0	1,832	13.2
Hispanic	1,581	75.9	244	11.7	173	8.3	86	4.1	2,084	15.0
Black	1,771	82.0	163	7.5	169	7.8	58	2.7	2,161	15.5
White	4,652	77.4	210	3.5	1,039	17.3	109	1.8	6,010	43.2
Birth weight (in grams)										
Less than 1,500	1,151	78.1	95	6.4	179	12.2	48	3.3	1,473	10.7
$\geq$ 1,500 and <2,500	1,647	80.1	106	5.2	259	12.6	43	2.1	2,055	14.7
2,500 or more	7,890	75.9	647	6.2	1,505	14.5	351	3.4	10,393	74.6
Plurality										
Twin	1,643	81.9	54	2.7	271	13.5	37	1.8	2,005	14.4
Non-twin (single birth or other multiple births)	9,045	75.9	794	6.7	1,672	14.0	405	3.4	11,916	85.6
Mother's highest grade										
Less than 4 years of high school	2,115	77.2	346	12.6	191	7.0	169	3.1	2,821	19.7
Four years of high school	3,241	76.9	280	6.6	572	13.6	222	2.9	4,315	30.3
Some college	2,214	77.3	112	3.9	460	16.1	129	2.7	2,915	20.6
Four or more years of college	2,868	75.8	87	2.3	691	18.3	156	3.6	3,802	27.2
Not classifiable	250	78.1	23	7.2	29	9.1	42	5.6	344	2.3

[1] A complete is defined as having at least completed a parent interview. The parent interview is considered complete if the interviewer successfully passed through Section CD (child development).

NOTE: Percents may not sum to 100 due to rounding. Source of information for child characteristics was the birth record. Source of information for final case status was the field management system. Percentages are unweighted data. All are row percentages, except for Total column. Race/ethnicity here represents the race/ethnicity of the child, as derived from the mother's race/ethnicity on the birth record (and the father's race/ethnicity on the birth record, for Chinese and American Indian or Alaska Native births).

SOURCE: U.S. Department of Education, National Center for Education Statistics, Early Childhood Longitudinal Study, Birth Cohort, (ECLS-B), Nine-Month Data Collection, 2001–02.

5.5 Quality Assurance for Data Collection Activities

Several types of quality control methods were used during the 9-month data collection. Those methods included validation interviews (a short telephone re-interview with parents), home visit observations, the contractor's home office review of the completed child assessment booklets, and feedback (notes to field staff about specific performance issues) for the hard-copy instruments and videotape. The following sections describe these aspects of quality control methods that were used.

5.5.1 Nine-Month Validations

For about 10 percent of the completed parent CAPI instruments, respondents were recontacted by telephone and asked a limited number of questions to validate the original interview and home visit. Because the father questionnaires were typically completed as self-administered instruments (98 percent of the resident father questionnaires and 79 percent of the nonresident father questionnaires), it was determined that validation of these questionnaires was not necessary; concerns about whether the father was actually the one who filled out the questionnaire were not supported by any evidence (e.g., in the course of prompting by telephone, interviewers did not report encountering any instances in which a member of the household volunteered that the father questionnaire was completed by someone else; telephone interviewers conducting the validation reinterviews with the parent respondents did not hear about father questionnaires completed by others; data entry staff reviewing the completed father questionnaires did not see any signs that would suggest the documents were completed by someone else); and interviews to validate self-administered questionnaires are not common practice.

For the home visits, cases were randomly selected for validation, stratified by the number of interviews completed by each field interviewer. For example, if a field interviewer completed only eight home visits, one of those eight cases was selected for validation. Each field interviewer's validation rate was set to equal at least 10 percent of her or his completed work. Cases selected for validation that were unreachable by phone were replaced with substitute validation cases. No in-person validation was done. A total of 941 validations were conducted for the home visit. The field supervisors completed home visit validations for 125 cases (13 percent), but because the supervisors had higher priorities during the first half of the data collection period (i.e., field supervisors were addressing severe production and cost problems for the home visits), the validation activities lagged behind home visit activities. Beginning in June 2002, the telephone validations were conducted by the contractor's Telephone Research Center staff.

The responsibility for doing validation was moved from the field supervisors to the Telephone Research Center to reduce the elapsed time between the home visit date and the completed validation date. Validations were conducted in both English and Spanish.

A formal script was developed for the validation. For the home visit, it included questions such as whether the 9-month parent respondent remembered the interview and if the field interviewer used a computer. The respondent was also asked if the interview was conducted in person or over the telephone, the length of the home visit, whether the BSF-R was administered, and whether the child's physical measurements were taken. Respondents were asked to provide feedback on the BSF-R administration and the physical measurements, and to share other thoughts they had about the home visit experience. Types of feedback included comments about the interviewer, about the survey, and about the child's performance (e.g., "I had no idea my child could do some of the things she did with the interviewer"). The validation form included specific case information including the field interviewer's name; date of the interview; respondent's name, address, and telephone number; respondent's relationship to the child; and the child's name and gender. The validation interviews took about 5 minutes to complete over the telephone.

Discrepancies between the interview data and the validation data were reviewed by the supervisor. Discrepancies between data for validation cases conducted by the Telephone Research Center were reviewed by project staff. An example of a discrepancy is a difference in the date of the home visit as reported in CAPI and by the respondent during validation. If the discrepancy could not be resolved by additional information (for example, the respondent did not acknowledge having been contacted to participate in the study), more cases completed by the field interviewer were selected for validation. Of the 243 field interviewers who ever worked during the 9-month data collection, there were no discrepancies discovered during validation that required pulling additional case work for the field interviewer. Interviews completed by all field interviewers who were active during 9-month data collection were successfully validated (i.e., selected data about a case pulled for validation was confirmed by the respondent in all instances.)

5.5.2 Home Visit Observations

During the 9-month data collection, each field interviewer was observed during a home visit at least once by his or her supervisor or a home office project staff member, in accordance with the data quality plan (revised September 30, 2002). Field interviewers who scored low in their BSF-R certification in training were targeted for an early observation. Most of the observations (73 percent) were completed by the end of February 2002, 4.5 months after the start of the study. During the home visit, observers completed an in-person observation form that covered specific aspects of both the parent CAPI instrument and the child assessments. Field interviewers were rated on a scale of 1 to 5 with 1 being unacceptable and 5 being excellent. On overall administration the field interviewer rating averaged 4.3. On each of the individual components the average rating was 4.6 for the parent CAPI instrument, 4.2 for the BSF-R, 4.1 for the NCATS, and 4.4 for the physical measurements. Observers who were certified in administering or scoring the BSF-R also completed a BSF-R review form that calculated the accuracy of the BSF-R administration and scoring, using the same form that was used for certification in training. The review and critique form was completed for 87 percent of the observations. For these observations, the average percentage administration score on the BSF-R was 95 (range 76–100 percent), and the average percentage scoring accuracy was 97 (range 76–100 percent).

After the home visit, the observers usually met with the field interviewer to share their observations and provide feedback on specific aspects of the home visit. Examples of the type of feedback given include comments on how adequately the interviewer addressed parent concerns, critiques of the order of component administration and of the positioning of the child relative to the interviewer during the BSF-R, and observations on the quality of the NCATS videotaping. In some instances interviewers were directed to review sections of the interviewer manual to improve their understanding of the correct protocol; in others, interviewers were instructed to complete an additional practice BSF-R on videotape and send it to Westat for review.

5.5.3 Audiotapes of Home Visits

In addition to observations, all of the field interviewers were instructed to submit an audiotape of the parent CAPI instrument from their third home visit. The field supervisors critiqued the audiotapes and provided appropriate feedback to the field interviewers regarding their interviewing techniques. About 75 percent of interviewers had audiotapes reviewed.[21] They were critiqued using a standard form collecting information on the parent interview (for example, accuracy of reading questions as worded). The main outcomes of this review were to reinforce proper interviewing technique and to discourage improper technique.

[21] Interviewers did not submit audiotapes for review if their third case was observed, or if they completed fewer than three interviews.

5.5.4 Quality Control Procedures for the Direct Child Assessments

The quality of the direct child assessments was monitored and evaluated in three ways: quality review of the home visit, quality review of recordkeeping in the Child Activity Booklet, and establishment of reliability of administration of the child assessments, or, in the case of the NCATS, reliability of scoring.

Home Visit Quality Review. Interviewers' administration of all aspects of the home visit was reviewed and evaluated during the home visit observation (described above). Additional procedures were implemented within Westat and in the field to continuously assess interviewers' performance. This included having field interviewers score and critique videotaped administrations of the BSF-R in order to ensure the reliability of their administrations. In addition, after case materials were receipted at Westat, each NCATS videotape was reviewed for quality. The contents of the Child Activity Booklet were also reviewed for quality and accuracy, including recordkeeping for the BSF-R, the NCATS and the physical measurements. Finally, NCATS coders were required to establish reliability in coding on a weekly basis. Each of these procedures is described below.

Quality Review of the Child Activity Booklet. The Child Activity Booklet contains the administration instructions and record forms for the direct child assessments including the BSF-R, the NCATS, and the physical measurements. For approximately the first 3 months of data collection, all completed Child Activity Booklets were reviewed and checked for thoroughness, accuracy, and consistency upon receipt at Westat by design staff and computer-assisted data entry (CADE) staff. Design staff reviewed the first 100 receipted Child Activity Booklets, and they developed guidelines for CADE staff review. CADE staff reviewed each page of the booklet and noted missing or out-of-range values and other discrepancies. A standard form was used to record these problems. Feedback based on this form was forwarded to field supervisors on an as-needed basis whenever consequential problems were identified. For example, miscounting the number of credits or no credits on the core item set of one of the subscales of the BSF-R could cause the basal or ceiling item sets not to be administered as required. Interviewers who repeatedly made systematic errors, such as forgetting to fill in the assessment start/stop time or failing to complete comments page about modifications and skips, were notified through their field supervisors. At the beginning of the field period, when the first 1,440 cases were reviewed, approximately 25 percent (out of 180 interviewers) had recordkeeping problems. (This percentage also includes such nonconsequential errors as not including the start/stop times for the child activities. Errors of consequences affected the child's score.) Over the entire data collection, errors declined sharply, so that only 2 out of 144 interviewers were identified as having problems by the end of data collection, and these “errors” were attributable to missing physical measurements on a case due to the child’s lack of cooperation.

Administration and Scoring of the BSF-R. In addition to being observed during an in-person home visit, each interviewer was required to complete four BSF-R video scoring assignments over the course of data collection, at a rate of approximately one each 3 to 4 months. These assignments helped to track the scoring accuracy and reliability of each interviewer and served as a review of the administration procedures for the interviewers. The first videotape was sent out one month after the study went into the field. After interviewers watched and scored the tape using the standard BSF-R review form, they returned the forms to the home office, and their scores were calculated. If they achieved the dual criteria of 85 percent correct in their review of administration and 90 percent accuracy in their scoring for BSF-R items (i.e., correctly noted credit vs. no credit for the child’s performance), then they continued data collection. If they did not achieve these criteria, they were instructed to review feedback on the items they failed before continuing to work (as detailed below). Overall, the percentage of field staff who met the passing criteria on the first review assignment was 86 percent. Performance on the review tapes continued to improve as data collection continued, with passing scores for the second tape at 91 percent, the third tape at 94 percent, and the last tape at 96 percent.

Field interviewers (24 total) who did not meet these criteria on the first review assignment were sent a detailed email explaining the mistakes that they made. After reviewing the feedback from the first tape, they carried on with data collection, and a second videotaped BSF-R was sent shortly thereafter. Six interviewers did not pass on the second review. A conference call was held with them to determine how best to correct the problems. After the conference call, interviewers were asked to videotape the administration of the BSF-R during their next home visit and send in the videotape for review at Westat. They were to stop work on all other cases until notified by Westat. If the videotaped administration met the certification criteria, the interviewer was allowed to resume data collection. Of the six interviewers who were required to videotape a BSF-R administration, four returned to work immediately after their videos were reviewed, one was given additional retraining and was re-certified by a field supervisor, and one was released after two failed attempts at re-certification.

Administration and Videotape Quality of the NCATS. In addition to the in-home observation by a member of the supervisory staff, which included review of the NCATS administration, a videotape quality control form was filled out by the NCATS coders for each tape received at the home office. Major NCATS administration or taping errors were addressed immediately by contacting supervisors with specific feedback for the interviewer, and minor errors were addressed by sending written feedback to supervisors every month. Major errors were those that resulted in loss of data due to items not being codeable, such as “Caregiver gets the child’s attention before beginning the task, at the start of the teaching interaction.” This item would not be codeable if the interviewer failed to capture the first few seconds of the interaction when the parent introduced the NCATS activity to the child. Minor errors had no consequences for coding and included such problems as not having a complete voice stamp on the tape or not recording the language the mother used on the tape label. Approximately 33 percent of interviewers received feedback and corrective training concerning their NCATS administration over the course of the data collection period; feedback was much more frequent in the early months. Videotape quality control scores averaged 86 percent for all interviewers over the entire field period.

Improvement was generally seen immediately after interviewers received feedback from the field supervisors during the weekly supervisor-interviewer progress calls. The success of this fast feedback system is reflected in the fact that there were no cases where the NCATS was not scoreable because of missing items due to poor videotape quality or poor administration. That said, there were approximately 8 percent of cases that did not have NCATS scores, primarily due to parent refusal to allow videotaping, short tapes that lasted less than 45 seconds and were therefore uncodeable, the presence of more than one caregiver (defined by NCAST at the University of Washington as uncodeable), and camera malfunctions.

Physical Measurements. In addition to the home visit observation by supervisors and other Westat staff, the information in the Child Activity Booklet for each set of physical measurements was reviewed for accuracy by the computer-assisted data entry (CADE) staff. Identifiable errors were reported on a weekly basis to field supervisors who in turn would follow up with interviewers and give detailed feedback regarding these errors. (For further details about corrections to physical measurement errors, please see section 3.1.4.) Identifiable errors included recording the child’s length in inches instead of centimeters, a difference between repeat measurements that was greater than 5 percent, failure to tare the child’s weight, and misplacing the decimal points in the space provided in the record form. Although there was a high percentage of physical measurements recordkeeping errors at the beginning of the data collection (most of the interviewers’ record-keeping problems mentioned above were in physical measurements), these problems were eliminated with a field memo reminding interviewers to check all measurements to make sure they were within 5 percent of each other. The memo also reminded interviewers to use the decimal point already on the record form and to record the measurements to the nearest tenth. By the end of the data collection period, detectable errors were eliminated from the physical measurements record form entirely.

5.5.5 Assessor Effects

The results of IRT reliability analyses are summarized in section 3.1.1.2. This section presents the results of hierarchical linear modeling (HLM) analyses that were conducted to assess interviewer effects on the BSF-R. The BSF-R is an individually administered assessment. Therefore, the quality of the administration may introduce unwanted sources of variance into children's performance on the BSF-R.

A three-level multilevel analysis (Goldstein, 1995; Bryk and Raudenbush, 1992) was carried out to examine sources of variance. Table 5-7 presents maximum likelihood estimates of the components of variance for supervisor (level 3), interviewer (level 2), and child (level 1) for the mental scale of the BSF-R. That is, with children nested within interviewers, nested within supervisors. The addition of these variables progressively reduces the residual variance at level 1.

These results suggest that on the BSF-R mental scale, interviewers account for as much as 14 percent of the total variance in mental scale scores. This implies that as much as 14 percent of the total variance depends on the particular interviewer selected to conduct the assessment. However, since interviewers work in different geographical areas, interviewer differences are confounded with any real differences between areas. True-score differences between children account for fully 86 percent of the total variance and are confined to Level 1 of the model.

Table 5-7. Components of variance associated with the child, interviewer, and supervisor for the BSF-R mental scale, scored with publisher item calibrations, 9-month data collection: 2001–02

Partitioning the variance in mental scale scores						
	Multilevel Model					
	Unconditional (Partitions the variance between levels)		Age[1]		Age and BSID age[2]	
Variance	Estimate	Percent	Estimate	Percent	Estimate	Percent
Total	0.93478	100.0	0.55676	100.0	0.54957	100.0
Children	0.81406	87.1	0.47926	86.1	0.47135	85.8
Interviewers	0.10714	11.5	0.07553	13.6	0.07668	14.0
Supervisors	0.01358	1.5	0.00197	0.4	0.00154	0.3

[1] Includes child's age as an explanatory variable at the child level of the model (level 1).
[2] Includes both child's chronological age and BSID age as explanatory variables (level 1). BSID age is child's chronological age adjusted for prematurity, following procedures recommended by the National Center for Health Statistics (NCHS).
SOURCE: U.S. Department of Education, National Center for Education Statistics, Early Childhood Longitudinal Study, Birth Cohort (ECLS-B), Nine-Month Data Collection, 2001–02.

Similar analyses were conducted for the BSF-R motor scale and are presented in table 5-8.

Table 5-8. Components of variance associated with the child, interviewer, and supervisor for the BSF-R motor scale, scored with publisher item calibrations, 9-month data collection: 2001–02

Partitioning the variance in motor scale scores						
	Multilevel Model					
	Unconditional (Partitions the variance between levels)		Age[1]		Age and BSID age[2]	
Variance	Estimate	Percent	Estimate	Percent	Estimate	Percent
Total	1.55702	100.0	0.79274	100.0	0.77775	100.0
Children	1.42045	91.2	0.74587	94.1	0.73133	94.0
Interviewers	0.09721	6.2	0.04103	5.2	0.04178	5.4
Supervisors	0.03936	2.5	0.00584	0.7	0.00464	0.6

[1] Includes child's age as an explanatory variable at the child level of the model (level 1).
[2] Includes both child's chronological age and BSID age as explanatory variables (level 1). BSID age is child's chronological age adjusted for prematurity, following procedures recommended by the National Center for Health Statistics (NCHS).
SOURCE: U.S. Department of Education, National Center for Education Statistics, Early Childhood Longitudinal Study, Birth Cohort (ECLS-B), Nine-Month Data Collection, 2001–02.

Similar to the mental scale, HLM analyses indicate that interviewer differences do exist, although they do not appear to be generally a major source of error. Table 5-8 shows that on the BSF-R motor scale, interviewers account for only 5.4 percent of the total variance in motor scale scores suggesting that interviewers were more reliable on the administration and scoring of the motor scale than on the mental scale.

5.6 NCATS Coder Training and Certification

NCATS videotapes were returned to the home office where they were coded by certified NCATS coding staff according to the NCATS coding system, developed by Dr. Kathryn Barnard at the University of Washington School of Nursing. Information about the NCATS is available from the Nursing Child Assessment Satellite Training (NCAST) web site *http://www.ncast.org*.

Among the key reasons why the NCATS was chosen was that the University of Washington conducts standardized training for individuals to become certified NCATS trainers. In addition, inter-lab reliability could be obtained between the University of Washington and the contractor's coders, which would ensure the comparability of the contractor's NCATS coding with both the University of Washington and other studies that have used the NCATS, such as the Early Head Start Research and Evaluation Project.

In order to train NCATS trainers, four staff members from Westat's Child and Family Studies group attended the standardized training for NCATS trainers at the University of Washington. One staff member was designated to be the lead trainer on the NCATS, with the other three serving as assistant trainers. Of these four staff members, one was the coding supervisor for the entire 9-month data collection, one assumed the role of liaison between the NCATS coding staff and the child development group, but left about two-thirds through the data collection effort, and one assumed that role as her replacement. The remaining staff member was a backup trainer in case there was a high attrition rate for coders and multiple coder training sessions needed to be conducted. Training multiple trainers was also helpful to establish a consensual understanding of the coding system that formed a strong foundation of knowledge for the coding process.

The entire training period took 5 days. Westat developed procedures for screening candidates. At the beginning of the first day, candidates were prescreened in a three-part process that began with a personal interview and with a (non-standardized) quiz of general observational skills. Following the initial interview, candidates were presented with a set of 6 still photos and line drawings. They were allowed to review each picture for up to 5 minutes. Each picture was then followed by up to 10 questions that asked about various details in the pictures. The candidate was not allowed to refer back to the picture when answering the questions. This is a procedure that is commonly included as a subtest in some aptitude tests, for example, statewide aptitude tests for admission to police officer candidate training.

These pictures and their questions were pilot tested within the contractor's child development group among individuals who had previously been trained on the NCATS (in previous rounds of field testing), and it was found that those who scored highest on this picture-based observational quiz also scored high on the NCATS. A criterion of 90 percent was set on this observational ability screener on the basis of empirical findings of the contractor's pilot testing.

Candidates who reached 90 percent accuracy on the observation test were then given an overall introduction to the ECLS-B and participated in a self-paced, self-instruction screening. Candidates reviewed the NCATS training videotapes, reviewed the NCATS manual, and completed the exercises in the manual. The goal was to invite those candidates who scored at 85 percent agreement with NCATS standards on these exercises to participate in the formal training. All candidates who scored 90 percent or higher on the observational ability screener also scored 90 percent or higher on the NCATS self-instruction. At the completion of the self-instruction, candidates were re-interviewed to determine whether they would be interested in work as a coder.

Candidates who obtained 90 percent on the observational ability screener and 90 percent on the self-instruction were invited to participate in the formal classroom-style NCATS training. Formal NCATS training lasted 3 days, culminating in the official reliability testing scored by the NCATS training staff at the University of Washington. Eighteen candidates passed NCATS reliability at 90 percent (within 3 attempts) and received research certification from the NCATS staff at the University of Washington. These coders began coding ECLS-B NCATS videotapes as soon as they received their certification.

5.6.1 Quality Control Procedures for NCATS Coders

In any study that requires large scale coding to be done over a long period of time, there is always the chance that coders will drift from the coding standard and become unreliable in their coding. In order to prevent this from happening to the NCATS coding on the ECLS-B two steps were undertaken. First, as mentioned above, the “passing grade” for coder certification was raised from 85 percent to 90 percent agreement with the standard, called the “research certification.” This was done upon the recommendation of Dr. Barnard who found that the extra 5 percent leads to more reliable coding in the long run, suggesting that coders who can score that high during training are less likely to drift from the standard, or even if they do drift are less likely to fall below the acceptable criterion.

The second step that was taken to prevent coder drift over the long term of the data collection was the establishment of a reliability system. There are two possible ways that reliability could be maintained. One way is to do intra-lab reliabilities with the contractor’s coders. It is important, however, that the ECLS-B NCATS results be comparable with the standard maintained by the University of Washington. It was determined that reliability should be measured against the standards set by NCATS coders at the University of Washington exclusively and to forgo intra-lab reliability. Therefore, NCATS videotapes from approximately 2.5 percent of cases (n = 219) were selected at random and were sent, over the data collection period, to NCATS coders at the University of Washington. These procedures conformed to the original plan for reliability review. However, the University of Washington coders returned scores for only 171 videotapes, due to resource constraints.

A certified coder at NCAST at the University of Washington coded the tapes. This information was data entered by Westat. On a weekly basis and on a day selected at random during the week, Westat’s coders were required to code one of these reliability tapes. They had to score 85 percent or higher to continue coding. Coders who fell below this criterion were given targeted feedback and asked to code additional reliability tapes until a score of 85 percent or better was obtained. Coders were not able to code any videotapes from the ECLS-B home visits until they had regained this level of 85 percent agreement with the tapes coded by the University of Washington over the coding period.

NCATS coding began in December 2001 with 18 certified coders who completed approximately 50 NCATS videotaped interactions each day (2.6 interactions per coder). In February, 2002, when the ECLS-B sample size was reduced, 8 coders, selected because they scored lowest on reliability, were released from the study. By the end of the data collection period, 4 full-time and 2 part-time coders remained, averaging approximately 25 tapes per week each. In all, a total of 9,340 videotapes were coded over the year of data collection. The average inter-lab reliability between the contractor's coders and the University of Washington's certified coder was 86 percent.

REFERENCES

American Association for Public Opinion Research. (2000). *Standard Definitions: Final Dispositions of Case Codes and Outcome Rates for Surveys.* Ann Arbor, Michigan: AAPOR.

Bryk, A., and Raudenbush, S. W. (1992). *Hierarchical Linear Models for Social and Behavioral Research: Applications and Data Analysis Methods*. Newbury Park, CA: Sage.

Council of American Survey Research Organizations. (1982). *On the Definition of Response Rates: A Special Report of the CASRO Task Force on Completion Rates.* Port Jefferson, NY: CASRO.

Goldstein, H. (1995). *Multilevel Statistical Models.* New York: Halstead.

U.S. Department of Education, National Center for Education Statistics. (2002). *NCES Statistical Standards* (NCES 2003-601). Washington, DC: U.S. Department of Education.

U.S. Department of Education, National Center for Education Statistics. (Forthcoming [a]). *Early Childhood Longitudinal Study, Birth Cohort (ECLS-B) Design and Operations Report for the Nine-Month Data Collection*. Washington, DC: U.S. Department of Education.

U.S. Department of Education, National Center for Education Statistics. (Forthcoming [b]). *Early Childhood Longitudinal Study, Birth Cohort (ECLS-B) Methodology Report for the Nine-Month Data Collection, Volume 2: Sampling*. Washington, DC: U.S. Department of Education.

This page intentionally left blank.

6. DATA PREPARATION

As described in chapter 2, the ECLS-B 9-month data collection obtained data through a variety of means, including parent computer-assisted personal (CAPI) interviews and questionnaires, father questionnaires, direct child assessments during an in-person home visit, and field staff observation of the child's behavior and home setting during the home visit. The data preparation required depended on the mode of data collection. The parent CAPI instrument and staff observations of the children and the home setting were conducted using computer-assisted interviewing (CAI) techniques. The parent self-administered questionnaire and the father questionnaires (resident and nonresident fathers) were paper (hard-copy) forms that respondents generally completed themselves. The direct child assessments were completed by the interviewers who filled out a hard-copy child activity booklet as they conducted the child assessments. As part of the child assessments, the interviewers videotaped the NCATS, a short interaction between the parent respondent and the child. The following sections describe the data preparation activities for these three modes of data collection in more detail.

6.1 Coding and Editing Specifications for Computer-Assisted Interviewing

The designing requirements for a computer-assisted interview require decisions about edit specifications to be made early in the design process. Editing specifications consisting of both acceptable ranges and logic consistency checks were preprogrammed into the CAI questionnaires. The following sections describe the coding and editing that were conducted on the parent CAPI instrument.

6.1.1 Range Specifications

Within the parent CAPI instrument, respondent answers were subjected to both “hard” and “soft” range edits during the interviewing process. A “soft range” is one that represents the reasonable expected range of values but does not include all possible values. Responses outside the soft range were confirmed with the respondent and entered a second time. For example, the number of hours each week a child attends a day care center on a regular basis had a soft range of 1 to 50. A value outside this range could be entered and confirmed as correct by the interviewer as long as it was within the hard range of values (1 to 70).

“Hard ranges” are those that have a finite set of parameters for the values that can be entered into the computer, for example, “1–14 months” for the age child began feeding himself or herself. Out-of-range values for closed-ended questions were not accepted. If the respondent insisted that a response outside the hard range was correct, the interviewer could enter the information in a comments data file. Data preparation and project staff reviewed these comments. Out-of-range values were accepted and entered into the data file if the comments supported the response.

6.1.2 Consistency Checks (Logical Edits)

Consistency checks, or logical edits, examine the relationship between responses to ensure that they do not conflict with one another or that the response to one item does not make the response to another item unlikely. For example, the Child Care Arrangements (CC) section asks how old the child was when he or she first started to receive regular care outside the home. The response is compared with the child’s age as calculated in the Introduction (IN) section. If the child’s age when he or she first received care is entered as greater than the child’s current age, an error message is displayed and the interviewer is instructed to correct the data. When logical errors such as these occurred during a session, the interviewer saw a message requesting verification of the last response and a resolution of the discrepancy.

6.1.3 Coding

Additional coding was required for some of the items collected in the CAI instruments. These items included “Other, specify” text responses and occupation. Interviewers keyed verbatim responses to these items. Once the data were keyed, staff were trained to code these data using coding manuals designed by Westat and the National Center for Education Statistics (NCES) to support the coding process. This section describes the coding activities for the CAI instruments.

6.1.3.1 Review of “Other, specify” Items

The parent CAPI instrument contained 48 “Other, specify” items. The responses to these open-ended items were reviewed to determine if they should be coded into one of the existing response categories. During the data collection, when a respondent selected an “other” response in the parent CAPI instrument, the interviewer entered the text into a “specify” overlay that appeared on the screen. The data preparation staff reviewed these text “specify” responses and, where appropriate, coded them into one of the existing response categories. If a text “specify” response for which there was no existing response category occurred frequently enough, new codes (response categories) were added, if appropriate. For some items, particularly those in the Child Health (CH) section of the parent CAPI instrument, a sufficient number of related responses remained in the “Other, specify” after upcoding that a new category was added. For example, the category "Steroids/Anti-Inflammatories" was added as a response to the question about how the child's asthma was treated. In all, 23 new categories were added.

6.1.3.2 Parent Occupation Coding

Occupations were coded using the *Standard Occupational Classification Manual* (Executive Office of the President, Office of Management and Budget, 2000). The occupation codes were collapsed into the following 23 codes, plus 2 additional categories for unemployed/retired and uncodeable. Uncodeable answers consist of responses such as “my father's occupation” or “none of your business”—answers which were unusable for coding purposes.

1. Management Occupations
2. Business and Financial Operations Occupations
3. Computer and Mathematical Science Occupations
4. Architecture and Engineering Occupations
5. Life, Physical, and Social Science Occupations
6. Community and Social Services Occupations
7. Legal Occupations
8. Education, Training, and Library Occupations
9. Arts, Design, Entertainment, Sports, and Media Occupations
10. Health Care Practitioners and Technical Occupations
11. Health Care Support Occupations

12. Protective Service Occupations

13. Food Preparation and Serving Related Occupations

14. Building and Grounds Cleaning and Maintenance Occupations

15. Personal Care and Service Occupations

16. Sales and Related Occupations

17. Office and Administrative Support Occupations

18. Farming, Fishing, and Forestry Occupations

19. Construction and Extraction Occupations

20. Installation, Maintenance, and Repair Occupations

21. Production Occupations

22. Transportation and Material Moving Occupations

23. Military Specific Occupations

See chapter 7, section 7.5.3 for a detailed description of the types of occupations covered by each of the above codes.

Occupation coding began with an autocoding procedure using a computer string match program developed for the National Household Education Surveys (NHES) Program. The program searched the responses for strings of text for each record/case and assigned an appropriate code. About 44 percent of the cases were autocoded.

Cases that could not be coded using the autocoding system were coded manually using a customized utility program designed for coding occupations. The customized coding utility program brought up each case for coders to assign the most appropriate codes. In addition to the text strings, other information, such as main duties, highest level of education, income, name of the employer, and industry was available for the coders. The coders used this information to ensure that the occupation code assigned to each case was appropriate.

Verification of coding is an important tool to ensure quality control and to strengthen coder training. One hundred percent of autocoded and manually coded cases were verified for this first round of data collection. As a verification step, two coders independently assigned codes (i.e., double-blind coding) to occupation cases. In the case of autocoding, one coder independently assigned occupation codes and his or her assignments were compared to the autocode. A data manager adjudicated disagreements between the initial code and the verification code.

6.1.3.3 Race/Ethnicity Coding

Race is a "code all that apply" question, so multiple races can be accommodated. Often specific countries of origin were provided instead of a race. As in the ECLS-K, this country of origin information was used to assign race and ethnicity as follows:

- **Asian** includes Asian, American Asian, Asian Indian, Asian Laotian, Bengali or Bangladeshi, Cambodian, Chinese, Chinese-American, Hindu, Hmong, India, Indian-American, Indian Guyana, Indian Hindu, Indian-Muslim, Indian South African, East Indian, Japanese, Korean, Mongolian, Oriental, Pakistani, Siamese, Sri Lanka, Thai(land), and Vietnamese.

- **Native Hawaiian or other Pacific Islander** includes Native Hawaiian, Guamanian, Samoan, Carolinian, Fijian, Kosraen, Melanesian, Micronesia, Northern Mariana Islander, Palauan, Papua New Guinean, Ponapean (Pohnpelan), Polynesian, Solomon Islander, Tahitian, Tarawa Islander, Tokelauan, Tongan, Trukese (Chuukese) and Yapese.

- **Black or African American** includes African (American/origin) Black, Ethiopian, Congolese, West Indies/Indian, Jamaican, Bahamian, Haitian, South African, and Trinidadian.

- **White** includes Afghan, Arab, Armenian, Cajun, Caucasian, Croatian, Danish, Dutch, Egyptian, English, French, German, Greek, Hungarian, Italian (alone or with American, Spanish, French, German, Polish), Irish, Iranian, Iraqi, Israeli, Jewish, Jordanian, Lebanese, Mid-Eastern, Norwegian, Persian, Polish, Romanian, Russian, Slavic, Spanish, Syrian, Turkish, White, and Yemeni.

- **American Indian or Alaska Native** includes Apache, Cherokee (Cherici), Chickasaw, Guatemalan Indian, Mayan, Hispanic and American Indian/Native American, Lakota, Mexican Indian, Dominican Indian, Indian from Santo Domingo, and Ogala.

- **Hispanic** includes Bolivian, Brazilian, Chilean, Central American, Colombian, Costa Rican, Cuban, Dominican, Ecuadorean, Guatemalan, Honduran, Latina/o, Latin, Latin

American, Mexican, Michocana, Mestizo, Moreno, Nicaraguan, Panamanian, Peruvian, Puerto Rican, Salvadoran, Spanish, South American, and Venezuelan.

When a race was identified that had not been accounted for, the case was set aside for the data manager to assign using other information in the parent CAPI instrument, such as primary language or comments made by interviewers. See chapter 7, section 7.5.1 for details on how the race variables were coded and how the race/ethnicity composite was created.

6.1.3.4 Language Coding

The coding scheme used by ECLS-K for coding languages was used. Essentially, specific languages were assigned to larger familial categories. Below is a list of the ECLS-K rules that were also used for the coding of language in the ECLS-B:

- **Chinese:** Mandarin, Cantonese, Taiwanese
- **Filipino:** Tagalog, Ilocano, Cebuano, Dambales
- **Spanish:** Castellano
- **African:** African, Amharic, Afrikaans, Ethiopian, Senegalese, Ghanaian, Nigerian, Swahili, Tigrenian, Twi, Zulu, Edo, Yoruba, Efik, Mein
- **Eastern European:** Albanian, Bosnian, Bulgarian, Croatian, Hungarian, Latvian, Lithuanian, Macedonian, Romanian, Russian, Serbian, Slovakian, Ukranian, Yugoslavian, Yiddish
- **Indian Subcontinental:** Singhalese, Nepali, Tibetan, Urdu, Pakistani, Bengali, Hindi, Telegu, Kannada, Tamil, Punjabi, Marathi, Gujrati, Malayalam, Bengladeshi
- **South East Asian:** Indonesian, Thai, Malaysian, Cambodian, Khmer, Laotian, Lao, Hmong, Mien
- **Pacific Islander:** Samoan, Guamanian, Micronesian, Tongan, Marshalese, Hawaiian, Lau, Ponapaien

When a language was identified that had not been accounted for, these cases were set aside for the data manager to assign using other information in the parent CAPI instrument including comments made by the interviewer.

6.1.3.5 Partially Completed Parent Interviews

A "completed" parent instrument was defined by whether the first three sections of the CAPI instrument (Introduction [IN], Family Structure [FS], and Child Development [CD]) were completed by the respondent. Only completed interviews were retained in the final data file. A small number of these interviews (15) terminated the parent CAPI instrument after the CD section but before the end of the instrument. These interviews were included in the data file. All instrument items after the interview termination point were set to −9 for "Not ascertained."

6.1.3.6 Household Roster in the Parent CAPI Instrument

Several tests were run on the household roster to look for missing or inaccurate information. There were essentially three general types of roster tests performed to determine which cases required editing. First, the relationship of an individual to the focal child was compared to the individual's listed age and sex. Second, households with more than one mother or more than one father were scrutinized for errors. While it is possible to have more than one mother in a household—for example, a household could contain one biological and one foster mother of the focal child—such cases warranted closer inspection. Corrections were made wherever clear errors and a clear resolution existed. Lastly, the relationship of an individual to both the focal child and the respondent was examined, as there were cases in which the relationship of an individual to the focal child conflicted with his status as the spouse/partner of the respondent. For example, in a household containing a child's grandparents but not his or her parents, the grandmother may be designated as the "mother" figure, and the grandfather thus becomes the "father" by virtue of his marriage to the grandmother. These cases were examined but left unchanged. Both the original—and correct (grandfather)—relationship data and the new "parent-figure" designation (father) that had been constructed were kept. There were 112 cases with at least one of these types of inconsistencies (such as a male mother or two mothers in the household) that were examined more closely; 65 of these cases were corrected, 47 required no change. At the end of this process all data were cleaned or deemed appropriate.

6.2 Coding and Editing Specifications for Hard-Copy Instruments and NCATS Tapes

As noted earlier, the parent self-administered questionnaires, the resident and nonresident father questionnaires, and the child assessments were completed on hard-copy forms. In addition, videotapes were returned to the home office. When these materials were returned to Westat from the field, coders recorded the receipt of these forms into the Home Office Management, Entry, and Receipt System (HOMER), a project-specific forms tracking system. Coders reviewed the questionnaires and child activity booklet to ensure data readability for transfer into an electronic format. Feedback was provided to field staff, where necessary, based on these reviews. The visual review included checking all forms for appropriate signatures, ensuring that all instruments were complete, and checking that all necessary items were included in the case folder. Once this review process was completed, the receipt clerks sent the instruments to the data entry staff to be transferred to a computer-assisted data entry (CADE) program. The data entry staff performed a second visual review which included changing (“upcoding”) any “Other, specify” responses that actually fit within the available response categories of the question. Additional detail on these steps is provided below.

The NCATS tapes that were receipted were sent to the NCATS coding staff to be coded onto NCATS forms. After the tapes had been coded, the completed NCATS forms were forwarded to data entry staff and underwent the same data entry and review procedures as the hard-copy instruments.

6.2.1 Receipt Control

In the first data collection wave, more than 80,000 documents were expected. The project-specific HOMER developed in the field test was used to monitor them. HOMER was initially loaded with the identifying information, such as identification numbers for each child and the instruments that were expected for each child in the sample. As data were collected in the field, field interviewers completed transmittal forms for each case indicating case identification number and which questionnaires and/or case instruments were to be mailed to the home office. Once data collection started, receipt control clerks compared the transmittal forms to the questionnaires and other instruments included in the case folder for accuracy and completeness. Feedback was then given to the field staff, where appropriate. The identification number on each form was matched against the identification numbers in HOMER to verify that the appropriate number of forms for each case were returned. The forms were then logged into HOMER. Once forms were logged in, the data were “caded” (entered with computer assistance) and verified. Occupations were coded and the data were edited. The following sections describe the cading, verification, coding and editing processes for hard-copy questionnaires.

6.2.2 Coding

The hard-copy questionnaires required occupation coding, the review and possible upcoding of all "Other, specify" text responses, and a quick visual review of particular questions in each questionnaire. The quick visual review was to ensure that the questionnaire values were accurate, complete, consistent across variables, and that the numbers were in the appropriate measurement unit prior to converting data to an electronic format. The staff were trained on the coding procedures; coding manuals supported the process. Senior coders verified all coding differences and "Other, specify" changes. For this first round of data collection, the verification rate was set at 100 percent for each coder.

6.2.2.1 Review of "Other, specify" Items

The self-administered questionnaires contained 12 "Other, specify' items. These "Other, specify" text responses were reviewed by the data editing staff, and where appropriate, upcoded into one of the existing response categories. The small number of "specify" responses that remained after upcoding did not fit into any preexisting category and were of insufficient numbers to warrant an additional category. No new codes were added.

6.2.2.2 Parent Occupation Coding

The occupations of the nonresident and/or resident father in the nonresident father or resident father questionnaires were coded using the procedures described above in section 6.1.3.2

6.2.2.3 Physical Measurement Coding

The physical measurements consisted of length, weight, and middle upper arm circumference (MUAC) for all children. In addition, for children who were very low birth weight, head circumference was obtained. Each of these measurements was obtained twice as a way to ensure accuracy. Acceptable ranges were set for each of these measurements. If a measurement fell outside the acceptable range, it was referred to the data manager to be hand coded.

6.2.3 Data Entry

Westat data entry staff entered the data from each instrument in the CADE program. The data were entered a second time at the rate of 100 percent for the first 2 weeks to verify, and subsequently at 10 percent once data entry staff achieved a less than 5 percent error rate. New staff, once trained, would start at 100 percent verification until they attained an error rate of below 5 percent. The results of the two passes were compared and differences identified. The hard-copy form was pulled and examined to determine what corrections had to be made to the data. These corrections were re-entered resulting in an accuracy rate exceeding 99 percent.

6.2.4 Data Editing

The data editing process consisted of running range edits for soft and hard ranges, running consistency edits, and reviewing frequencies of the results.

6.2.4.1 Range Specifications

Hard-copy range specifications set the parameters for high and low acceptable values for a question. Ranges were based on data collected from ECLS-K, the ECLS-B field test, and other studies. For open-ended questions, such as, “What was your age at the time your first child was born?”, high and low ranges were established as acceptable values. Data frequencies on the range of values were reviewed to identify any errors. Values outside the range were identified as errors and were printed for a data editor to review. Cases with range errors were identified, and the original response was updated. Data frequencies were then rerun and reviewed. This iterative process was repeated until no further range errors were found.

6.2.4.2 Consistency Checks (Logical Edits)

By programming logical edits between variables, consistency between variables not involved in a skip pattern was confirmed. For example, in the resident father questionnaire, the number of biological or natural children living outside the household could not be greater than the number of biological or natural children the father reported having. These logical edits were run on the whole database after all data entry and range edits were complete. The logical edits were run separately for each form. All of the data were combined into one large data file, and data frequencies were produced. The frequencies were reviewed to ensure the data remained logically consistent within the form. When an inconsistency was found, the case was identified, and the inconsistency was printed on paper for an editor to review. The original value was replaced with a corrected value, and the case was then rerun through the consistency edits. Once the case passed the consistency edits, it was appended back into the main data set. The frequencies were then rerun and reviewed. This was an iterative process; it was repeated until no further inconsistencies were found.

6.2.4.3 Frequency and Cross-Tabulation Review

Frequencies and cross-tabulations were run to determine consistency and accuracy across the various forms and matched against the data in the Field Management System. If discrepancies could not be explained, no changes were made to the data. For example, in the resident father questionnaire, an item asked about “how much is earned before taxes and other deductions.” If the father gave a response of $60,000 per year, but the parent respondent also answered (in the parent CAPI instrument) that the father earns $120,000 per year, then the response was left as recorded by the father because the discrepancy could not be resolved.

REFERENCE

Executive Office of the President, Office of Management and Budget (2000). *Standard Occupational Classification Manual* (NTIS No. SOC-HB-2). Springfield, VA: National Technical Information Service.

7. DATA FILE CONTENT AND COMPOSITE VARIABLES

This chapter describes the content of the Early Childhood Longitudinal Study, Birth Cohort (ECLS-B) Nine-Month Public-Use Data File and focuses largely on the composite variables that have been created.

There is one child-level 9-month public-use data file or catalog, as noted in chapter 1. The 9-month file contains one record for each child. Only cases with a completed parent computer-assisted personal interview (CAPI) are included in the data file, with a total number of 10,688 cases in the file. Most of these cases also will have a completed child assessment (n = 10,221). In addition, a subset of cases with a completed parent interview and child assessment also will have a completed father self-administered questionnaire (n = 6,937). The 9-month composites also are stored in the child catalog.

Each child-level record in the 9-month public-use data file contains data from various respondents associated with the child (the child himself/herself, the primary caregiver, one or more fathers),[1] as well as data from the birth certificate record. For example, a child record will include data from the parent interview (parent CAPI instrument and parent self-administered questionnaire), the resident father questionnaire (formally titled "Questionnaire for Fathers and Other Important People"), the child assessments, and the birth certificate record, if these components are available. For households with sampled twins, a separate child-level record for each twin in the twin pair is included in the data file.

The file, named Child9mP.dat, is stored in the root directory of the CD-ROM as an ASCII file. However, it is strongly recommended that users access the data using the electronic code book (ECB) software available on the CD-ROM rather than accessing the ASCII file directly. The ECB allows data users to view variable frequencies, tag variables for extraction, and create the SAS, SPSS for Windows, or Stata code needed to create an extract each file for analysis. The child-level data file on the ECB is referred to as a "catalog." Instructions for using the CD-ROM and ECB are provided in chapter 8. Appendix B on the CD-ROM contains the record layout for the child catalog.

This chapter is divided into seven sections. Sections 7.1 through 7.3 focus on the description of identification variables, missing values, and variable names. Section 7.4 provides details on the birth

[1] Respondents to the father component instruments may not be fathers. Eligible respondents include the father, but also parent figures that are partners of the parent component respondent; some may be female.

certificate data. Section 7.5 includes information about the creation of composite variables, and section 7.6 contains a table of the composite variables. Finally, section 7.7 describes masked variables.

7.1 Identification Variables

The 9-month public-use data file contains a case identification variable (I_ID) that uniquely identifies each child record. The case identification number (I_ID) has six digits and is randomly assigned to each record. If a twin is sampled, there is a separate record for each twin in the twin pair. A twin identification variable (I_TWINID) is used to identify the case identification number of (I_ID) of the twin record. I_TWINID identifies the case identification number of all cases sampled as twins, including those who are no longer in the child's household. Use X1TWIN to determine whether there is data in the file for a twin case.

A crosstabulation of the case identification variable (I_ID) and the twin identification variable (I_TWINID) can be used to identify families with sampled twins, as well as the case identification numbers of the two child records associated with the twin pair (see section 7.5.4 for a description of other composite variables that can be used to identify twin cases). If a twin is sampled, a crosstabulation of these two variables will show the case identification of one twin in the I_ID field, and the case identification number of the second twin in the I_TWINID field. For example, the case identification number 032168 may appear in the I_ID field, and case identification number 065719 may appear in the I_TWINID field. In this example, data associated with the one twin can be found by selecting the child record with the case identification number (I_ID) of 032168, and data associated with the second twin can be found by selecting the child record with the case identification number (I_ID) of 065719. For singleton cases, the case identification number of the sampled child will appear in I_ID field, and the I_TWINID field will be empty (system missing).

7.2 Missing Values

Most variables in the ECLS-B data use a standard scheme for missing values. Codes are used to indicate item nonresponse, legitimate skips, and unit nonresponse (see exhibit 7-1).

Exhibit 7-1. Standard missing values codes for the ECLS-B 9-month variables: 2001–02

−1	Not applicable, including legitimate skips
−7	Refused (a type of item nonresponse)
−8	Don’t know (a type of item nonresponse)
−9	Not ascertained (a type of item nonresponse)
(blank)	System missing, including unit nonresponse

SOURCE: U.S. Department of Education, National Center for Education Statistics, Early Childhood Longitudinal Study, Birth Cohort, Nine-Month Data Collection, 2001–02.

The “Not applicable” code (−1) has two purposes. Its primary purpose is to indicate that a respondent did not answer the question due to skip instructions within the instrument or external reasons that led a respondent not to participate. In the parent CAPI instrument, where the parent or guardian was a respondent, a “Not applicable” is coded for questions that were not asked of the respondent because of a previous answer given. For example, the parent CAPI instrument question about whether the twin is living in the household (IN010) is not asked when the case is not part of the twin sample. A “Not applicable” code is also used in the direct child assessments. For example, a child’s head circumference was measured only when the child had a very low birth weight (less than 1,500 grams). If a child had a normal or moderately low birth weight, then the measurement of the child’s head circumference is coded “Not applicable.” For the parent and father data where the instruments are self-administered, a “Not applicable” is coded for questions that the respondent left blank because the written directions instructed them to skip the question due to a certain response on a previous question.

Another use of the “Not applicable” code is the circumstance in which it is not known whether a respondent would have answered a question series following a lead question. One example of this use of “Not applicable” is question 40 on the resident father questionnaire. Question 40 asks whether the respondent has a high school diploma or equivalent, such as a GED. If the answer is “Yes,” the questionnaire skips to question 41, which asks whether the respondent has a high school diploma or a GED. If the answer is “No,” the questionnaire skips to question 42. If question 40 was left blank by the respondent, question 41 is coded “Not applicable.”

The “Refused” code (−7) indicates that the respondent specifically told the interviewer that he or she would not answer the question. This, along with the “Don’t know” code and the “Not ascertained” code, indicates item nonresponse. The “Refused” code rarely appears in the parent and father data where the instruments are self-administered because it indicates the respondent specifically wrote something on the questionnaire indicating an unwillingness to answer the question.

The “Don’t know” code (−8) indicates that the respondent specifically told the interviewer that he or she does not know the answer to the question (or in rare cases on the self-administered questionnaires, “I don’t know” was written in for the question). For questions where “Don’t Know” is one of the options explicitly provided (e.g., questions 81 through 83 of the resident father questionnaire), a “−8” will not be coded for those that chose this option; instead the “Don’t know” response will be coded as indicated in the value label information for that question.

The “Not ascertained” code (−9) indicates that the respondent left a question blank that he or she should have answered. For the parent and father self-administered questionnaires, this is the primary code for item nonresponse. For data outside the self-administered questionnaires (e.g., direct assessment scores), a “−9” means that a value was not ascertained or could not be calculated due to nonresponse. For example, as noted in section 3.1.2, a value of “−9” is assigned to a Nursing Child Assessment Teaching Scale (NCATS) subscale composite score if more than 25 percent of the items that make up the NCATS subscale are missing.

“System missing” appears as a blank when viewing code book frequencies and in the ASCII data file. System missing codes (blanks) in the 9-month data file indicate that an entire instrument or assessment is missing due to unit nonresponse. An example of “System missing” is the child’s father not completing the resident father questionnaire. In this case, all questions from the resident father questionnaire will be blank (System missing). These may be translated to another value when the data are extracted into specific processing packages. For instance, SAS will translate these blanks into periods (“.”) for numeric variables.

Depending on the research question being addressed, cases with missing values may need to be recoded. It is advised that users crosstabulate all lead questions (e.g., question CC015 in the parent CAPI instrument asks whether the child received child care from a relative) and followup questions (e.g., question CC070 in this instrument asks about the hours of care the child received from a relative) before proceeding with any recodes or use of the data. It is important to do this because some cases may have missing values for certain components that are of interest to the user.

Missing values for composite variables were coded using the following coding rules. If a particular composite was inappropriate for a given household—as the variable X1MOMID (the household roster number for the resident mother) was for a household with no resident mother—that variable was given a value of “−1” (Not applicable). In instances where a variable was appropriate, but complete information to construct the composite was not available, the composite was given a value of “−9” (Not ascertained). The “Refused” and “Don’t know” codes were not used for the composites.

7.2.1 Exceptions to the Standard Scheme for Missing Values

It should be noted that some variables on the 9-month file are exceptions to the standard scheme for missing values described above. These variables include the Bayley Short Form–Research Edition (BSF-R) scores and the birth certificate data. Missing values for the BSF-R scores are coded as System missing codes (blanks) in the data file. For the birth certificate variables included in the file, there are various codes associated with missing data. For example, several different ways in which missing data are labeled include “Unknown,” “Unknown or not stated,” “Not stated,” “Not ascertained,” “Not reported,” “Not on certificate,” and “Not classifiable.” In addition, the codes associated with these various labels vary among birth certificate variables. For example, “Unknown or not stated” may take on a value of “9” for one birth certificate variable, and a value of “−9” for another. Therefore, it is important for users to carefully note how missing values are handled for different birth certificate variables when conducting analyses of these data. For more information on missing data for birth certificate variables, users are directed to the National Center for Health Statistics (NCHS) web site (http://www.cdc.gov/nchs/births.htm#micro-data) for technical documentation of birth certificate variables on the NCHS natality file. In particular, users should consult the public-use data file documentation, *2001 Natality Technical Appendix*, and the section on “Classification of Data” for a description of the individual items on the standard birth certificate (U.S. Department of Health and Human Services, National Center for Health Statistics, Revised February 2003).

7.2.2 Other Special Codes

Selected variables in the ECLS-B 9-month data set include special codes that are used when the answer to a question does not fit into the standard response scheme. For example, in response to the question “How much weight did you gain during pregnancy?”, mothers who report that they lost weight during pregnancy were given the code 995. Exhibit 7-2 lists the variables that include special codes, indicates the code used, and describes the code's meaning. Users should review these codes and decide how they want to handle them in their analyses.

Exhibit 7-2. Special codes used in the ECLS-B 9-month data collection: 2001–02

Variable name	Description	Special code	Meaning of code
P1AGFORM	P1 CD030 Age first fed formula (mths)	995	Child never fed formula
P1AGFINF	P1 CD045 Age first fed finger food (mths)	995	Child has not yet been given finger food
P1AGFOOD	P1 CD055 Age when first fed solid food (mths)	995	Child has not yet started eating solid food
P1AGSIT	P1 CD080 Age when first sat alone without support (mths)	95	Child not yet able to sit alone without support
P1AGCRWL	P1 CD090 Age when started crawling (mths)	95	Child has not crawled yet
P1AGSTND	P1 CD100 Age pulled self to stand (mths)	95	Child has not pulled self to standing yet
P1AGWALK	P1 CD110 Age walking with help (mths)	95	Child has not walked with help yet
P1NMVS3T	Number of prenatal visits in third trimester	95	Respondent gave birth before 3rd trimester
P1GNPRG	P1 FH057 Weight gain during pregnancy	995	Respondent lost weight during pregnancy
P1BFSEEC	P1 BF165 Number of times father saw child in last 3 months	95	Father has seen child every day or almost every day in last 3 months
P1AGEBM	P1 RI020 Age of respondent's biological mother	1	Mother no longer living
P1AGEBF	P1 RI045 Age of respondent's biological father	1	Father no longer living

SOURCE: U.S. Department of Education, National Center for Education Statistics, Early Childhood Longitudinal Study, Birth Cohort (ECLS-B), Nine-Month Data Collection, 2001–02.

7.3 Variable Naming Conventions

All variables on the 9-month data file include a combination of letters and numbers of up to eight characters. The first two characters (referred to here as prefixes) of the variable names indicate the data source (e.g., parent CAPI instrument, resident father questionnaire) and the data collection point (e.g., round 1). For example, the variable P1PARTNR (question FS030 in the parent CAPI instrument) asks the respondent whether she or he has a spouse or partner who lives in the household. The variable name prefix (P1) indicates both the source of the question ("P" for parent CAPI instrument) and the data collection point ("1" for round 1). In another example, the variable F1RELCH (question 1 of the resident father questionnaire) asks the respondent to report his or her relationship to the focal child. The variable name prefix (F1) indicates the source of the question ("F" for resident father questionnaire) and the data collection point ("1" for round 1).

Exhibit 7-3 lists the variable prefixes that are used throughout the catalog. The prefixes listed here are in two categories: (1) 9-month variables, and (2) weights. In general, variable names start with the following prefixes.

Exhibit 7-3. Prefixes for 9-month variables and weights: 2001–02

9-Month Variables	
P1	Data collected/derived from 9-month parent CAPI instrument and parent self-administered questionnaire
C1	Data collected/derived from 9-month direct child assessments
F1	Data collected/derived from 9-month resident father (and other important people) questionnaire
N1	Data collected/derived from 9-month nonresident father questionnaire
R1	Data collected/derived from 9-month interviewer remarks questionnaire (IRQ) and interviewer observation of the child and home setting
BC	Data collected/derived from the birth certificate
X1	9-month composite variables
Weights	
W1R	Child-level weight variable for parent data from the 9-month wave
W1C	Child-level weight variable for combined child/parent data from the 9-month wave
WIF	Child-level weight variable for combined father/parent data from the 9-month wave
W1FC	Child-level weight variable for combined father/child/parent data from the 9-month wave

SOURCE: U.S. Department of Education, National Center for Education Statistics, Early Childhood Longitudinal Study, Birth Cohort, Nine-Month Data Collection, 2001–02.

7.4 Birth Certificate Data

As noted previously, the ECLS-B used birth certificates of live births in 2001 as the sampling frame (see chapter 4 for more information). This sampling procedure provided access to data on prenatal and neonatal characteristics at the time of the child's birth. Children's birth certificates contained a variety of useful data about the children, their mothers, and their families. They contained information on the date of birth, birth weight, gender, plurality (e.g., if multiple birth and birth order), and prematurity as well as on parents' education, parents' race and ethnicity (including Hispanic origin), and mother's marital status. Birth certificates also included information on the mother's pregnancy history, prenatal care, medical and other risk factors during this pregnancy, and complications during labor and birth. Information on the obstetric procedures and delivery method was also available. Health characteristics of children, such as congenital anomalies and abnormal conditions of the baby and the baby's APGAR[2] score, were included.

Selected variables from the birth certificates were included in the data set, and the birth certificate data were appended to the child record[3] of the ECLS-B 9-month restricted-use data file. Birth certificate variables included in the 9-month restricted-use data file are presented in exhibit 7-4. Because of disclosure risk, the majority of the birth certificate items were suppressed in the public-use data file. The presence of birth certificate data in the public-use data file would allow users to match records in this file with 2001 birth certificate data that are publicly available from the National Center for Health Statistics (NCHS). Therefore, in order to prevent matching of records between these files, all data except the following were removed from the public-use data file: APGAR score, whether the birth was vaginal or by cesarean section, mother's age, mother's education, mother's marital status, and child's birth weight. Among these data, the latter four are composites of survey and birth certificate data. Furthermore, the data on mode of delivery are collapsed from the original birth certificate data, which gave additional details such as whether forceps were used, whether there had been a previous cesarean section, and so forth. The suppression of these details was also done for confidentiality reasons, to prevent matching the ECLS-B survey data with the NCHS birth certificate data. See table 7-7 for a list of variables that were modified or suppressed.

---

[2] The acronym "APGAR" represents the five qualities upon which newborns are rated: Appearance, Pulse, Grimace, Activity, and Respiration.

[3] The birth certificate is an external document that the National Center for Education Statistics (NCES) did not administer, collect, or transcribe. In some instances, the birth certificate variables provided by the National Center for Health Statistics may be discrepant from similar variables collected by NCES (e.g., whether the mother had an ultrasound during pregnancy, or whether the mother drank alcohol during pregnancy). This could be due to the fact that they were administered under different circumstances and may have had different procedures for administering, collecting, and/or transcribing data. It could also be due to the fact that any data collection has measurement error that can result in discrepancies between variables that seemingly would have similar values.

Variables in the data file that come from the birth certificate are identified with the variable name prefix BC. It is important to note that there are 80 cases in the data file for which birth certificate data are missing. For these cases, the birth certificate variables are “System missing” (blanks) in the data file. For 70 of these cases, the birth certificate data are missing because the children were sampled from hospital records rather than birth certificates (two states did not permit sampling children from birth certificates). Therefore, the birth certificate data were not available for this group of children. For 10 other cases, the birth certificate data are missing because the birth certificate numbers sampled from initial birth certificate data received from the states did not later match birth certificate numbers in the final, edited natality file produced by NCHS. This could occur, for example, if two birth certificates were initially inadvertently filed for a single baby, and one was later voided from the file as a duplicate record.

Users also should be aware that there are various codes associated with missing data for birth certificate variables included in the file. For example, several different ways in which missing data are labeled include “Unknown,” “Unknown or not stated,” “Not stated,” “Not ascertained,” “Not reported,” “Not on certificate,” and “Not classifiable.” Furthermore, the codes associated with these various labels vary among birth certificate variables. For example, “Unknown or not stated” may take on a value of “9” for one birth certificate variable, and a value of “−9” for another. It is also important to note that some birth certificate variables have higher numbers of missing values than others because not all states collect all information on the birth certificate. Therefore, users should carefully note how missing values are handled for different birth certificate variables when conducting analyses of these data. For more information on missing data for birth certificate variables, users are directed to the NCHS web site (http://www.cdc.gov/nchs/births.htm#micro-data) for technical documentation of birth certificate variables on the NCHS natality file. In particular, users should consult the public-use data file documentation, *2001 Natality Technical Appendix,* and the section on “Classification of Data” for a description of the individual items on the standard birth certificate (U.S. Department of Health and Human Services, National Center for Health Statistics, Revised February 2003).

Exhibit 7-4. ECLS-B 9-month birth certificate (BC) data: 2001–02

Variable name	Variable label
Date of birth	
BCDOBYY	BC year of birth
BCDOBMM	BC month of birth
BCDOBDD	BC day of birth
Characteristics of place of birth	
BCRESSTS	BC resident status
BCPLDEL	BC place or facility of delivery
BCBRTATT	BC attendant at birth
BCSTOCFP	BC state of occurrence (FIPS)
BCCNOCFP	BC county of occurrence (FIPS)
BCCNOCSZ	BC population size-occurrence county
BCSTRSFP	BC state of residence (FIPS)
BCCNRSFP	BC county of residence (FIPS)
BCCNRSSZ	BC population size-residence county
Demographic characteristics of mother	
BCMOMAGE	BC age of mother
BCMOMHSP	BC Hispanic origin of mother
BCMOMRCI	BC race of mother imputation flag
BCMOMRC	BC race of mother
BCMOMED	BC education of mother detail
BCMOMMAI	BC mother marital status imputed
BCMOMMAR	BC marital status of mother
BCMBRPLC	BC place of birth of mother recode
Maternal/neonatal history	
BCADQCAR	BC care adequacy recode – Kessner index
BCNLBNL	BC number of live births, now living
BCNLBND	BC number of live births, now dead
BCNOTERM	BC number of other terminations
BCMONPRE	BC month pregnancy prenatal care began
BCNMPRVS	BC total number of prenatal visits
BCLMPMM	BC month last normal menses began
BCLMPDD	BC day last normal menses began
BCLMPYY	BC year last normal menses began
Demographic characteristics of father	
BCFTHAGE	BC age of father
BCFTHHSP	BC father of Hispanic origin
BCFTHRC[1]	BC race of father

See notes at end of exhibit.

Exhibit 7-4. ECLS-B 9-month birth certificate (BC) data: 2001–02—Continued

Variable name	Variable label
Gestation	
BCGESTMM	BC clinical gestation estimate used flag
BCGSTMMI	BC gestation imputation flag
BCGESTWK	BC gestation-detail in weeks
BCCLNGST	BC clinical estimate of gestation
Child characteristics	
BCSEX	BC sex
Plurality	
BCPLURI	BC plurality imputation flag
BCPLURAL	BC plurality
Child health status at birth	
BCFMAPS	BC five minute APGAR score
BCBRTHWT[2]	C1 child birth weight (grams-birth cert)
Method of delivery	
BCVAGINL	BC delivery-vaginal
BCVGBACS	BC delivery-vaginal after prev c-section
BCPRIMAC	BC delivery-primary c-section
BCREP_C	BC delivery-repeat c-section
BCFORCEP	BC delivery-forceps
BCVACUUM	BC delivery-vacuum
Medical risk factors for pregnancy	
BCANEMIA	BC med risk factor-anemia
BCCARDIC	BC med risk factor-cardiac disease
BCLUNG	BC med risk factor-acute/chron lung dis
BCDIABTS	BC med risk factor-diabetes
BCHERPES	BC med risk factor-genital herpes
BCHYDRA	BC med risk factor-(oligo)hydramnios
BCHEMO	BC med risk factor-hemoglobinopathy
BCCHYPER	BC med risk factor-hypertension/chronic
BCPHYPER	BC med risk-hypertension dur preg
BCECLAMP	BC med risk factor-eclampsia
BCINCRVX	BC med risk factor-incompetent cervix
BCPRV4KG	BC med risk-previous 4000+ gram
BCPRVPTM	BC med risk factor-prev preterm or small
BCRENAL	BC med risk factor-renal disease
BCRH_SNS	BC med risk factor-rh sensitization
BCUTERIN	BC med risk factor-uterine bleeding
BCOTHMR	BC med risk factor-other

See notes at end of exhibit.

Exhibit 7-4. ECLS-B 9-month birth certificate (BC) data: 2001–02—Continued

Variable name	Variable label
Other risk factors for pregnancy	
BCTOBACC	BC tobacco use during pregnancy
BCNUMCIG	BC avg number cigs per day during preg
BCALCOHL[3]	BC alcohol use during pregnancy
BCNUMDRK[3]	BC number drnks per week during preg
BCWTGAIN	BC weight gain during pregnancy (pounds)
Obstetric procedures	
BCAMNIO	BC obstetric proc-amniocentesis
BCMONITR	BC obstetric proc-elect fetal monitor
BCINDUCT	BC obstetric proc-labor induced
BCSTIMUL	BC obstetric proc-labor stimulated
BCTOCOL	BC obstetric proc-tocolysis
BCULTRAS	BC obstetric proc-ultrasound
BCOTHOB	BC obstetric proc-other
Complications of labor and/or delivery	
BCFEBRL	BC labor complc-febrile(>100 degree f)
BCMCNM	BC labor complc-meconium, moderate/heavy
BCRUPT	BC labor complc-membrane rupture>12 hours
BCABRPT	BC labor complc-abruptio placenta
BCPREVIA	BC labor complc-placenta previa
BCEXCBLD	BC labor complc-other excessive bleeding
BCSEIZR	BC labor complc-seizures during labor
BCPRECIP	BC labor complc-precip lbr (<3 hours)
BCPROLNG	BC labor complc-prolonged lbr >20 hours
BCDYSFNC	BC labor complc-dysfunctional labor
BCBREECH	BC labor complc-breech/malprst
BCCEPHLO	BC labor complc-cephalopelvic disprprtn
BCCORD	BC labor complc-cord prolapse
BCANESTH	BC labor complc-anesthetic complic
BCDSTRSS	BC labor complc-fetal distress
BCOTHLBR	BC labor complc-other
Abnormal conditions of the newborn[3]	
BCCHANEM	BC newborn cond-anemia hct<39/hgb <13
BCINJURY	BC newborn cond-birth injury
BCCHFAS	BC newborn cond-fetal alcohol syndrome
BCHYALIN	BC newborn cond-hyaline membrane disease
BCMECSYN	BC newborn cond-meconium aspiration syn
BCVENL30	BC newborn cond-assisted ventlation <30m
BCVENM30	BC newborn cond-assisted ventlation>=30m
BCCHSEIZ	BC newborn cond-seizures
BCOTHNC	BC newborn cond-other specify

See notes at end of exhibit.

Exhibit 7-4. ECLS-B 9-month birth certificate (BC) data: 2001–02—Continued

Variable name	Variable label
Congenital anomalies of the child[3]	
BCANECPH	BC congen anomaly-anencephalus
BCSPINA	BC congen anomaly-spina bifida
BCHYDCPH	BC congen anomaly-hydrocephalus
BCMCRCPH	BC congen anomaly-microcephalus
BCNRVS	BC congen anomaly-other centrl nerv.
BCHEART	BC congen anomaly-heart malformations
BCCIRCUL	BC congen anomaly-other circl/resp
BCRECTAL	BC congen anomaly-rectal atresia/stnosis
BCTRACHE	BC congen anomaly-tracheo/esophageal
BCOMPHL	BC congen anomal-omphalocele-gastrosch
BCGASTRO	BC congen anomaly-other gastrointest
BCGENITL	BC congen anomal-malformed genitalia
BCRNLAGE	BC congen anomaly-renal agenesis
BCUROGEN	BC congen anomaly-other urogenital
BCCLFTLP	BC congen anomaly-cleft lip/palate
BCACTYLY	BC congen anomaly-(poly/syn/a)dactyly
BCCLUBFT	BC congen anomaly-club foot
BCHERNIA	BC congen anomaly-diaphragmatic hernia
BCMUSCLO	BC congen anomal-other musculoskeletal
BCDOWNS	BC congen anomaly-Downs syndrome
BCCHROM	BC congen anomaly-other chromosomal
BCOTHCON	BC congen anomaly-other no category

[1] Father race is often missing, and the NCHS does not impute for these missing data.
[2] This birth certificate variable is included in the data file with the other physical measurements collected for the child assessment.
[3] Users should be aware that these variables tend to be underreported. For more information concerning the quality of the birth certificate data, users are directed to the National Center for Health Statistics (NCHS) web site and the following two documents: (1) the *2001 Natality Technical Appendix* (U.S. Department of Health and Human Services, National Center for Health Statistics, Revised February 2003), available at http://www.cdc.gov/nchs/data/techap01.pdf and (2) the NCHS annual report *Births, Final Data, 2001* (U.S. Department of Health and Human Services, National Center for Health Statistics, 2002), available at http://www.cdc.gov/nchs/data/nvsr/nvsr51/nvsr51_02.pdf.
SOURCE: U.S. Department of Education, National Center for Education Statistics, Early Childhood Longitudinal Study, Birth Cohort, Nine-Month Data Collection, 2001–02.

7.5 Composite Variables

To facilitate analysis of the survey data, composite variables were created and added to the child data file. Most composite variables were created using two or more variables, each of which is named in the text that explains the composite variable. Other composite variables are recodes of single variables. In addition, certain composite variables were created using data from the Field Management System (FMS) that are not included in the data file.

Composite variables based on the child assessment include physical measurements (child length, weight, middle upper arm circumference, and head circumference), scores for the NCATS, and scores for the BSF-R. Variables constructed from the parent CAPI instrument and the resident father questionnaire include parent identifiers, parent demographics, household composition, household income and poverty, child care, and child demographics. Variables developed from the nonresident father questionnaire include nonresident father education and employment status.

In section 7.6, all the composite variables included on the 9-month data file are listed. All basic child demographic variables are presented first, followed by child care variables and variables created from the child assessment. Household composition variables are listed next, followed by demographics for parents (resident mother and father characteristics are followed by nonresident biological father characteristics). Following parent demographic variables, derived and indicator variables are listed in the composite table. Once the user identifies the composites of interest, he or she can refer to exhibit 8-60 for instructions on accessing the variables from the ECB.

7.5.1 Child Composite Variables

There are many child composite variables on the child catalog. As noted earlier, section 7.6 at the end of this chapter lists and describes all of the composites. Some of these composite variables that may be of particular interest to users are described in further detail here. A detailed description of composite variables developed for the NCATS and the BSF-R child assessment components is included in chapter 3.

7.5.1.1 Child's Age in Months at Assessment (X1ASAGE)

On the 9-month data file, child's age at the time of the child assessment is defined in two different ways in composite variables X1ASAGE and X1BSIDAG. The composite variable X1ASAGE defines child's age at assessment in terms of the child's chronological age in decimal months.[4] For most analyses, X1ASAGE is the child's age variable that analysts should use. The second composite variable, X1BSIDAG, defines the child's age at assessment adjusted for prematurity and is also described later in this section.

The composite variable X1ASAGE was calculated as follows. If the child's date of birth on the birth record was determined to be correct by the parent respondent, then the child's age was calculated by determining the number of days between the date when the child completed the ECLS-B direct child assessments and the child's date of birth indicated on the birth record. If the child's date of birth on the birth record was determined to be incorrect by the parent respondent, then the child's age was calculated by determining the number of days between the date when the child completed the direct child assessment and the child's date of birth reported in the parent CAPI instrument. The total number of days was then divided by 30 to calculate the age in months.

The date of the direct child assessments came from the following source variables in order of priority: (1) the assessment date in the Child Activity Booklet (C1V1DATE, C1V2DATE), (2) the assessment date on the front cover of the Child Activity Booklet (C1CADATE), (3) the date of the parent interview from the FMS, and (4) the date of the child assessments from the FMS. If there was more than one child assessment visit, the child's assessment age was calculated as the average of the child's assessment age calculated from the first home visit and the child's assessment age calculated from the last home visit of the 9-month data collection.

7.5.1.2 Child's Age in Months at Assessment Adjusted for Prematurity (X1BSIDAG)

For X1BSIDAG, the child's chronological age is adjusted for children born 3 or more weeks premature and is also presented in decimal months. X1BSIDAG is calculated in the parent CAPI application and is upcoded from the parent CAPI application for inclusion in the child data file as a

[4] The decimal month is calculated by dividing the child's age in days by the average number of days in a month (i.e., 365/12) and rounding the resulting number to one decimal place. For example, a child who is 270 days old (approximately 9 months old) will have a decimal age of 8.9 months (round [270 / (365/12)] = 8.9). The range of this variable is from 6.2 to 22.3.

composite variable. The child's age adjusted for prematurity is calculated as a function of the child's date of birth, the date of the child assessment, and the child's prematurity. Babies born at or after 37 weeks were considered full-term, babies born before 37 weeks were considered preterm, and babies born before 32 weeks were considered very premature. If there was more than one child's age at assessment adjusted for prematurity calculated in the parent CAPI application as a function of more than one home visit, then X1BSIDAG is calculated as the average of the first and the last child's age at assessment adjusted for prematurity calculated in the parent CAPI application. In general, X1BSIDAG is not recommended for use in analyses. X1BSIDAG is included in the data file primarily for analysts who are interested in creating developmental index scores for the BSF-R. For more detailed information regarding the calculation of X1BSIDAG in the parent CAPI application, users are referred to *Early Childhood Longitudinal Study, Birth Cohort (ECLS-B), Methodology Report for the Nine-Month Data Collection, Volume 1: Psychometric Characteristics* (U.S. Department of Education, National Center for Education Statistics, forthcoming).

7.5.1.3 Child's Race/Ethnicity (X1CHAMIN, X1CHASN, X1CHPCIL, X1CHBLCK, X1CHWHT, X1CHHISP, X1CHMLRC, X1CHRACE)

The composites for the child's race/ethnicity are presented in the 9-month data file in two ways: (1) as dichotomous variables for each race/ethnicity category (American Indian (X1CHAMIN), Asian (X1CHASN), Pacific Islander (X1CHPCIL), Black (X1CHBLCK), White (X1CHWHT), Hispanic (X1CHHISP), and more than one race, unspecified (X1CHMLRC)) from the parent CAPI instrument data; or (2) as a single race/ethnicity composite taken from the parent CAPI instrument data (X1CHRACE).

Parent respondents were allowed to indicate that the child belonged to one or more of 14 race categories. These categories include (1) White, (2) Black or African American, (3) American Indian or Alaska Native, (4) Asian Indian, (5) Chinese, (6) Filipino, (7) Japanese, (8) Korean, (9) Vietnamese, (10) Other Asian, (11) Native Hawaiian, (12) Guamanian or Chamorro, (13) Samoan, and (14) Other Pacific Islander. From these responses, a series of five dichotomous race variables were created that indicated separately whether the child belonged to each of five main specified race groups, including White, Black, Asian (including Asian Indian, Chinese, Filipino, Japanese, Korean, Vietnamese, and Other Asian), and Native Hawaiian or other Pacific Islander (including Native Hawaiian, Guamanian or Chamorro, Samoan, and Other Pacific Islander). In addition, one more dichotomous variable was created for parent respondents who had simply indicated that the child was multiracial without specifying a race (e.g., biracial). The retention of the dichotomous variables in the file allows users to analyze children's race in a variety of ways.

Data were collected on ethnicity as well. Specifically, respondents were asked whether the child was of Hispanic or Latino origin. Using the six race dichotomous variables and the Hispanic ethnicity variable, the race/ethnicity composite variable (X1CHRACE) was created. The categories were White, non-Hispanic; Black or African American, non-Hispanic; Hispanic, race specified; Hispanic, no race specified; Asian, non-Hispanic; Native Hawaiian or other Pacific Islander, non-Hispanic; American Indian or Alaska Native, non-Hispanic; and more than one race specified, non-Hispanic. A child was classified as Hispanic if a parent respondent indicated the child's ethnicity was Hispanic regardless of whether a race was identified and what that race was.

It should be noted that for any given analysis, users will need to decide how they want to define race/ethnicity. For example, as discussed in section 4.1.1, 873 completed cases were classified as American Indian for sampling purposes, but only 286 were classified as American Indian for the composite of child's race (X1CHRACE). Users should further consult section 4.1.1 for more information regarding differences between sampling variables and data that were collected in the parent CAPI application regarding race/ethnicity.

7.5.1.4 Child's Length (X1CHLENG)

To obtain good measurements, each child's length was measured twice. For the length composite X1CHLENG, if the two length values (i.e., C1CHLGT1 and C1CHLGT2) were less than 5 percent apart (calculated as C1CHLGT1 – C1CHLGT2 / C1CHLGT1 X 100), the average of the two length values was computed and used as the composite value. Otherwise, if the two length values were equal to or greater than 5 percent apart, the length value that was closest to the weighted average length for the group of children in the ECLS-B 9-month sample of the same age (in months) and birth weight status (normal, moderately low, and very low birth weight) was used as the composite value. If both child length measurement values from the child assessment were missing, then the composite was set to −9 (Not ascertained).

7.5.1.5 Child's Weight (X1CHWGHT)

Each child's weight was also measured twice.[5] For the weight composite (X1CHWGHT), if the two weight values (i.e., C1CHWGT1 and C1CHWGT2) were less than 5 percent apart (calculated as C1CHWGT1 – C1CHWGT2 / C1CHWGT1 X 100), the average of the two values was computed and used as the composite value. Otherwise, if the two weight values were equal to or greater than 5 percent apart, the weight value that was closest to the weighted average weight for children in the ECLS-B 9-month sample of the same age (in months) and birth weight status (normal, moderately low, and very low birth weight) was used as the composite value. If both weight values from the child assessment were missing, then the composite is set to −9 (Not ascertained).

7.5.1.6 Child's Middle Upper Arm Circumference (X1MUAC)

Each child's middle upper arm circumference (or MUAC) was also measured twice. For the middle upper arm circumference composite (X1CHMUAC), if the two MUAC values (i.e., C1MUAC1 and C1MUAC2) were less than 5 percent apart (calculated as C1MUAC1 – C1MUAC2 / C1MUAC1 X 100), the average of the two values was computed and used as the composite value. Otherwise, if the two MUAC values were equal to or greater than 5 percent apart, the MUAC value that was closest to the weighted average MUAC for the group of children in the ECLS-B 9-month sample of the same age (in months) and birth weight status (normal, moderately low, and very low birth weight) was used as the composite value. If both MUAC values from the child assessment were missing, then the composite is set to –9 (Not ascertained). See section 3.1.4 for how the measurement was obtained (multistep process with length of arm and midpoint measurements) and specific analytic concerns.

7.5.1.7 Child's Head Circumference (X1CHCRFM)

For very low birth weight children only, each child's head circumference was measured twice. For the head circumference composite (X1CHCRFM), if the two circumference values of the child's head (i.e., C1CHHC1 and C1CHHC2) were less than 5 percent apart (calculated as C1CHHC1 – C1CHHC2 / C1CHHC1 X 100), the average of the two values was computed and used as the composite

[5] It is important to note that child's weight was not measured the same way for all children. Some children (less than one percent) were not weighed with an adult, but were weighed alone on the scale. In addition, there is one case where information about the child's weight from a recent doctor's visit was used instead of a weight measurement taken by the interviewer during the child assessment.

value. Otherwise, if the two circumference values of the child’s head were equal to or greater than 5 percent apart, the head circumference value that was closest to the weighted average head circumference for the group of children in the ECLS-B 9-month sample of the same age (in months) was used as the composite value. For very low birth weight children, if both circumference values of the child’s head from the child assessment were missing, then the composite is set to –9 (Not ascertained). For children with a normal or moderately low birth weight, the composite is set to –1 (Not applicable).

7.5.1.8 Primary Nonparental Child Care Arrangement (X1PRIMNW)

There are several composite variables in the file that can be used to describe children’s child care arrangements based on information from the parent CAPI instrument. One of these composites (X1PRIMNW) was created to indicate the primary, nonparental, individual child care arrangement in which the child spent the most hours per week at the time of the 9-month home visit. The values for this variable are as follows:

- 0 = No nonparental care
- 1 = Relative care in child’s home
- 2 = Relative care in another home
- 3 = Relative care, location varies
- 4 = Nonrelative care in child’s home
- 5 = Nonrelative care in another home
- 6 = Nonrelative care, location varies
- 7 = Center-based program
- 8 = Multiple care arrangements with same hours in each

To obtain the composite, the hours for relative care (CC070), nonrelative care (CC170), and center-based care (CC260) were compared to select the primary care arrangement with the most number of hours. If the number of hours of either relative or nonrelative care (given in CC070 and CC170) was higher than hours of center-based care, the variable indicating location of care (child’s home/other home/varies) for that type was examined using parent CAPI instrument items CC045 and CC145. If location of care was missing, then X1PRIMNW was coded as −9 (Not ascertained); if location of care was not missing, then X1PRIMNW was coded 1, 2, 3, 4, 5, or 6 depending on the type with the most hours (relative/nonrelative) and the location of care (child's home/other home/varies). Otherwise, if the number of hours of care in center-based programs (CC260) was higher than for relative or nonrelative care, then X1PRIMNW was coded as 7. If the number of hours of care was equal for two or more types of care, X1PRIMNW was coded as 8. If the indicator variables for regular receipt of relative, nonrelative, and center-based care (CC015, CC120, and CC220) were all equal to 2 (No regular receipt), then X1PRIMNW was coded as 0 (No nonparental care).

It should be noted that it is possible to have missing data for the primary child care arrangement (X1PRIMNW), but have information on the number of hours of child care a child has (X1HRSCAR). This is because there must be information about the location of care in order to have a valid value for X1PRIMNW.

7.5.1.9 Hours Per Week in Child Care (X1HRSCAR)

Another child care composite variable indicates the number of hours per week the child spent in regularly scheduled nonparent child care. X1HRSCAR indicates the total number of hours per week the focal child spent in all primary and secondary care arrangements at the time of the 9-month parent CAPI instrument. The variable combines hours in child care arrangements in which the child spent the most time with hours from additional regular child care arrangements.

X1HRSCAR was coded as follows. If any of the indicator variables for regular receipt of relative, nonrelative, and center-based care (CC015, CC120, and CC220) were equal to 2 (No regular receipt), the number of hours for that type of care was coded to 0. If the regular receipt of care variables were refused or unknown, then the number of hours for that type of care was coded as −9 (Not ascertained). Also, if the regular receipt variable was coded as 1 (Yes), but the hours given was refused or unknown, then the number of hours for that type of care was coded as −9 (Not ascertained). Otherwise, if the indicator for regular receipt of care was equal to 1 (Yes), and the hours given were greater than or equal to 0, then the number of hours for that type of care was coded as the number of hours given in CC070, CC170, and CC260.

The composite also includes hours spent with additional regularly scheduled providers of care of the same type. This was done to include child care arrangements such as those in which two different relatives cared for the child on a regular basis or two different child care programs were attended. For each type of care, if the care receipt variables indicated no care of that type, or if the number of providers of that type of care (CC035, CC140, and CC240) was equal to 1, then additional hours were coded to 0. Otherwise, if the number of providers or the number of additional hours (CC118, CC218, and CC308) was refused or unknown, then the number of additional hours was coded as −9 (Not ascertained). Otherwise, if the number of additional providers was greater than 1 (CC035, CC140, and CC240), the number of additional hours was coded to equal the appropriate number of additional hours variables in the instrument (CC118, CC218, and CC308).

This process was followed three times, once each for relative, nonrelative, and center-based care. If any of the three primary caregiver hour variables or the three additional hour variables was missing, then the total number of hours was coded as −9 (Not ascertained). Otherwise the total number of hours in regularly scheduled child care was coded as the sum of the six hour variables. If the child did not receive regular care of any type (i.e., all the indicator variables for regular receipt of relative, nonrelative, and center-based care (CC015, CC120, and CC220) were equal to 2), X1HRSCAR is set to –1 (Not applicable).

7.5.1.10 Earliest Age Child Received Nonparental Care (X1AGCARE)

Another child care composite (X1AGCARE) indicates the earliest age in months the child first began any type of nonparental care on a regular basis. A value of 0 on X1AGCARE indicates that the child was less than 1 month of age when he or she first received care. The composite is coded as follows: If all the indicator variables for regular receipt of relative, nonrelative, and center-based care (CC015, CC120, and CC220) were equal to 2 (No regular receipt) and all the indicator variables for ever receiving regular care from a relative, nonrelative, and center-based program (CC020, CC125, and CC225) were equal to 2 (No regular receipt ever), X1AGCARE is coded as −1 (Not applicable).

For the remaining cases, if any of the indicator variables for regular receipt of care of any type (CC015, CC120, and CC220) was equal to 1 (Yes, regular receipt), or if any of the indicator variables for ever receiving care of any type (CC020, CC125, and CC225) was equal to 1 (Yes, regular receipt ever), X1AGCARE is coded as the single youngest age in months the child first received care of any type (CC025, CC130, and CC230). Otherwise, if any of the indicator variables for regular receipt of care or ever receiving care of any type was equal to 1 (Yes), but the age in months the child first received care of that type was missing, X1AGCARE is coded as −9 (Not ascertained).

7.5.2 Family and Household Composite Variables

Many composites were created to capture information about the sampled children's family and household characteristics. These composite variables are generally of two types. Some composite variables describe household-level characteristics such as household composition, the number of siblings of the focal child, the number of household members under the age of 18 years, the number of household members aged 18 years and older, as well as the family's food security status, socioeconomic status, and poverty status. Other family/household composite variables identify the type of mother/female guardian and/or father/male guardian residing in the household and demographic characteristics (e.g., age, race/ethnicity, education level, and occupation) associated with the child's parent(s)/guardian(s).

Several of these composites, which may be commonly used by analysts of the 9-month data, are described below. All of the family and household composites are listed and described in table 7-6.

7.5.2.1 Number of Siblings (X1NUMSIB)

The composite X1NUMSIB indicates the total number of siblings (full, step-, adoptive, or foster) with whom the child lived in the household (FS055 and FS060). Siblings were identified through the respondents' stated relationship of the sibling to the focal child. In addition, any child that was reported to be a child of the focal child's parent/guardian was considered a sibling of the focal child.

7.5.2.2 Parents' and Household Members' Age (X1HFAGE, X1HMAGE, X1LESS18, and X118OVER)

There are several composite variables in the file that refer to the ages of adults and children in the household. These are X1HFAGE (age of resident father), X1HMAGE (age of resident mother), X1LESS18 (total number of people in the household under age 18, including focal child, parents, siblings, and other children), and X118OVER (total number of people in the household age 18 or older, including parents, other adults, siblings, and other children), The resident father may be the birth, adoptive, step-, or foster parent of the focal child, or the partner of the resident mother. The resident mother is similarly defined. The ages of these persons in the household were collected during the 9-month parent CAPI instrument in the household roster.

7.5.2.3 Food Security Status

Food security status of the children's families was assessed based on responses to the 18 food security questions (HF020a through HF080) in the 9-month parent CAPI instrument. The questions measured a wide range of food insecurity and reduced food intake issues. Composites were created based on three scales calculated from these responses: Household Food Security Scale, Adult Food Security Scale, and Children's Food Security Scale. Calculations of the Household Food Security Scale composites and the Adult Food Security Scale composites were carried out in accordance with the standard methods described in *Guide to Measuring Household Food Security, Revised 2000* (U.S. Department of Agriculture, Food and Nutrition Service, 2000). Calculations of the Children's Food Security Scale composites were carried out in accordance with the standard methods described in *Measuring Children's Food Security in U.S. Households, 1995-99* (U.S. Department of Agriculture, Economic Research Service, 2002). The items and the food security scales based on them have been validated using both ethnographic and statistical methods (Hamilton et al., 1997a; Hamilton et al., 1997b; Radimer, 1990; Radimer, Olson, and Campbell, 1990; Radimer et al., 1992; Wehler, Scott, and Anderson, 1992). Analysis of the ECLS-B data conducted by Mark Nord (U.S. Department of Agriculture) using statistical methods based on the Rasch measurement model (Fischer and Molenar, 1995; Wright, 1977; Wright, 1983) indicated that use of the standard benchmark household scores was appropriate.

Food security scales were calculated for all household members (based on both adult-referenced and child-referenced questions), for adults, and for children. For each scale, three measures are presented: a continuous scale score, a categorical score, and a raw score. For scale and raw scores, higher scores indicate more severe food insecurity. The analytic advantages of each type of variable are described below.

7.5.2.4 Food Security Status: Continuous Measures (X1FSSCAL, X1FSADSC, and X1FSCHSC)

X1FSSCAL is the scale score presentation of the Household Food Security items. This is a measure of the severity of food insecurity or hunger experienced in the household in the previous 12 months. It is a continuous, interval-level measure of food insecurity and is appropriate for linear models. This scale score is a Rasch transformation of the raw score (X1FSRAW). Valid values range from 1.4 to 13.0, with higher values indicating more severe food insecurity. Under Rasch-model assumptions, the scale score for families that affirm no items (raw score = 0) is indeterminate. It is less than the lowest measured value (1.4), but its precise value is unknown and may vary substantially among families. X1FSSCAL for such cases is assigned a value of –6. If these cases (a substantial majority of all cases) are included in linear models, appropriate methods must be used to take into account this indeterminacy (see U.S. Department of Agriculture, Food and Nutrition Service, 2000).

X1FSADSC is similar to X1FSSCAL, but is the Adult Food Security scale score. This is a measure of the severity of food insecurity or hunger experienced by adults in the household in the previous 12 months. Valid values range from 1.7 to 11.1, with higher values indicating more severe food deprivation. The scale score is undefined for households that affirmed no adult-referenced items (see discussion of X1FSSCAL above).

X1FSCHSC is the Children's Food Security scale score. This is a measure of the severity of food insecurity or hunger experienced by children in the household in the previous 12 months. Valid values range from 4.1 to 12.2, with higher values indicating more severe food deprivation. The scale score is undefined for households that affirmed no child-referenced items (see discussion of X1FSSCAL above).

7.5.2.5 Food Security Status: Categorical Measures (X1FSSTAT, X1FSADST, and X1FSCHST)

X1FSSTAT is a categorical measure of Household Food Security status formed by dividing X1FSSCAL into three ordered categories: food secure, food insecure without hunger, and food insecure with hunger. X1FSSTAT is appropriate for comparing prevalence rates of food insecurity and hunger across subpopulations, but comparisons may not be valid unless the ages of children are similarly distributed within each of the subpopulations. X1FSSTAT also can be used as a categorical variable in associative models. When interpreting food security statistics, users should remember that food security status is a household-level characteristic. In most households classified as food insecure with hunger, the children in the household were not hungry.

X1FSADST is a categorical measure of Adult Food Security status that identifies households as food secure, food insecure without hunger, and food insecure with hunger among adults. This variable is appropriate for comparing food security status of adults and other adult family members across subpopulations of children.

X1FSCHST is a two-category measure of Children's Food Security status that identifies households with and without hunger among children at some time during the 12 months prior to the survey. This variable is appropriate for comparing prevalence rates of hunger among children across subpopulations. There were few households that reported hunger among children (n = 28, 0.3 percent), so the analytic utility of this variable is limited. Even this small number may nevertheless provide insight into characteristics of these most severely food-insecure households and the young children in them. However, for analytic purposes, other categories of children's food insecurity delineated by less severe thresholds (based on children's food security raw scores or scale scores) may be useful. For example, Nord and Bickel (2001) suggested a threshold of 2 or more affirmative responses to child-referenced items as representing reduced quality and variety of children's diets. When interpreting children's food security statistics, users should remember that these variables represent conditions among all children in the household and may not reflect experiences of the child in the ECLS-B study if there are other children in the household.

7.5.2.6 Food Security Status: Raw Scores (X1FSRAW, X1FSADRA, and X1FSCHRA)

The Household Food Security raw score, X1FSRAW, is a count of affirmative responses to the 18 food security items. Valid values range from 0 to 18. The Adult Food Security raw score, X1FSADRA, is a count of affirmative responses to household- and adult-referenced items. Valid values range from 0 to 10. The Children's Food Security raw score, X1FSCHRA, is a count of affirmative responses to child-referenced items. Valid values range from 0 to 8. Responses to items skipped because of screening are assumed to be "No." All three raw scores indicate the number of food-insecure conditions reported by the household. Thus, higher scores indicate more severe food insecurity. Families with no valid responses (0.08 percent of all households) are coded as missing (−9). Missing item responses of families with at least one valid response (0.16 percent of households had one more such missing responses) are imputed as negative responses.

7.5.2.7 Socioeconomic Status and Poverty (X1FTHSCR, X1MOMSCR, X1SESL, X1SESQ5, X1INCOME, X1POVRTY)

Socioeconomic status (SES) is a measure of social standing (see Duncan [1961] for more information on socioeconomic status)**.** It was computed at the household level using data from the parent CAPI instrument and the resident father questionnaires. The SES variable reflects the socioeconomic status of the household at the time of the 9-month parent CAPI instrument (2002). The components used to create the measure of SES were as follows:

- Father/male guardian’s education;
- Mother/female guardian’s education;
- Father/male guardian’s occupation;
- Mother/female guardian’s occupation; and
- Household income.

Occupations were coded using the *Standard Occupational Classification Manual* (Executive Office of the President, Office of Management and Budget, 2000). The occupation codes were collapsed into 23 aggregated categories, with one additional category for unemployed, retired, disabled, and unclassified workers (see section 6.1.3 for a description of the procedures used to code occupation and section 7.5.3 for a description of the 23 aggregated occupation categories). Occupation was recoded to reflect the average of the 1989 General Social Survey (GSS) prestige score. This was computed as the average of the corresponding prestige scores for the 2000 Census occupational categories covered by the ECLS-B occupation. Table 7-6 provides details on the prestige score values (X1FTHSCR, X1MOMSCR).

The variables were collected as follows:

1. **Parents’ education.** The information about parents’ education was collected in the 9-month parent CAPI instrument and the resident father questionnaire. Household

members identified by the respondent as a mother/female guardian or father/male guardian of the focal child, or a spouse/partner of the mother/female guardian or father/male guardian, were considered parents of the focal child.

2. **Parents' occupation.** The information about parents' occupation was collected in the 9-month parent CAPI instrument and the resident father questionnaire.

3. **Income.** The information about income was collected in the 9-month parent CAPI instrument. Broad-range and detailed-range income questions were asked of all participants. The broad range classifies household income as $25,000 and less per year, or as greater than $25,000. The detailed range classifies household income as shown in table 7-1. Households that were determined to meet the size and income criteria related to poverty shown in table 7-2 were asked to report income to the nearest $1,000. (For simplicity, this is called exact income.) Because not all households were asked to report exact income, the midpoint of the detailed income range was used to compute the SES composite variable.

Table 7-1. Levels of the detailed income range in the 9-month data collection: 2001–02

Detailed income range	Total household income
1	$5,000 or less
2	$5,001 to $10,000
3	$10,001 to $15,000
4	$15,001 to $20,000
5	$20,001 to $25,000
6	$25,001 to $30,000
7	$30,001 to $35,000
8	$35,001 to $40,000
9	$40,001 to $50,000
10	$50,001 to $75,000
11	$75,001 to $100,000
12	$100,001 to $200,000
13	$200,001 or more

SOURCE: U.S. Department of Education, National Center for Education Statistics, Early Childhood Longitudinal Study, Birth Cohort, Nine-Month Data Collection, 2001–02.

Table 7-2. Households asked to report income to the nearest $1,000 in the 9-month data collection: 2001–02

Household size	Total household income
Two or three	$15,000 or less
Four or five	$20,000 or less
Six or seven	$25,000 or less
Eight	$30,000 or less
Nine or more	$35,000 or less

SOURCE: U.S. Department of Education, National Center for Education Statistics, Early Childhood Longitudinal Study, Birth Cohort, Nine-Month Data Collection, 2001–02.

Not all parents who completed the parent CAPI instrument and the resident father questionnaire responded to every question, and not all fathers completed the resident father questionnaire. Therefore, there were missing values for some of the components of the SES composite variable. Only a small percentage of values for the education and occupation variables were missing; a larger proportion of households had missing values for the detailed income range (see table 7-3).

Table 7-3. Item missing data for SES source variables in the 9-month data collection: 2001–02

Variable	Number missing	Percent
Mother’s education	39	0.37
Father’s education	114	1.07
Mother’s occupation	44	0.41
Father’s occupation	190	1.79
Detailed income range	950	8.93

SOURCE: U.S. Department of Education, National Center for Education Statistics, Early Childhood Longitudinal Study, Birth Cohort, Nine-Month Data Collection, 2001–02.

The following procedure was used to impute missing values for each component of the SES composite variable.

Data were imputed using a hot deck methodology. In hot deck imputation, the value reported by a respondent for a particular item is assigned or “donated” to a “similar” person who failed to respond to that question. Auxiliary information known for both donors and nonrespondents is used to form groups of persons having similar characteristics. These groups of similar respondents and nonrespondents are called “imputation cells.” The imputed value for a case with a missing value is taken from a randomly selected donor among the respondents within the cell.

Imputation cells were defined by respondent characteristics that potentially were the best predictors of the variables to be imputed. These relationships had been determined previously by CHAID (Chi-squared Automatic Interaction Detector) analyses of base year data from the Early Childhood Longitudinal Study, Kindergarten Class of 1998–99 (ECLS-K). Missing values for the education, occupation, and detailed income range variables were imputed by the hot deck method for all households with at least one parent present. Hot deck imputation was done in a sequential order, separately, by type of household (female single parent, male single parent, and both parents present). For households with both parents present, the mother’s and father’s variables were imputed separately. Imputed as well as reported values were used to define imputation cells; missing values for donor characteristics were treated as a separate category. No imputed value was used as a donor. No donor was used more than once. The order of hot deck imputation for all the variables was from the lowest percent missing to the highest. For more information on hot-deck imputation procedures, see Little and Rubin (2002).

Hot deck imputation was implemented for three types of households (female single parent, male single parent, and both parents present) defined by the composite variable X1HFAMIL (described below). Households that could not be categorized into one of the three types were labeled as "other" households. An "other" type of household includes those in which there was no biological, adoptive, step-, or foster mother or father, and the child lived with a relative (e.g., aunt, grandmother) or nonrelative. Less than one percent of households were of this "other" type in the ECLS-B 9-month sample. For these households, respondent information related to the components of SES were extracted from the parent CAPI instrument and used to calculate the SES composite variable.

Occupation imputation involved two steps. First, the labor force status of the parent was imputed (i.e., whether the parent was employed). Then the parent's occupation was imputed only for those parents whose status was identified as employed either through the parent CAPI instrument or the first imputation step. The detailed income range was imputed in two steps: first for cases where the broad income range was known, and second for cases where it was unknown.

For households where at least one parent was present, the order of hot deck imputation was as follows:

1. Mother's occupation;
2. Mother's labor force status;
3. Mother's education;
4. Father's occupation;
5. Father's education;
6. Father's labor force status;
7. Detailed income range, where the broad income range was known; and
8. Detailed income range, where the broad income range was unknown.

At this point, all of the missing values had been imputed for the three types of households identified above. However an exact income value was still required to construct the SES composite. The midpoint of the detailed income range was assigned for this purpose to households.

The log of the detailed income range midpoint was then used to compute the SES composite. This value does not vary widely within the levels of the detailed income range, so the midpoint was a reasonable choice. It was used only for the purpose of computing the SES composite and was not retained in the data file.

Table 7-4 shows results for imputing component variables for SES. Each component is listed on the left, with the number missing and number imputed in the right-hand columns. There was relatively little missing data for education and occupation, particularly for mothers. However, there was substantial missing data for income. As shown in table 7-4, all cases with missing data were successfully imputed.

Table 7-4. Summary of imputation results, 9-month data collection: 2001–02

SES component	Number missing	Number imputed
Mother’s education	39	39
Father’s education	114	114
Mother’s labor force status	32	32
Father’s labor force status	301	301
Mother’s occupation	44	44
Father’s occupation	190	190
Detailed income range	950	950

SOURCE: U.S. Department of Education, National Center for Education Statistics, Early Childhood Longitudinal Study, Birth Cohort, Nine-Month Data Collection, 2001–02.

Once the components of the SES variable were imputed, their corresponding *z*-scores or normalized values were computed. The standardization of scores put them on the same numeric scale, allowing them to be averaged. The expression of *z*-score z_{hi} for the *h*-th component in the *i*-th household is

$$z_{hi} = \frac{x_{hi} - \bar{x}_w}{se(\bar{x}_w)},$$

where

x_{hi}	is	the value of the *h*-th SES component for the *i*-th household*;*
$\bar{x}_w$	is	the weighted mean of x_{hi} ; and
$se(\bar{x}_w)$	is	the standard error of $\bar{x}_w$.

Thus, each component was converted to a *z*-score with mean of 0 and a standard deviation of one. For income, the component x_i is the logarithm of the income for *i*-th household. The logarithm of income was used because the distribution of the logarithm of income is less skewed than the direct income values. The SES value for the *i*-th household was then computed as

$$SES_i = \frac{\sum_{h=1}^{m_i} z_{hi}}{m_i},$$

where m_i is number of nonmissing SES components for the *i*-th household. X1SESL is the continuous variable for the SES composite that ranges from –2.10 to 2.25. As described, the SES composite is the average of up to five measures, each of which was standardized to have a mean of 0 and a standard deviation of 1, hence the negative values. For analyses that require a continuous SES measure, such as multiple regressions, X1SESL is the variable to use. A categorical SES variable (X1SESQ5) was created that contains the quintile for the value of the composite SES for the child. Quintile 1 represents the lowest SES category and quintile 5 represents the highest SES category. The quintiles were computed at the child level using the 9-month parent weight (W1R0).

Note that for households with only one parent present, not all the components were defined. In these cases, the SES was computed by averaging the available components.

The imputed detailed income range variable (X1INCOME) was also used to create a household-level poverty variable (X1POVRTY) that defines 100 percent of poverty. Income was compared to Census poverty thresholds for 2001, which vary by household size. Table 7-5 shows the detailed income categories used in the ECLS-B parent CAPI instrument for determining whether to ask a more detailed question about income to the nearest $1,000. For comparison, the table also shows weighted poverty thresholds from the Census.[6] Households in which the income fell below the

[6] The ECLS-B provides an approximate, but not exact measure of poverty. Income category thresholds used in the parent CAPI instrument are similar, but not identical to those from weighted Census averages.

appropriate threshold were classified as poor. For example, if a household contained two members, and the household income was lower than $11,920, then the household was considered to be below 100 percent of the poverty threshold.

Table 7-5. ECLS-B and Census poverty thresholds for 2001

Household size	ECLS-B income categories	Census weighted average thresholds for 2001[1]
2	Less than or equal to $15,000	$11,920
3	Less than or equal to $15,000	$14,128
4	Less than or equal to $20,000	$18,104
5	Less than or equal to $20,000	$21,405
6	Less than or equal to $25,000	$24,195
7	Less than or equal to $25,000	$27,517
8	Less than or equal to $30,000	$30,627
9+	Less than or equal to $35,000	$36,286

[1] U.S. Department of Commerce, U.S. Census Bureau, Current Population Survey, "Poverty 2001" (Available: http://www.census.gov/hhes/poverty/threshld/thresh01.html).

SOURCE: U.S. Department of Education, National Center for Education Statistics, Early Childhood Longitudinal Study, Birth Cohort, Nine-Month Data Collection, 2001–02.

7.5.2.8 Parent Education (X1PARED, X1FTHED, and X1MOMED)

There are three parent education composites in the file. These are X1PARED (the highest level of education for the child's parents who reside in the household), X1FTHED (father's highest level of education), and X1MOMED (mother's highest level of education). The variables include only birth adoptive, step-, or foster parents residing in the household. For example, if the child did not have birth, adoptive, step-, or foster parents but was living with another relative (such as an aunt) who served as the parent respondent, the education of the relative and his or her spouse were not used in the creation of the composites. At 9 months of age, less than one percent of the children did not have at least one parent in the household. Moreover, the information on these other adults is still available in the data file as P1REDU (for respondent) and P1SPEDUC (for respondent's spouse/partner).

If both parents resided in the household, X1PARED was the highest value for education level from either the mother/guardian in X1MOMED or the father/guardian in X1FTHED. For example, if the child's biological mother and biological father both lived in the household, X1PARED was set equal to the highest value for education level found in X1MOMED and X1FTHED. If the household only had one parent, then X1PARED was equal to either X1MOMED or X1FTHED depending on which parent or guardian resided with the child. For example, if the birth mother lived in the household and there was no birth, adoptive, step-, or foster father in the household, X1PARED was set equal to the education level of the birth mother indicated in X1MOMED. If the education data for either of the parents were missing[7] they were imputed, and the composite X1PARED was created based on both the reported and imputed data.

7.5.2.9 Parent Race/Ethnicity (X1HFAMIN, X1HFASN, X1HFPCIL, X1HFBLCK, X1HFWHT, X1HFHISP, X1HFMLRC, X1HFAIBC, X1HFASBC, X1HFPCBC, X1HFBLBC, X1HFWHBC, X1HFHSBC, X1HFRACE, X1HMAMIN, X1HMASN, X1HMPCIL, X1HMBLCK, X1HMWHT, X1HMHISP, X1HMMLRC, X1HMAIBC, X1HMASBC, X1HMPCBC, X1HMBLBC, X1HMWHBC, X1HMHSBC, X1HMRACE)

The composites for the parent's race/ethnicity are presented in the ECLS-B files in three ways: (1) as dichotomous variables for each race/ethnicity category for the father: American Indian (X1HFAMIN), Asian (X1HFASN), Pacific Islander (X1HFPCIL), Black (X1HFBLCK), White (X1HFWHT), Hispanic (X1HFHISP), more than one race unspecified (X1HFMLRC) and for the mother: (American Indian (X1HMAMIN), Asian (X1HMASN), Pacific Islander (X1HMPCIL), Black (X1HMBLCK), White (X1HMWHT), Hispanic (X1HMHISP), more than one race, unspecified (X1HMMLRC) from the parent CAPI instrument and birth certificate data; (2) as dichotomous variables for race/ethnicity categories for the father: American Indian (X1HFAIBC), Asian (X1HFASBC), Pacific Islander (X1HFPCBC), Black (X1HFBLBC), White (X1HFWHBC), Hispanic (X1HFHSBC) and for the mother: American Indian (X1HMAIBC), Asian (X1HMASBC), Pacific Islander (X1HMPCBC), Black (X1HMBLBC), White (X1HMWHBC), Hispanic (X1HMHSBC) indicating whether birth certificate data were used to code father or mother race/ethnicity, and (3) as a single race/ethnicity composite created from the dichotomous variables for each race/ethnicity category for fathers (X1HFRACE) and mothers (X1HMRACE).

The composites for race/ethnicity for the parents were calculated in the same way as those for the child, except that birth certificate data was used to supplement parent-reported race/ethnicity for residential biological mothers and fathers. That is, data on parent race/ethnicity come from the parent

[7] Missing data were due to "Refused" or "Don't know" answers from respondents.

CAPI instrument, and in cases where parent race/ethnicity was missing from the parent CAPI instrument for residential biological mothers and fathers, birth certificate data were used.

Parent respondents were allowed to indicate that they belonged to one or more of 14 race categories. These categories include (1) White, (2) Black or African American, (3) American Indian or Alaska Native, (4) Asian Indian, (5) Chinese, (6) Filipino, (7) Japanese, (8) Korean, (9) Vietnamese, (10) Other Asian, (11) Native Hawaiian, (12) Guamanian or Chamorro, (13) Samoan, and (14) Other Pacific Islander. From these responses, a series of five dichotomous race variables were created that indicated separately whether the respondent belonged to each of five main specified race groups, including White, Black, Asian (including Asian Indian, Chinese, Filipino, Japanese, Korean, Vietnamese, and Other Asian), and Native Hawaiian or other Pacific Islander (including Native Hawaiian, Guamanian or Chamorro, Samoan, and Other Pacific Islander). In addition, one more dichotomous variable was created for parent respondents who had simply indicated that they were multiracial without specifying a race (e.g., biracial). For residential biological mothers and fathers whose race was refused, don't know, missing or unspecified in the parent CAPI instrument, birth certificate data, if available, were used to code the five dichotomous race variables. The variables X1HFAIBC, X1HFASBC, X1HFPCBC, X1HFBLBC, and X1HFWHBC (for fathers) and X1HMAIBC, X1HMASBC, X1HMPCBC, X1HMBLBC, and X1HMWHBC (for mothers) indicate whether birth certificate data were used to code the five dichotomous race variables (1 = Yes, 2 = No).

Parent race/ethnicity was obtained for all parents and spouses of respondent parents, but may or may not have been collected for a parent's boyfriend or girlfriend. For example, in a family with a birth mother and stepfather the race/ethnicity of both parents was obtained. However, in a family with a birth mother and the mother's boyfriend, if the mother's boyfriend was not identified as a spouse or partner of the mother, the race/ethnicity of the mother was obtained but that of the boyfriend was not. In addition, the race composites are calculated only for birth, adoptive, step-, and foster parents, and not for nonparent relatives and nonrelatives.

7.5.2.10 Parent Identifiers and Household Composition (X1MOMTYP, X1FTHTYP, X1MOMID, X1FTHID, X1HPARNT, X1HFAMIL)

The construction of parent identifiers and the household composition variables from the parent CAPI instrument data was done as follows. Individuals identifying themselves as the child's mother/father were located within the household roster, and their relationship to the child (biological, adoptive, foster, step-, partner of respondent, or unknown) was established. For households containing more than one father or mother, a hierarchy was used to designate the "current" or residential mother and/or father. The biological parent, if present, was always the current mother or father. In the absence of a biological parent, the current mother/father designation was assigned to the adoptive, step-, foster/guardian, partner (including household members defined as spouses/partners of the respondent, but not identified by the respondent as mothers/female guardians or fathers/male guardians), or "unknown-type" parent.

The type of mother/father in the household established by the procedure outlined above is identified in the composite variables X1MOMTYP (type of resident mother; biological, adoptive, foster, step-, partner of father, partner of relative/nonrelative father figure, or unknown) and X1FTHTYP[8] (type of resident father; biological, adoptive, foster, step-, partner of mother, partner of relative; nonrelative mother figure, or unknown). If there were no household members that could be identified as one of the parent types outlined above, the composite variables were set to equal 7 (no resident father/mother). The person numbers of the household members identified in X1MOMTYP and X1FTHTYP are provided in the composite variables X1MOMID and X1FTHID respectively. Person number refers to the number each household member has on the household roster list. Household members are listed in the order they are reported by the respondent (see section 6.1.3 for more information about the household roster).

After the resident parents were identified and the composite variables X1MOMTYP and X1FTHTYP were constructed, the parent demographic variables (including age, race/ethnicity, education, occupation, and employment status) were then constructed for all parents identified in X1MOMTYP and X1FTHTYP.[9]

---

[8] Users should be aware that there are 10 cases where there is a biological father in the household, but someone other than this biological father completed the resident father questionnaire. This is, the mother indicated that the biological father lives in the household (X1FTHTYP = 1), but the first question on the resident father questionnaire (F1RELCH – "What is your relationship to the child?") reveals that someone other than the biological father completed the resident father questionnaire. These cases have a father weight (W1F0) assigned to them. Users can decide whether or not to include these cases in their analyses.

[9] Prior to finalizing the data, 53 cases were identified that consisted of relative/nonrelative female respondents to the parent CAPI interview. Originally, these cases had a value of X1MOMTYP = no resident mother. The composite X1MOMTYP was revised, and a new category (Rel/nonrel resp–mother figure) was added to include these cases. Of these 53 cases, 24 were identified as having partners who completed the resident father questionnaires. Originally these cases had a value of X1FTHTYP = no resident father. The composite X1FTHTYP was revised and a new category (Rel/nonrel resps partner–father figure) was added to include these cases. In addition, 2 cases were identified that consisted of relative/nonrelative male respondents. Originally, these cases had a value of X1FTHTYP = no resident father. The composite X1FTHTYP was revised and a new category (Rel/nonrel resp–father figure.) was added to include these cases. The parent demographic composite variables (age, race/ethnicity,education, occupation, and employment status), however, were not revised for these cases. To obtain the syntax necessary for updating these demographic composites to reflect the changes to X1FTHTYP and X1MOMTYP, go to http://nces.ed.gov/ecls.

Information about parents in the household, along with household size and presence or absence of grandparents, siblings, and other relatives, was used to construct the household composition variables X1HPARNT (classification of focal child's parents who reside in the household) and X1HFAMIL (family type categories using both parent and sibling information).

For example, for X1HFAMIL, composite values are as follows:

- 1 = two parents and sibling(s)
- 2 = two parents, no siblings
- 3 = one parent and sibling(s)
- 4 = one parent, no siblings
- 5 = other

Parent figures not identified by X1MOMTYP and X1FTHTYP were placed in the "other" category for this composite. Likewise, for the composite X1HPARNT, parent figures were placed in categories 8 or 9 for related and unrelated guardians, respectively.[10]

7.5.3 Occupation Codes Used in the ECLS-B

This section describes the aggregated categories that were used for coding occupation in the ECLS-B. Occupations were coded using the *Standard Occupational Classification Manual* (Executive Office of the President, Office of Management and Budget, 2000). The occupation codes were collapsed into the following 23 codes, plus one additional code for unemployed/retired/disabled/unclassified workers (see section 6.1.3 for a description of the procedures used to code occupation).

1. Management Occupations

This category includes senior-level and middle management occupations and occupations that directly support management. Senior-level managers are persons concerned with policymaking, planning, staffing, directing, and/or controlling

[10] It should be noted that because the composite construction identifies only one resident mother or one resident father, same-sex parents are not readily identified in the composites themselves. Two approaches can be used to identify these couples. First, the user should search the relationship variables (P1REL_01, etc.) to identify households in which more than one person is identified as a father/mother to the focal child. Second, since not all same-sex partners identify themselves as "mother" or "father" to the focal child, the user should also search for households in which the respondent is the child's parent and the respondent's spouse/partner (identified from P1SPSPTR) is the same sex as the respondent.

activities. Middle managers include persons who plan, organize, or direct and/or control activities at the operational level. Workers in this category are not directly concerned with the fabrication of products or with the provision of services. Other officials and administrators include legislators, education administrators, construction managers, operations specialty managers, and agricultural managers.

2. Business and Financial Operations Occupations

This category includes senior- and middle-level business and financial occupations. Senior-level business operations specialists are persons concerned with wholesale and retail buying, negotiators of contracts and labor relations, management analysts, and emergency management specialists. The middle-level business occupation category includes claims adjusters, compliance officers, convention planners, and logisticians. Financial specialists include accountants, auditors, analysts, and examiners.

3. Computer and Mathematical Science Occupations

This category includes occupations concerned with the fields of computer and mathematical science. The computer science field includes computer and systems software programmers, computer support specialists, and database and network administrators. The mathematical sciences positions in this category include actuaries, operations research analysts, statisticians, and mathematical scientists.

4. Architecture and Engineering Occupations

This category includes occupations concerned with applying principles of architecture and engineering in the design and construction of buildings, equipment and processing systems, highways and roads, and land utilization. Aerospace, agricultural, biomedical, chemical, civil, computer hardware, electrical, environmental, industrial, mechanical, and marine engineers are all included in this category.

5. Life, Physical, and Social Science Occupations

This category includes those engaged primarily in the application of scientific principles to research and development. Life sciences are those including biology, agriculture and medicine. Physical science includes chemistry and physics, while the social sciences focus on economics, psychology, sociology, history, and geography. In addition, this category includes the support technicians for all of these fields of research and study.

6. Community and Social Services Occupations

This category includes occupations concerned with the social needs of people and basic and applied research in the social sciences. For example, counselors, social workers, community and social service specialists, and religious workers are included in this category.

7. Legal Occupations

This category includes occupations directly related to the legal system, including judges, lawyers, paralegals and legal assistants, court reporters, law clerks, and title examiners.

8. Education, Training, and Library Occupations

This category includes those who teach at higher education institutions and at other postsecondary (after high school) institutions, such as vocational institutions. This category also includes prekindergarten and kindergarten teachers, elementary and secondary teachers, special education teachers, instructional coordinators, and adult education teachers (outside postsecondary). In addition, vocational and educational counselors, librarians, curators, and archivists are included here.

9. Arts, Design, Entertainment, Sports, and Media Occupations

This category includes occupations concerned with creating and executing artistic works in a personally interpreted manner by painting, sculpturing, drawing, engraving, etching and other methods; creating designs for products and interior decorations; designing and illustrating books, magazines, and other publications; writing; still, motion picture and television photography/filming; producing, directing, staging, acting, dancing, and singing in entertainment; and participating in sports and athletics as a competitor or player and administering and directing athletic programs.

10. Health Care Practitioners and Technical Occupations

This category includes health care professionals who diagnose and treat patients. In addition to physicians, dentists, and veterinarians, this category includes optometrists, podiatrists, and other diagnosing and treating professionals, such as chiropractors, hypnotherapists, and acupuncturists. This category also includes occupations concerned with the maintenance of health, the prevention of illness and the care of the ill through the provision and supervision of nursing care; compounding drugs; planning food service or nutritional programs; providing assistance to physicians; and the provision of therapy and treatment as directed by physicians.

Health technologists and technicians, for example, clinical laboratory technologists and technicians, dental hygienists, radiologic technicians, licensed practical nurses (LPNs), and other health technologists are also included in this category.

11. Health Care Support Occupations

This category includes occupations concerned with providing aid and assistance to medical and therapy staff. For example, home health aides, physical therapy assistants or aides, massage therapists, medical assistants, and veterinary assistants are included here.

12. Protective Service Occupations

This category includes occupations providing protective services. Examples are fire fighters, fire inspectors, police officers, correctional officers, animal control workers, and security guards.

13. Food Preparation and Serving Related Occupations

This category includes occupations concerned with the preparation, cooking, and service of food. Some examples are chefs and head cooks, fast food and counter workers, bartenders, waiters, and dishwashers.

14. Building and Grounds Cleaning and Maintenance Occupations

This category includes occupations concerned with the maintenance and upkeep of buildings and grounds. It includes janitors, maids, landscape gardeners, pesticide handlers, and tree and shrub trimmers and pruners.

15. Personal Care and Service Occupations

This category includes occupations providing personal services to individuals. Some examples are nonfarm animal caretakers, entertainment attendants, personal appearance workers, transportation, tourism and lodging attendants, and child care workers.

16. Sales and Related Occupations

This category includes occupations involving in selling goods or services, purchasing commodities and property for resale, and conducting wholesale or retail business.

17. Office and Administrative Support Occupations

This category includes occupations involving preparing, transcribing, transferring, systematizing, and preserving written communications and records; collecting accounts; gathering and distributing information; operating office machines and data processing equipment; operating switchboards; distributing mail and messages; and other support and clerical duties such as bank telling, data entry keying, and so forth.

18. Farming, Fishing, and Forestry Occupations

This category is concerned with the production, propagation (breeding/growing), gathering and catching of animals, animal products, and plant products (timber, crop, and ornamental); the provision of services associated with agricultural production; and game farms, fisheries and wildlife conservation. “Other agricultural and related occupations” include occupations concerned with the production and propagation of animals, animal products, plants and products (crops and ornamental).

19. Construction and Extraction Occupations

This category includes occupations that normally are performed at a specific site, which will change over time, in contrast to production workers, where the work is usually at a fixed location. Construction workers include those in overall construction, brickmasons, stonemasons, carpenters, electricians, drywall installers, paperhangers and painters, and so forth. The extractive occupation category includes oil well drillers, mining machine operators, and so on.

20. Installation, Maintenance, and Repair Occupations

This category includes occupations that are concerned with installation, adjustment, maintenance, part replacement, and repair of tools, equipment, and machines.

21. Production Occupations

This category includes both precision and manufacturing production. Precision production includes occupations concerned with performing production tasks that require a high degree of precision or attainment of rigid specification and operating plants or large systems. Examples are tool and die makers, pattern and model makers, machinists, jewelers, engravers, and so on. Also included are some food-related workers including butchers and bakers. Plant and system operators include water and sewage, gas power, chemical, petroleum, and other plant or system operators. Manufacturing production includes occupations concerned with setting up, operating, and tending of machines and hand production work usually in a factory or other fixed place of business.

22. Transportation and Material Moving Occupations

This category includes occupations concerned with operating and controlling equipment used to facilitate the movement of people or materials and the supervising of those workers.

23. Military-Specific Occupations

This category includes occupations that are unique to the military. Artillery and missile officers, special forces officers, and infantry soldiers are included in this category. Many military occupations are similar to occupations in the other categories, and they are classified in those groups. For example, a cook in the military would be classified under the Food Preparation and Serving Related Occupation category.

90. Unemployed, Retired, Disabled, or Unclassified Workers

This category includes persons who are unemployed, have retired from the work force, or are disabled. It also includes unclassified occupations that do not fit in the categories above.

7.5.4 Derived and Indicator Composite Variables

To facilitate methodological research, 18 derived variables and 8 indicator variables are included on the 9-month data file. All derived and indicator variables are listed and described in table 7-6.

The derived variables include a series of dichotomous variables that indicate whether the child was sampled as part of certain populations that were oversampled in the ECLS-B. These include X1TWSAMP (Child sampled as part of twin), X1AISAMP (Child sampled as part of American Indian population), X1ASISMP (Child sampled as part of the Asian and Pacific Island population), and X1CHNSMP (Child sampled as part of Chinese population). In addition, derived variables for the child's birth weight status (X1BTHWGT; 1 = Normal birth weight, 2 = Moderately low birth weight, and 3 = Very low birth weight) and the child's multiple birth status (X1MBRTST; 1 = Singleton, 2 = Twin, 3 = Higher order) are included in the data file.

The derived variable X1HHURBN classifies the location where the parent interview took place as urban or rural using definitional criteria from the 2000 census. The values for this composite variable are: 1 = Urban, inside UA, 2 = Urban, inside UC, 3 = Rural, where "UA" refers to an urbanized area and "UC" refers to an urban cluster (i.e., less densely populated than a "UA" or urbanized area).

Several derived variables indicate the language in which components of the 9-month data collection were conducted and whether an interpreter was used. These include the language in which the parent CAPI instrument was conducted (X1LNGPRT), the language used by parent respondent in the NCATS teaching task (X1LNGNCT), and whether an interpreter was used to conduct the parent CAPI instrument (X1INTRST). In addition, the interviewer identification number (X1INTVID) is included in the data file.

Eight indicator variables were added to the data file to identify the status of various components of the 9-month data collection. Several of these indicator variables simply identify the presence or absence (1 = Yes, present, 2 = No, not present) of components, including the parent CAPI instrument (X1STATPI), the parent self-administered questionnaire (X1STTPSQ), and the resident father questionnaire (X1STTRFQ).

Other indicator variables have more detailed values and combine information from other composite variables in the data file. An indicator variable for the status of the BSF-R mental assessment

(X1STBMTL) includes five categories: 1 = BSF-R mental administered, no modifications, 2 = BSF-R mental administered, with modifications, 3 = BSF-R mental administered, unknown modification status, 4 = Not scoreable,[11] 5 = Not administered. Similar categories are used for an indicator variable for the status of BSF-R motor assessment (X1STBMTR). For the NCATS tape of the parent-child interaction, four categories are identified in the indicator variable X1STTNCT: 0 = No NCATS tape, 1 = NCATS tape coded, 2 = NCATS tape not coded, foreign language, 3 = NCATS tape not coded, uncodeable.[12]

Finally, an indicator variable for the status of the physical measurements (X1STATPM) has three values that are based on the presence or absence of the child's length and weight measurements. These values are (1) Both length and weight measurements are present, (2) Only one measurement (length or weight) is present, and (3) Neither length nor weight measurements are present. Because the child's length and weight measurements were considered to be most useful to users of the physical measurements taken during the child assessment, the child's middle upper arm circumference and head circumference measurements were not included in the definition of the status of the physical measurements indicator variable.

7.6 Composite Table

Table 7-6 describes the composite and derived variables that are on the ECLS-B 9-month child catalog. Please note that a few of the variables specified in the "derived from" column are not included in the final data sets. For example, several composite variables use information from the Field Management System (FMS), and these FMS variables are not included in the data file.

The "derived from" column also contains the item numbers from the questionnaire, which help in identifying the items that were used in the creation of these composites. This information allows a user to decide if he or she would like to use the composite based on how it was defined. Some variables in table 7-6 have been recoded or suppressed. Reasons for these data changes are discussed in section 7.7. The new recoded categories are noted for applicable variables in table 7-6.

[11] For more information on cases determined to be "Not scoreable," refer to section 3.1.1 in chapter 3 of this manual.
[12] The determination of whether a tape could be coded was made at Westat. If a coder judged a tape uncodeable, the coder completed a checklist to indicate the reasons why (see variables C1LNGUAG, C1LESSMN C1TRIAD, C1PRQUAL, C1EQPROB, C1NONTOY, and C1NOCDOT). Because the coder could check more than one reason, the sum of the reasons why a tape is not codeable exceeds the number of tapes that are not codeable as indicated in X1STTNCT and in C1ABLCDE.

Table 7-6. ECLS-B 9-month composite variables

Variable name	Category	Description	Derived from	Output values
X1ASAGE	Child	Child's age at the time the direct child assessment occurred. This age is in decimal months.	P1CHDOBC (IN135), P1CHDOBM (IN140), P1CHDOBD (IN142), P1CHDOBY (IN144), Preloaded variables for child's date of birth from birth certificate (P1CBIRTH, P1CHMBIR, P1CHYBIR, not on data file), date of the child assessment (C1V1DATE, C1V2DATE, C1CADATE), date of the parent interview (FMS), date of the child assessment (FMS)	Continuous
X1BSIDAG	Child	Focal child's age adjusted for prematurity (BSID age). This age is in decimal months.	Variable upcoded from parent CAPI application (not on data file)	Continuous
X1CHSEX	Child	Sex of focal child	P1CHLSEX (Birth certificate data), P1CHSEX (IN145)	1 = Male, 2 = Female
X1CHRACE	Child	Race and ethnicity of the focal child	RACE1, RACE2, RACE3, RACE4, RACE5, RACE6 (Variables coded in parent interview based on FS085), and P1HSP_01 through P1HSP_19 (FS075)	1 = White, non-Hispanic, 2 = Black or African American, non-Hispanic, 3 = Hispanic, race specified, 4 = Hispanic, no race specified, 5 = Asian, non-Hispanic, 6 = Native Hawaiian or other Pacific Islander, non-Hispanic, 7 = American Indian or Alaska Native, non-Hispanic, 8= More than 1 race, non-Hispanic

See note at end of table.

Table 7-6. ECLS-B 9-month composite variables—Continued

Variable name	Category	Description	Derived from	Output values
X1CHAMIN	Child	Child is Native American	RACE1 (Variable coded in parent interview based on P1AIS_01 through P1AIS_19, and P1AIU_01 through P1AIU_19 (FS085)).	1 = Yes, 2 = No
X1CHASN	Child	Child is Asian	RACE2 (Variable coded in parent interview based on P1ASN_01 through P1ASN_19, P1CHN_01 through P1CHN_19, P1FIL_01 through P1FIL_19, P1JPN_01 through P1JPN_19, P1KOR_01 through P1KOR_19, P1VTN_01 through P1VTN_19, P1AAS_01 through P1AAS_19, and P1ASU_01 through P1ASU_19 (FS085)).	1 = Yes, 2 = No
X1CHPCIL	Child	Child is Pacific Islander	RACE4 (Variable coded in parent interview based on P1NHI_01 through P1NHI_19, P1GC_01 through P1GC_19, P1SAM_01 through P1SAM_19, P1PIS_01 through P1PIS_19, and P1PIU_01 through P1PIU_19 (FS085)).	1 = Yes, 2 = No
X1CHBLCK	Child	Child is African American	RACE3 (Variable coded in parent interview based on P1BLK_01 through P1BLK_19 (FS085)).	1 = Yes, 2 = No
X1CHWHT	Child	Child is White	RACE5 (Variable coded in parent interview based on P1WH_01 through P1WH_19 (FS085)).	1 = Yes, 2 = No

See note at end of table.

Table 7-6. ECLS-B 9-month composite variables—Continued

Variable name	Category	Description	Derived from	Output values
X1CHHISP	Child	Child is Hispanic	P1HSP_01 through P1HSP_19 (FS075)	1 = Yes, 2 = No
X1CHMLRC	Child	Child is more than one race, unspecified	RACE6 (Variable coded from parent interview based on P1ARS_01 through P1ARS_19 (FS085), and P1ARU_01 through P1ARU_19 (FS089)).	1 = Yes, 2 = No
X1PRIMNW	Indirect Child	Primary, nonparental child care arrangement in which the child currently spends the most hours per week	P1RELNOW (CC015), P1RHRS (CC070), P1NRNOW (CC120), P1NHRS (CC170), P1CTRNOW (CC220), P1CHRS (CC260)	0 = No nonparental care 1 = Relative care in child's home 2 = Relative care in another home 3 = Relative care, location varies 4 = Nonrelative care in child's home 5 = Nonrelative care in another home 6 = Nonrelative care, location varies 7 = Center-based program 8 = Multiple care arrangements with same hours in each
X1HRSCAR	Indirect Child	Total number of hours per week the focal child currently spends in all primary and secondary, nonparental child care arrangements	P1RELNOW (CC015), P1RELNUM (CC035), P1RHRS (CC070), P1RHROTH (CC118), P1NRNOW (CC120), P1NRNUM (CC140), P1NHRS (CC170), P1NHROTH (CC218), P1CTRNOW (CC220), P1CTRNUM (CC240), P1CHRS (CC260), P1CHROTH (CC308)	Continuous
X1AGCARE	Indirect Child	The age in months when the focal child first began any type of nonparental care on a regular basis	P1RELNOW (CC015), P1REVER (CC020), P1RAGEMM (CC025), P1NRNOW (CC120), P1NREVER (CC125), P1NAGEMM (CC130), P1CTRNOW (CC220), P1CEVER (CC225), P1CAGEMM (CC230)	Continuous

See note at end of table.

Table 7-6. ECLS-B 9-month composite variables—Continued

Variable name	Category	Description	Derived from	Output values
X1CHPREM	Indirect Child	Prematurity of child (in days)	BCGESTWK (gestation detail in weeks from the birth certificate)	Continuous
X1NCATTM	Direct Child	Time it took parent to teach the NCATS task to the focal child. This time is in seconds.	C1LENGTH (NCATS CADE)	Continuous
X1CHLENG	Direct Child	Child's 9-month composite length in centimeters	C1CHLGT1, C1CHLGT2	Continuous
X1CHWGHT	Direct Child	Child's 9-month composite weight in kilograms	C1CHWGT1, C1CHWGT2	Continuous
X1CHMUAC	Direct Child	Child's 9-month composite MUAC (Middle Upper Arm Circumference) in centimeters	C1MUAC1, C1MUAC2	Continuous
X1CHCRFM	Direct Child	Child's 9-month composite head circumference in centimeters (for very low birth weight children only)	C1CHHC1, C1CHHC2	Continuous

See note at end of table.

Table 7-6. ECLS-B 9-month composite variables—Continued

Variable name	Category	Description	Derived from	Output values
X1HPARNT	Family/HH	Classification of focal child's parents who reside in the household	P1RSREL (IN035), X1MOMTYP, X1FTHTYP (composites)	1 = Biological mother and biological father, 2 = Biological mother and other father (step-, adoptive, foster), 3 = Biological father and other mother (step-, adoptive, foster), 4 = Biological mother only, 5 = Biological father only, 6 = Two adoptive parents, 7 = Single adoptive parent or adoptive parents and stepparent, 8 = Related guardian(s), 9 = Unrelated guardian(s)
X1NUMSIB	Family/HH	Total number of siblings with whom the focal child lives, including anyone reporting him/herself as the child of the focal child's foster parent/guardian	P1RSREL (IN035), P1REL_01 through P1REL_19 (FS040)	Continuous
X1HFAMIL	Family/HH	Family type categories using both parent and sibling information	P1RSREL (IN035), P1NRREL (IN060), P1REL_01 through P1REL_19 (FS040), P1NRS_01 through P1NRS_19 (FS065), X1MOMTYP, X1FTHTYP, X1NUMSIB, X1HPARNT (composites)	1 = Two parents and sibling(s), 2 = Two parents, no siblings, 3 = One parent and sibling(s), 4 = One parent, no siblings, 5 = Other
X1MARSTA	Family/HH	Marital status of parent(s) in household	P1MOMREL (IN040), P1DADREL (IN045), P1MARSTS (MH005), X1HPARNT (Composite)	1 = Married, 2 = Separated, 3 = Divorced, 4 = Widowed, 5 = Never married, 6 = No biological or adoptive parents in household

See note at end of table.

Table 7-6. ECLS-B 9-month composite variables—Continued

Variable name	Category	Description	Derived from	Output values
X1LESS18	Family/HH	Total number of household members younger than 18 years old	P1AGE_01 through P1AGE_19 (FS015)	Continuous
X118OVER	Family/HH	Total number of household members age 18 years and older	P1AGE_01 through P1AGE_19 (FS015)	Continuous
X1HTOTAL	Family/HH	Total number of household members	Household roster	Continuous
X1POVRTY	Family/HH	Poverty indicator	P1HHINCS (HI015), P1HHINCY (HI020), X1HTOTAL (composite), 2001 Census-defined thresholds	1 = Below poverty threshold, 2 = At or above poverty threshold
X1INCOME	Family/HH	Household income	P1HHINCS (HI015)	Categorical
X1SESL	Family/HH	Socioeconomic scale	X1INCOME, X1FTHED, X1MOMED, X1FTHSCR, X1MOMSCR (all composites)	Continuous
X1SESQ5	Family/HH	Quintile indicator for X1SESL	X1SESL (composite)	1 = First quintile (lowest), 2 = Second quintile, 3 = Third quintile, 4 = Fourth quintile, 5 = Fifth quintile (highest)

See note at end of table.

Table 7-6. ECLS-B 9-month composite variables—Continued

Variable name	Category	Description	Derived from	Output values
X1FSRAW	Family/HH	Household food security raw score, a simple count of the number of food security items affirmed by the parent	P1WRRYFD (HF020a), P1FDLST (HF020b), P1BLCDML (HF020c), P1FEWTYP (HF020d), P1CHBLML (HF020e), P1CHFOOD (HF022), P1SKML (HF025), P1HOSKML (HF030), P1EATLSS (HF035), P1NOTEAT (HF040), P1LOSTWT (HF045), P1SKFDDY (HF050), P1HOSKDY (HF055), P1CHMLSZ (HF060), P1CHSKML (HF065), P1HOCHSK (HF070), P1CHHNGR (HF075), P1CHSKDY (HF080)	Continuous
X1FSSCAL	Family/HH	Household food security scale score. This is a measure of the severity of food insecurity or hunger experienced in the household in the previous 12 months	P1WRRYFD (HF020a), P1FDLST (HF020b), P1BLCDML (HF020c), P1FEWTYP (HF020d), P1CHBLML (HF020e), P1CHFOOD (HF022), P1SKML (HF025), P1HOSKML (HF030), P1EATLSS (HF035), P1NOTEAT (HF040), P1LOSTWT (HF045), P1SKFDDY (HF050), P1HOSKDY (HF055), P1CHMLSZ (HF060), P1CHSKML (HF065), P1HOCHSK (HF070), P1CHHNGR (HF075), P1CHSKDY (HF080)	Continuous

See note at end of table.

Table 7-6. ECLS-B 9-month composite variables—Continued

Variable name	Category	Description	Derived from	Output values
X1FSSTAT	Family/HH	A categorical measure of household food security status that identifies households as food secure, food insecure without hunger, food insecure with hunger (moderate), and food insecure with hunger (severe)	P1WRRYFD (HF020a), P1FDLST (HF020b), P1BLCDML (HF020c), P1FEWTYP (HF020d), P1CHBLML (HF020e), P1CHFOOD (HF022), P1SKML (HF025), P1HOSKML (HF030), P1EATLSS (HF035), P1NOTEAT (HF040), P1LOSTWT (HF045), P1SKFDDY (HF050), P1HOSKDY (HF055), P1CHMLSZ (HF060), P1CHSKML (HF065), P1HOCHSK (HF070), P1CHHNGR (HF075), P1CHSKDY (HF080)	1 = Food secure, 2 = Food insecure without hunger, 3 = Food insecure with hunger
X1FSADRA	Family/HH	Adult food security raw score, a simply count of the number of household- and adult-referenced food security items affirmed by the parent	P1WRRYFD (HF020a), P1FDLST (HF020b), P1BLCDML (HF020c), P1SKML (HF025), P1HOSKML (HF030), P1EATLSS (HF035), P1NOTEAT (HF040), P1LOSTWT (HF045), P1SKFDDY (HF050), P1HOSKDY (HF055)	Continuous
X1FSADSC	Family/HH	Adult food security scale score. This is a measure of the severity of food insecurity of hunger experienced by adults in the household in the previous 12 months	P1WRRYFD (HF020a), P1FDLST (HF020b), P1BLCDML (HF020c), P1SKML (HF025), P1HOSKML (HF030), P1EATLSS (HF035), P1NOTEAT (HF040), P1LOSTWT (HF045), P1SKFDDY (HF050), P1HOSKDY (HF055)	Continuous

See note at end of table.

Table 7-6. ECLS-B 9-month composite variables—Continued

Variable name	Category	Description	Derived from	Output values
X1FSADST	Family/HH	A categorical measure of adults' food security status that identifies households as food secure, food insecure without hunger, and food insecure with hunger among adults	P1WRRYFD (HF020a), P1FDLST (HF020b), P1BLCDML (HF020c), P1SKML (HF025), P1HOSKML (HF030), P1EATLSS (HF035), P1NOTEAT (HF040), P1LOSTWT (HF045), P1SKFDDY (HF050), P1HOSKDY (HF055)	1 = Food secure, 2 = Food insecure without hunger among adults, 3 = Food insecure with hunger among adults
X1FSCHRA	Family/HH	Children's food security raw score, a simple count of the number of child-referenced food security items affirmed by the parent	P1FEWTYP (HF020d), P1CHBLML (HF020e), P1CHFOOD (HF022), P1CHMLSZ (HF060), P1CHSKML (HF065), P1HOCHSK (HF070), P1CHHNGR (HF075), P1CHSKDY (HF080)	Continuous
X1FSCHSC	Family/HH	Children's food security scale score. This is a measure of the severity of food insecurity or hunger experienced by children in the household in the previous 12 months	P1FEWTYP (HF020d), P1CHBLML (HF020e), P1CHFOOD (HF022), P1CHMLSZ (HF060), P1CHSKML (HF065), P1HOCHSK (HF070), P1CHHNGR (HF075), P1CHSKDY (HF080)	Continuous
X1FSCHST	Family/HH	A categorical measure of children's food security status that identifies households with hunger among children at some time during the 12 months prior to the survey	P1FEWTYP (HF020d), P1CHBLML (HF020e), P1CHFOOD (HF022), P1CHMLSZ (HF060), P1CHSKML (HF065), P1HOCHSK (HF070), P1CHHNGR (HF075), P1CHSKDY (HF080)	1 = Food secure or food insecure without hunger among children, 2 = Food insecure with hunger among children
X1LANGST	Family/HH	The primary language other than English that is spoken in child's home.	P1ANYLNG (HE005), P1ENGLIS (HE015), P1PRMLNG (HE020), P1PRMLOS (HE022; not on data file)	1 = Non-English language, 2 = English language

See note at end of table.

Table 7-6. ECLS-B 9-month composite variables—Continued

Variable name	Category	Description	Derived from	Output values
X1FTHID	Family/HH	Household roster number of resident father	P1RSREL (IN035), P1NRREL (IN060), P1REL_01 through P1REL_19 (FS040), P1NRS_01 through P1NRS_19 (FS065)	Continuous
X1FTHTYP	Family/HH	Type of resident father; indicates whether the birth, adoptive, step- or foster father of the focal child resides in the household with the focal child	P1RSREL (IN035), P1DADREL (IN045), P1PARTNR (FS030), P1REL_01 through P1REL_19 (FS040), P1DAD_01 through P1DAD_19 (FS050), F1RELCH (Resident father (and Other Important People) SAQ)	1 = Birth father, 2 = Adoptive father, 3 = Step father, 4 = Foster father, 5 = Rel/nonrel resp–father figure, 6 = Mothers partner–father figure, 7 = Rel/nonrel resps partner–father figure, 8 = Father, unknown type, 9 = No resident father
X1HFAGE	Family/HH	Age of resident father	P1RSDOBM (IN031), P1RSDOBY (IN033), P1RSAGEY (IN033b), P1RSREL (IN035), P1AGE_01 through P1AGE_19 (FS015), P1PARTNR (FS030), P1SPDOBM (SI031), P1SPDOBY(SI033), F1BDTEMM, F1BDTEDD, F1BDTEYY (Resident father (and Other Important People) SAQ-Q33), date of parent interview	Continuous
X1HFRACE	Family/HH	Race and ethnicity of the father or male guardian in the household	RACE1, RACE2, RACE3, RACE4, RACE5, RACE6 (Variables coded in parent interview based on FS085), and P1HSP_01 through P1HSP_19 (FS075)).	1 = White, non-Hispanic, 2 = Black or African American, non-Hispanic, 3 = Hispanic, race specified, 4 = Hispanic, no race specified, 5 = Asian, non-Hispanic, 6 = Native Hawaiian or other Pacific Islander, non-Hispanic, 7 = American Indian or Alaska Native, non-Hispanic, 8 = More than 1 race, non-Hispanic

See note at end of table.

Table 7-6. ECLS-B 9-month composite variables—Continued

Variable name	Category	Description	Derived from	Output values
X1HFAMIN	Family/HH	Father or male guardian in the household is American Indian	RACE1 (Variable coded in parent interview based on P1AIS_01 through P1AIS_19, and P1AIU_01 through P1AIU_19 (FS085)), BCFTHRC (father race from birth certificate)	1 = Yes, 2 = No
X1HFASN	Family/HH	Father or male guardian in the household is Asian	RACE2 (Variable coded in parent interview based on P1ASN_01 through P1ASN_19, P1CHN_01 through P1CHN_19, P1FIL_01 through P1FIL_19, P1JPN_01 through P1JPN_19, P1KOR_01 through P1KOR_19, P1VTN_01 through P1VTN_19, P1AAS_01 through P1AAS_19, and P1ASU_01 through P1ASU_19 (FS085)), BCFTHRC (father race from birth certificate)	1 = Yes, 2 = No
X1HFPCIL	Family/HH	Father or male guardian in the household is Pacific Islander	RACE4 (Variable coded in parent interview based on P1NHI_01 through P1NHI_19, P1GC_01 through P1GC_19, P1SAM_01 through P1SAM_19, P1PIS_01 through P1PIS_19, and P1PIU_01 through P1PIU_19 (FS085)), BCFTHRC (father race from birth certificate)	1 = Yes, 2 = No
X1HFBLCK	Family/HH	Father or male guardian in the household is Black	RACE3 (Variable coded in parent interview based on P1BLK_01 through P1BLK_19 (FS085)), BCFTHRC (father race from birth certificate)	1 = Yes, 2 = No

See note at end of table.

Table 7-6. ECLS-B 9-month composite variables—Continued

Variable name	Category	Description	Derived from	Output values
X1HFWHT	Family/HH	Father or male guardian in the household is White	RACE5 (Variable coded in parent interview based on P1WH_01 through P1WH_19 (FS085)), BCFTHRC (father race from birth certificate)	1 = Yes, 2 = No
X1HFHISP	Family/HH	Father or male guardian in the household is Hispanic	P1HSP_01 through P1HSP_19 (FS075), BCFTHHSP (father ethnicity from birth certificate)	1 = Yes, 2 = No
X1HFMLRC	Family/HH	Father or male guardian in the household is more than one race, unspecified	RACE6 (Variable coded from parent interview based on P1ARS_01 through P1ARS_19 (FS085), and P1ARU_01 through P1ARU_19 (FS089)), BCFTHRC (father race from birth certificate)	1 = Yes, 2 = No
X1HFAIBC	Family/HH	Flag indicating that the biological father's race from the birth certificate was used to code X1HFAMIN	RACE1 (Variable coded in parent interview based on P1AIS_01 through P1AIS_19, and P1AIU_01 through P1AIU_19 (FS085)), BCFTHRC (father race from birth certificate)	1 = Yes, 2 = No
X1HFASBC	Family/HH	Flag indicating that the biological father's race from the birth certificate was used to code X1HFASN	RACE2 (Variable coded in parent interview based on P1ASN_01 through P1ASN_19, P1CHN_01 through P1CHN_19, P1FIL_01 through P1FIL_19, P1JPN_01 through P1JPN_19, P1KOR_01 through P1KOR_19, P1VTN_01 through P1VTN_19, P1AAS_01 through P1AAS_19, and P1ASU_01 through P1ASU_19 (FS085)), BCFTHRC (father race from birth certificate)	1 = Yes, 2 = No

See note at end of table.

Table 7-6. ECLS-B 9-month composite variables—Continued

Variable name	Category	Description	Derived from	Output values
X1HFPIBC	Family/HH	Flag indicating that the biological father's race from the birth certificate was used to code X1HFPCIL	RACE4 (Variable coded in parent interview based on P1NHI_01 through P1NHI_19, P1GC_01 through P1GC_19, P1SAM_01 through P1SAM_19, P1PIS_01 through P1PIS_19, and P1PIU_01 through P1PIU_19 (FS085)), BCFTHRC (father race from birth certificate)	1 = Yes, 2 = No
X1HFBLBC	Family/HH	Flag indicating that the biological father's race from the birth certificate was used to code X1HFBLCK	RACE3 (Variable coded in parent interview based on P1BLK_01 through P1BLK_19 (FS085)), BCFTHRC (father race from birth certificate)	1 = Yes, 2 = No
X1HFWHBC	Family/HH	Flag indicating that the biological father's race from the birth certificate was used to code X1HFWHT	RACE5 (Variable coded in parent interview based on P1WH_01 through P1WH_19 (FS085)), BCFTHRC (father race from birth certificate)	1 = Yes, 2 = No
X1HFHSBC	Family/HH	Flag indicating that the biological father's ethnicity from the birth certificate was used to code X1HFHISP	P1HSP_01 through P1HSP_19 (FS075), BCFTHHSP (father ethnicity from birth certificate)	1 = Yes, 2 = No
X1FTHED	Family/HH	Resident father's highest level of education	P1REDU (RI075), P1SPEDUC (SI050), F1HIGHGR (Resident Father (and Other Important People) SAQ-Q39)	1 = 8th grade or below, 2 = 9th to 12th grades, 3 = High school diploma/ equivalent, 4 = Voc/Tech program, 5 = Some college, 6 = Bachelor's degree, 7 = Graduate professional school/ no degree, 8 = Master's degree, 9 = Doctorate or professional degree

See note at end of table.

Table 7-6. ECLS-B 9-month composite variables—Continued

Variable name	Category	Description	Derived from	Output values
X1HFEMP	Family/HH	The work status of the father/male guardian in the household	P1WRKLWK (RI130), P1VACTN (RI135), P1PAIDHR (RI145), P1LKJOB4 (RI200), P1SWPBAG (RI205a), P1SWPRAG (RI205b), P1SWEMPD (RI205c), P1SWFRFH (RI205d), P1SWSNRS (RI205e), P1SWRDAD (RI205f), P1SWOTH (RI205g), P1SPWORK (SI065), P1SPVCTN (SI070), P1SPJBHR (SI080), P1SPLKWK (SI125), P1SPSWPA (SI130a), P1SPSWEA (SI130b), P1SPSWEM (SI130c), P1SPSWFF (SI130d), P1SPSWSR (SI130e), P1SPSWRA (SI130f), P1SPSWOT (SI130g), F1WRPTWK, F1VACTN, F1HRSWK, F1LOOKJB, F1PUBAGY, F1PRIAGY, F1EMPRES, F1FRIFAM, F1RESADS, F1WNTADS, F1FNWKOT, F1FNWKOS (not on data file) (Resident Father (and Other Important People) SAQ-Q46, 47, 54, 55)	1 = 35 hours or more per week, 2 = Less than 35 hours per week, 3 = Looking for work, 4 = Not in the labor force

See note at end of table.

Table 7-6. ECLS-B 9-month composite variables—Continued

Variable name	Category	Description	Derived from	Output values
X1FTHOCC	Family/HH	Resident father's occupation	P1EMPLN (RI180), P1WKINDS (RI185), P1JOBTYP (RI190), P1SPEMNM (SI105), P1SPWKIN (SI110), P1SPJBTY (SI115), F1WHOWRK, F1TYPWRK, F1KNDWRK (Resident Father (and Other Important People) SAQ-Q53) (variables not on data file), X1HFEMP (Composite)	1 = Management Occupations; 2 = Business and Financial Operations Occupations; 3 = Computer and Mathematical Occupations; 4 = Architecture and Engineering Occupations; 5 = Life, Physical, and Social Science Occupations; 6 = Community and Social Service Occupations; 7 = Legal Occupations; 8 = Education, Training, and Library Occupations; 9 = Arts, Design, Entertainment, Sports and Media Occupations; 10 = Healthcare Practitioner and Technical Occupations; 11 = Healthcare Support Occupations; 12 = Protective Service Occupations; 13 = Food Preparation and Serving Related Occupations; 14 = Building and Grounds Cleaning and Maintenance Occupations; 15 = Personal Care and Service Occupations; 16 = Sales and Related Occupations; 17 = Office and Administrative Support Occupations; 18 = Farming, Fishing and Forestry Occupations; 19 = Construction and Extraction Occupations; 20 = Installation, Maintenance and Repair Occupations; 21 = Production Occupations; 22 = Transportation and Material Moving Occupations; 23 = Military Specific Occupations; 90 = Unemployed, Retired, Disabled, Unclassified

See note at end of table.

Table 7-6. ECLS-B 9-month composite variables—Continued

Variable name	Category	Description	Derived from	Output values
X1FTHSCR	Family/HH	Father's occupation GSS prestige score	1989 GSS prestige scores, X1FTHOCC (Composite)	52.0 = Management Occupations; 48.0 = Business and Financial Operations Occupations; 58.4 = Computer and Mathematical Occupations; 59.7 = Architecture and Engineering Occupations; 61.2 = Life, Physical, and Social Science Occupations; 50.4 = Community and Social Service Occupations; 63.3 = Legal Occupations; 62.0 = Education, Training, and Library Occupations; 50.2 = Arts, Design, Entertainment, Sports and Media Occupations; 64.2 = Healthcare Practitioner and Technical Occupations; 41.4 = Healthcare Support Occupations; 50.7 = Protective Service Occupations; 27.1 = Food Preparation and Serving Related Occupations; 30.2 = Building and Grounds Cleaning and Maintenance Occupations; 33.2 = Personal Care and Service Occupations; 34.0 = Sales and Related Occupations; 37.7 = Office and Administrative Support Occupations; 31.7 = Farming, Fishing and Forestry Occupations; 36.4 = Construction and Extraction Occupations; 40.1 = Installation, Maintenance and Repair Occupations; 35.3 = Production Occupations; 36.5 = Transportation and Material Moving Occupations; 56.5 = Military Specific Occupations

See note at end of table.

Table 7-6. ECLS-B 9-month composite variables—Continued

Variable name	Category	Description	Derived from	Output values
X1MOMID	Family/HH	Household roster number of resident mother	P1RSREL (IN035), P1NRREL (IN060), P1REL_01 through P1REL_19 (FS040), P1NRS_01 through P1NRS_19 (FS065)	Continuous
X1MOMTYP	Family/HH	Type of resident mother; indicates whether the birth, adoptive, step-, or foster mother of the focal child resides in the household with the focal child	P1RSREL (IN035), P1MOMREL (IN040), P1PARTNR (FS030), P1REL_01 through P1REL_19 (FS040), P1MOM_01 through P1MOM_19 (FS045)	1 = Birth mother, 2 = Adoptive mother, 3 = Stepmother, 4 = Foster mother, 5 = Rel/nonrel resp–mother figure, 6 = Fathers partner–mother figure, 7 = Rel/nonrel resps partner–mother figure, 8 = Mother, unknown type, 9 = No resident mother
X1HMAGE	Family/HH	Age of resident mother	P1RSDOBC (IN030), P1RSDOBM (IN031), P1RSDOBY (IN033), P1RSAGEY (IN033b), P1RSREL (IN035), P1AGE_01 through P1AGE_19 (FS015), P1PARTNR (FS030), P1SPDOBM (SI031), P1SPDOBY (SI033), F1BDTEMM, F1BDTEDD, F1BDTEYY (Resident father (and Other Important People) SAQ)), date of parent interview (FMS), birth certificate data	Continuous

See note at end of table.

Table 7-6. ECLS-B 9-month composite variables—Continued

Variable name	Category	Description	Derived from	Output values
X1HMRACE	Family/HH	Race and ethnicity of the mother or female guardian in the household	RACE1, RACE2, RACE3, RACE4, RACE5, RACE6 (Variables coded in parent interview based on FS085), and P1HSP_01 through P1HSP_19 (FS075)).	1 = White, non-Hispanic, 2 = Black or African American, non-Hispanic, 3 = Hispanic, race specified, 4 = Hispanic, no race specified, 5 = Asian, non-Hispanic, 6 = Native Hawaiian or other Pacific Islander, non-Hispanic, 7 = American Indian or Alaska Native, non-Hispanic, 8 = More than 1 race, non-Hispanic
X1HMAMIN	Family/HH	Mother or female guardian in the household is American Indian	RACE1 (Variable coded in parent interview based on P1AIS_01 through P1AIS_19, and P1AIU_01 through P1AIU_19 (FS085)), BCMOMRC (mother race from birth certificate)	1 = Yes, 2 = No
X1HMASN	Family/HH	Mother or female guardian in the household is Asian	RACE2 (Variable coded in parent interview based on P1ASN_01 through P1ASN_19, P1CHN_01 through P1CHN_19, P1FIL_01 through P1FIL_19, P1JPN_01 through P1JPN_19, P1KOR_01 through P1KOR_19, P1VTN_01 through P1VTN_19, P1AAS_01 through P1AAS_19, and P1ASU_01 through P1ASU_19 (FS085)), BCMOMRC (mother race from birth certificate)	1 = Yes, 2 = No

See note at end of table.

Table 7-6. ECLS-B 9-month composite variables—Continued

Variable name	Category	Description	Derived from	Output values
X1HMPCIL	Family/HH	Mother or female guardian in the household is Pacific Islander	RACE4 (Variable coded in parent interview based on P1NHI_01 through P1NHI_19, P1GC_01 through P1GC_19, P1SAM_01 through P1SAM_19, P1PIS_01 through P1PIS_19, and P1PIU_01 through P1PIU_19 (FS085)), BCMOMRC (mother race from birth certificate)	1 = Yes, 2 = No
X1HMBLCK	Family/HH	Mother or female guardian in the household is Black	RACE3 (Variable coded in parent interview based on P1BLK_01 through P1BLK_19 (FS085)), BCMOMRC (mother race from birth certificate)	1 = Yes, 2 = No
X1HMWHT	Family/HH	Mother or female guardian in the household is White	RACE5 (Variable coded in parent interview based on P1WH_01 through P1WH_19 (FS085)), BCMOMRC (mother race from birth certificate)	1 = Yes, 2 = No
X1HMHISP	Family/HH	Mother or female guardian in the household is Hispanic	P1HSP_01 through P1HSP_19 (FS075), BCMOMHSP (mother ethnicity from birth certificate)	1 = Yes, 2 = No
X1HMMLRC	Family/HH	Mother or female guardian in the household is more than one race, unspecified	RACE6 (Variable coded from parent interview based on P1ARS_01 through P1ARS_19 (FS085), and P1ARU_01 through P1ARU_19 (FS089)), BCMOMRC (mother race from birth certificate)	1 = Yes, 2 = No

See note at end of table.

Table 7-6. ECLS-B 9-month composite variables—Continued

Variable name	Category	Description	Derived from	Output values
X1HMAIBC	Family/HH	Flag indicating that the biological mother's race from the birth certificate was used to code X1HMAMIN	RACE1 (Variable coded in parent interview based on P1AIS_01 through P1AIS_19, and P1AIU_01 through P1AIU_19 (FS085)), BCMOMRC (mother race from birth certificate)	1 = Yes, 2 = No
X1HMASBC	Family/HH	Flag indicating that the biological mother's race from the birth certificate was used to code X1HMASN	RACE2 (Variable coded in parent interview based on P1ASN_01 through P1ASN_19, P1CHN_01 through P1CHN_19, P1FIL_01 through P1FIL_19, P1JPN_01 through P1JPN_19, P1KOR_01 through P1KOR_19, P1VTN_01 through P1VTN_19, P1AAS_01 through P1AAS_19, and P1ASU_01 through P1ASU_19 (FS085)), BCMOMRC (mother race from birth certificate)	1 = Yes, 2 = No
X1HMPIBC	Family/HH	Flag indicating that the biological mother's race from the birth certificate was used to code X1HMPCIL	RACE4 (Variable coded in parent interview based on P1NHI_01 through P1NHI_19, P1GC_01 through P1GC_19, P1SAM_01 through P1SAM_19, P1PIS_01 through P1PIS_19, and P1PIU_01 through P1PIU_19 (FS085)), BCMOMRC (mother race from birth certificate)	1 = Yes, 2 = No
X1HMBLBC	Family/HH	Flag indicating that the biological mother's race from the birth certificate was used to code X1HMBLCK	RACE3 (Variable coded in parent interview based on P1BLK_01 through P1BLK_19 (FS085)), BCMOMRC (mother race from birth certificate)	1 = Yes, 2 = No

See note at end of table.

Table 7-6. ECLS-B 9-month composite variables—Continued

Variable name	Category	Description	Derived from	Output values
X1HMWHBC	Family/HH	Flag indicating that the biological mother's race from the birth certificate was used to code X1HMWHT	RACE5 (Variable coded in parent interview based on P1WH_01 through P1WH_19 (FS085)), BCMOMRC (mother race from birth certificate)	1 = Yes, 2 = No
X1HMHSBC	Family/HH	Flag indicating that the biological mother's ethnicity from the birth certificate was used to code X1HMHISP	P1HSP_01 through P1HSP_19 (FS075), BCMOMHSP (mother ethnicity from birth certificate)	1 = Yes, 2 = No
X1MOMED	Family/HH	Mother's highest level of education	P1REDU (RI075), P1SPEDUC (SI050), F1HIGHGR (Resident Father (and Other Important People) SAQ-Q39), BCMOMED (education of mother detail from birth certificate)	1 = 8th grade or below, 2 = 9th to 12th grades, 3 = High school diploma/ equivalent, 4 = Voc/Tech program, 5 = Some college, 6 = Bachelor's degree, 7 = Graduate professional school/ no degree, 8 = Master's degree, 9 = Doctorate or professional degree

See note at end of table.

Table 7-6. ECLS-B 9-month composite variables—Continued

Variable name	Category	Description	Derived from	Output values
X1HMEMP	Family/HH	The work status of the mother or female guardian in the household	P1WRKLWK (RI130), P1VACTN (RI135), P1PAIDHR (RI145), P1LKJOB4 (RI200), P1SWPBAG (RI205a), P1SWPRAG (RI205b), P1SWEMPD (RI205c), P1SWFRFH (RI205d), P1SWSNRS (RI205e), P1SWRDAD (RI205f), P1SWOTH (RI205g), P1SPWORK (SI065), P1SPVCTN (SI070), P1SPJBHR (SI080), P1SPLKWK (SI125), P1SPSWPA (SI130a), P1SPSWEA (SI130b), P1SPSWEM (SI130c), P1SPSWFF (SI130d), P1SPSWSR (SI130e), P1SPSWRA (SI130f), P1SPSWOT (SI130g), F1WRPTWK, F1VACTN, F1HRSWK, F1LOOKJB, F1PUBAGY, F1PRIAGY, F1EMPRES,F1FRIFAM, F1RESADS, F1WNTADS, F1FNWKOT, F1FNWKOS (not on data file) (Resident Father (and Other Important People) SAQ-Q46, 47, 54, 55)	1 = 35 hours or more per week, 2 = Less than 35 hours per week, 3 = Looking for work, 4 = Not in the labor force

See note at end of table.

Table 7-6. ECLS-B 9-month composite variables—Continued

Variable name	Category	Description	Derived from	Output values
X1MOMOCC	Family/HH	Mother's occupation	P1EMPLN (RI180), P1WKINDS (RI185), P1JOBTYP (RI190), P1SPEMNM (SI105), P1SPWKIN (SI110), P1SPJBTY (SI115), F1WHOWRK, F1TYPWRK, F1KNDWRK (Resident Father (and Other Important People) SAQ-Q53), (variables not on data file) X1HMEMP (Composite)	1 = Management Occupations; 2 = Business and Financial Operations Occupations; 3 = Computer and Mathematical Occupations; 4 = Architecture and Engineering Occupations; 5 = Life, Physical and Social Science Occupations; 6 = Community and Social Service Occupations; 7 = Legal Occupations; 8 = Education, Training and Library Occupations; 9 = Arts, Design, Entertainment, Sports and Media Occupations; 10 = Healthcare Practitioner and Technical Occupations; 11 = Healthcare Support Occupations; 12 = Protective Service Occupations; 13 = Food Preparation and Serving Related Occupations; 14 = Building and Grounds Cleaning and Maintenance Occupations; 15 = Personal Care and Service Occupations; 16 = Sales and Related Occupations; 17 = Office and Administrative Support Occupations; 18 = Farming, Fishing and Forestry Occupations; 19 = Construction and Extraction Occupations; 20 = Installation, Maintenance and Repair Occupations; 21 = Production Occupations; 22 = Transportation and Material Moving Occupations; 23 = Military Specific Occupations; 90 = Unemployed, Retired, Disabled, Unclassified

See note at end of table.

Table 7-6. ECLS-B 9-month composite variables—Continued

Variable name	Category	Description	Derived from	Output values
X1MOMSCR	Family/HH	Mother's occupation GSS prestige score	1989 GSS prestige scores, X1MOMOCC (Composite)	52.0 = Management Occupations; 48.0 = Business and Financial Operations Occupations; 58.4 = Computer and Mathematical Occupations; 59.7 = Architecture and Engineering Occupations; 61.2 = Life, Physical, and Social Science Occupations; 50.4 = Community and Social Service Occupations; 63.3 = Legal Occupations; 62.0 = Education, Training, and Library Occupations; 50.2 = Arts, Design, Entertainment, Sports and Media Occupations; 64.2 = Healthcare Practitioner and Technical Occupations; 41.4 = Healthcare Support Occupations; 50.7 = Protective Service Occupations; 27.1 = Food Preparation and Serving Related Occupations; 30.2 = Building and Grounds Cleaning and Maintenance Occupations; 33.2 = Personal Care and Service Occupations; 34.0 = Sales and Related Occupations; 37.7 = Office and Administrative Support Occupations; 31.7 = Farming, Fishing and Forestry Occupations; 36.4 = Construction and Extraction Occupations; 40.1 = Installation, Maintenance and Repair Occupations; 35.3 = Production Occupations; 36.5 = Transportation and Material Moving Occupations; 56.5 = Military Specific Occupations

See note at end of table.

Table 7-6. ECLS-B 9-month composite variables—Continued

Variable name	Category	Description	Derived from	Output values
X1HMWGHT	Family/HH	Resident mother/female guardian’s weight	C1WHADT1, C1WHADT2, C1ADWGT1, C1ADWGT2	Continuous
X1PARED	Family/HH	Highest level of education for the focal child’s parents or nonparental guardians who reside in the household. If only one parent or guardian resides in the household, then this variable reflects that parent’s education level.	X1FTHED, X1MOMED (composites)	1 = 8th grade or below, 2 = 9th to 12th grades, 3 = High school diploma/ equivalent, 4 = Voc/Tech program, 5 = Some college, 6 = Bachelor’s degree, 7 = Graduate professional school/ no degree, 8 = Master’s degree, 9 = Doctorate or professional degree
X1NRFED	Family/HH	Nonresident father’s highest level of education	P1BFEDUC (BF050), N1HIGHGR (Nonresident Father SAQ-Q26)	1 = 8th grade or below, 2 = 9th to 12th grades, 3 = High school diploma/ equivalent, 4 = Voc/Tech program, 5 = Some college, 6 = Bachelor’s degree, 7 = Graduate professional school/ no degree, 8 = Master’s degree, 9 = Doctorate or professional degree
X1NRFEMP	Family/HH	The work status of the non-resident father	N1WRPYWK, N1VACTN, N1HRSWK, N1LOOKJB, N1PUBAGY, N1PRIAGY, N1EMPRES, N1FRIFAM, N1RESADS, N1WNTADS, N1FNWKOT, N1FNWKOS (not on data file) (Nonresident father SAQ-Q29-33)	1 = 35 hours or more per week, 2 = Less than 35 hours per week, 3 = Looking for work, 4 = Not in the labor force
X1CHID	Family/HH	Household roster number of focal child	Household roster	Continuous

See note at end of table.

Table 7-6. ECLS-B 9-month composite variables—Continued

Variable name	Category	Description	Derived from	Output values
X1RESPID	Family/HH	Household roster number of parent respondent	P1RSREL (IN035), P1NRREL (IN060), household roster	Continuous
X1RSPREL	Family/HH	Respondent relationship to child	P1RSREL (IN035), P1MOMREL (IN040), P1DADREL (IN045)	1 = Biological mother, 2 = Other mother type, 3 = Biological father, 4 = Other father type, 5 = Non-parent relative, 6 = Non-relative
X1HHURBN	Derived	Indicates whether the household is located in an urban or rural area	Sampling Frame Data, 2000 Census of Population summary tapes containing urban/rural information	1 = Urban, inside UA, 2 = Urban, inside UC, 3 = Rural
X1HHREGN	Derived	Indicates the geographic region in which the sampled household is located	P1CSTATE	1 = Northeast: CT, ME, MA, NH, RI, VT, NJ, NY, PA; 2 = Midwest: IL, IN, MI, OH, WI, IA, KS, MN, MO, NE, ND, SD; 3 = South: DE, DC, FL, GA, MD,NC, SC, VA, WV, AL, KY, MS, TN, AR, LA, OK, TX; 4 = West: AZ, CO, ID, MT, NV, NM, UT, WY, AK, CA, HI, OR, WA
X1LNGPRT	Derived	Language in which parent interview conducted	R1PRNLNG	1 = English, 2 = Spanish, 3 = Chinese, 4 = Other Asian language, 5 = Other non-Asian language
X1LNGNCT	Derived	Language in which NCATS conducted	C1LNGCMP (not on data file)	1 = English, 2 = Spanish, 3 = Chinese, 4 = Other
X1BTHWGT	Derived	Child birth weight status	DBIRWT, BCBRTHWT	1 = Normal birth weight, 2 = Moderately low birth weight, 3 = Very low birth weight
X1TWSAMP	Derived	Child sampled as part of twin	W1VARSTR (Sampling frame variable not on file)	1 = Yes, 2 = No

See note at end of table.

Table 7-6. ECLS-B 9-month composite variables—Continued

Variable name	Category	Description	Derived from	Output values
X1AISAMP	Derived	Child sampled as part of American Indian population	CHRACE (Sampling frame variable not on file)	1 = Yes, 2 = No
X1ASISMP	Derived	Child sampled as part of Asian population	CHRACE (Sampling frame variable not on file)	1 = Yes, 2 = No
X1CHNSMP	Derived	Child sampled as part of Chinese population	CHRACE (Sampling frame variable not on file)	1 = Yes, 2 = No
X1INTVID	Derived	Interviewer ID	Parent CAPI application	Continuous
X1INTREF	Derived	Interview was initially refused	Field Management System (FMS)	1 = Yes, 2 = No
X1MBRTST	Derived	Multiple birth status	BCPLURAL	1 = Singleton, 2 = Twin, 3 = Higher order
X1INTRST	Derived	Interpreter Status	R1PRNLNG (IR120), R1PRLNOS (IR121) (not on data file), RIPRNTRN (IR125)	1 = Home interpreter, 2 = Community interpreter, 3 = No interpreter-bilingual interviewer, 4 = No interpreter-conducted in English
X1MTLMOD	Derived	Mental: Number of items modified	C1BAYMOD, C1MODI01-C1MODI34 (not on data file)	Continuous
X1MTRMOD	Derived	Motor: Number of items modified	C1BAYMOD, C1MODI01-C1MODI34 (not on data file)	Raw count of the number of BSF-R motor items modified
X1TWIN	Derived	Household has sampled twins	Twin flag (Sampling frame variable not on file), P1TWINHH (IN010)	0 = No twin in HH, 1 = Twin in HH

See note at end of table.

Table 7-6. ECLS-B 9-month composite variables—Continued

Variable name	Category	Description	Derived from	Output values
X1MTLSKP	Derived	Mental: Number of items skipped	C1NOTADM, C1NOTA01-C1NOTA53 (not on data file)	Continuous
X1MTRSKP	Derived	Motor: Number of items skipped	C1NOTADM, C1NOTA01-C1NOTA53 (not on data file)	Continuous
X1STATPI	Indicator	Status of parent CAPI instrument	P1RAISE (CD125), the final question in section CD of the parent CAPI instrument (A "complete" is defined by the respondent completing the instrument through section CD).	0 = No, 1 = Yes
X1STTPSQ	Indicator	Status of parent self-administered questionnaire	PSAQ CADE	0 = No, 1 = Yes
X1STBMTL	Indicator	Status of Bayley mental assessment	X1MTLMOD (composite), X1MTLSCL (composite), Bayley mental scale completed (FMS variable not on file)	1 = BSF-R mental administered, no modifications, 2 = BSF-R mental administered, with modifications, 3 = BSF-R mental administered, unknown modification status, 4 = Not scoreable, 5 = Not administered
X1STBMTR	Indicator	Status of Bayley motor assessment	X1MTRMOD (composite), X1MTRSCL (composite), BSF-R motor scale completed (FMS variable not on file)	1 = BSF-R motor administered, no modifications, 2 = BSF-R motor administered, with modifications, 3 = BSF-R motor administered, unknown modification status, 4 = Not scoreable, 5 = Not administered
X1STTNCT	Indicator	Status of NCATS assessment	C1ABLCDE, C1LNGUAG (not on data file)	0 = No NCATS tape, 1 = NCATS tape coded, 2 = NCATS tape not coded, foreign language, 3 = NCATS tape not coded, uncodeable

See note at end of table.

Table 7-6. ECLS-B 9-month composite variables—Continued

Variable name	Category	Description	Derived from	Output values
X1STATPM	Indicator	Status of physical measurements	C1CHLGT1, C1CHLGT2, C1CHWGT1, C1CHWGT2	1 = Both length and weight measurements are present, 2 = Only one measurement (length or weight) is present, 3 = Neither length nor weight measurements are present
X1STTRFQ	Indicator	Status of resident father questionnaire	FSAQ CADE	1 = Yes, 2 = No, 0 = Not eligible for resident father questionnaire
X1STTNRQ	Indicator	Status of nonresident father questionnaire	NRFSAQ CADE, FMS, P1DADREL (IN045), P1BFLVNG (MH050), P1DAD_01 through P1DAD_19 (FS050), P1FCM260 (CM260) (not on data file)	1 = Yes, 2 = No, nonresident father present but no mother permission, 3 = No, refusal/ incomplete by nonresident father, 4 = No, nonresident father ineligible (lack of contact), 5 = No because no nonresident father, 6 = No, respondent is not biological mother, do not ask about nonresident father
X1SYNDRM	Child	Presence of Down's syndrome, Turner's syndrome, or spina bifida	Combines P1DOWNS (child has Down's syndrome), P1TRNRS (child has Turner's syndrome), and P1SPNBFD (child has spina bifida)	1 = Child has Down's, Turner's, or spina bifida, 2 = Child does not have Down's, Turner's, or spina bifida
X1PRGCMB	Indirect child	Participation in early intervention programs	Combines P1PRGOT and P1PRGSD	1 = Child participates in early intervention program through school district or other source, 2 = Child does not participate in early intervention program
X1CHPCMB	Indirect child	Why child doesn't have CHIP	Combines P1CHIPOT, P1CHIPKN, P1CHIPJN, P1CHIPBR and P1CHIPSK	1 = Didn't know they qualified or how to join, child wasn't sick, or it was too much bother, 2 = Other reason

See note at end of table.

Table 7-6. ECLS-B 9-month composite variables—Continued

Variable name	Category	Description	Derived from	Output values
X1CVGCMB	Indirect child	Why no health coverage in past	Combines P1PASTO, P1PASTDV, P1PASTRE, P1PASTNO, P1PASTCT, and P1PASTRF	1 = Divorced, remarried, too costly, refused, or employer doesn't offer, 2 = Other reason
X1CRCMB	Indirect child	Prenatal care payment	Combines P1PNCOTH and P1NOPYMT	1 = No payment necessary or other source of payment, 2 = Not applicable/refused/not ascertained
X1LANGME	Family/HH	Language spoken at home is Middle Eastern	Combines P1LANG18 and P1LANG01	1 = Language spoken at home is Arabic or other Middle Eastern language, 2 = Not applicable/refused/not ascertained
X1LANGEU	Family/HH	Language spoken at home is European	Combines P1LANG19, P1LANG04, P1LANG15, P1LANG05, P1LANG06, P1LANG07, P1LANG10, and P1LANG11	1 = Language spoken at home is French, German, Greek, Italian, Polish, Portuguese, other Eastern European, or other Western European language, 2 = Not applicable/refused/not ascertained
X1LANGOA	Family/HH	Language spoken at home is Asian	Combines P1LANG03, P1LANG08, P1LANG09, P1LANG13, P1LANG20 P1LANG21,and P1LANG22	1 = Language spoken at home is Filipino, Japanese, Korean, Vietnamese, Indian subcontinent, southeast Asian, or Pacific Islander, 2 = Not applicable/refused/not ascertained
X1LANGOL	Family/HH	Language spoken at home is sign or other language	Combines P1LANG14, P1LANG16, P1LANG17, and P1LANG23	1 = Language spoken at home is sign language or other, 2 = Not applicable/refused/not ascertained
X1CGDCMB	Family/HH	How often does father change diapers	Combines P1CGDIAP and F1CGDIAP	If resident father responded on resident father questionnaire (F1CGDIAP), then X1CGDCMB takes this value, otherwise X1CGDCMB = P1CGDIAP

SOURCE: U.S. Department of Education, National Center for Education Statistics, Early Childhood Longitudinal Study, Birth Cohort, Nine-Month Data Collection, 2001–02.

7.7 Masked Variables

All the variables from the ECLS-B 9-month restricted-use file are included in the same order on the ECLS-B 9-month public-use file. For some of the variables, certain categories were modified. The value labels for those masked variables were updated from the restricted-use variables to reflect the new categories that were created during the masking process.

Variables on the restricted-use files were modified in different ways based on the disclosure analysis NCES conducted in order to protect the identity of the respondents and children. There are several types of modifications on the public-use files.

- Outliers are top- or bottom- coded to prevent identification of unique parents and children without affecting the overall data quality.
- Variables with too few cases and a sparse distribution are suppressed in the public-use files. The values for these variables were set to −2 and labeled "suppressed" in the ECB.
- For one group of variables, values were modified by "data swapping." This process removes a reported value and replaces it with a reported value from a different respondent for a subset of the records.
- Certain continuous variables are modified into categorical variables, and certain categorical variables have their categories collapsed in the public-use file. While this protects the cases from disclosure risk, these variables can still be used in all different kinds of analysis such as regression analysis.
- Ten new composites were added to the 9-month public-use file. These composites provide summary information about topics for which more detailed information has been suppressed in response to the disclosure risk analysis. See the end of table 7-6 for a description of these new composites.

There is a comment field in the variable frequency distribution view screen of the electronic code book that displays a comment for each masked variable indicating whether the variable from the restricted-use file has been recoded or suppressed. Variables that were recoded in any way during the data masking process display the comment, "These data recoded for respondent confidentiality." Variables that were suppressed on the public-use file for protection of the respondent or child from identification display the comment, "These data suppressed for respondent confidentiality" and all values for the variable are set to equal −2 for that variable.

Table 7-7 presents the list of masked variables. The tables display the variable name, variable label, and the comment displayed in the electronic code book indicating if the variable was recoded or suppressed. The table is sorted sequentially by the variable Field ID (see section 8.3.1.1 for how to use the variable Field ID.)

Table 7-7. Recoded and suppressed data on the ECLS-B 9-Month Public-Use Data File: 2001–02

Field ID	Variable	Field label	Comment
8	X1ASAGE	X1 CHILDS AGE AT TIME DIRECT ASSESSMENT	These data recoded for respondent confidentiality. Recoded 8.1 or less to 8.2 and 13.0 or more to 12.9.
11	X1CHRACE	X1 RACE/ETHNICITY - CHILD	These data recoded for respondent confidentiality.
15	X1CHASN	X1 CHILD – ASIAN	These data suppressed for respondent confidentiality.
16	X1CHPCIL	X1 CHILD – PAC ISLANDER	These data suppressed for respondent confidentiality.
18	X1CHMLRC	X1 CHILD – MULTIRACIAL UNSPECIFIED	These data suppressed for respondent confidentiality.
19	X1CHPREM	X1 PREMATURITY OF CHILD FROM BIRTH CERT	These data recoded for respondent confidentiality. Recoded 78 or more to 77.
23	X1PRIMNW	X1 PRIM CARE ARRNGMNT WHERE MOST HRS/WK	These data recoded for respondent confidentiality.
51	X1RSPREL	X1 RESPONDENTS RELATIONSHIP TO CHILD	These data recoded for respondent confidentiality.
53	X1MOMTYP	X1 TYPE RES MOTHER-BIRTH/ADOPT/STEP/FOSTER	These data recoded for respondent confidentiality.
54	X1HMAGE	X1 AGE OF RESIDENT MOTHER	These data recoded for respondent confidentiality.
55	X1HMRACE	X1 RACE/ETHNICITY–HH MOTHER/FEMALE GUARD	These data recoded for respondent confidentiality.
60	X1HMPCIL	X1 HH MTHR/FEMALE GUARD-PAC ISLANDER	These data suppressed for respondent confidentiality.
62	X1HMMLRC	X1 HH MOTHER/FEMALE GUARD-MULTIRACE UNSP	These data suppressed for respondent confidentiality.
63	X1MOMED	X1RES MOTHER HIGHEST EDUCATION LEVEL	These data recoded for respondent confidentiality.
65	X1MOMOCC	X1 HH MOTHER/FEMALE GUARD-OCCUPATION	These data suppressed for respondent confidentiality.
66	X1MOMSCR	X1 HH MOTHER-OCC GSS PRESTIGE SCORE	These data recoded for respondent confidentiality.
69	X1FTHTYP	X1 TYPE RES FATHER-BIRTH/ADOPT/STEP/FOSTER	These data recoded for respondent confidentiality.
70	X1HFAGE	X1 AGE OF RESIDENT FATHER	These data recoded for respondent confidentiality. Recoded 17 or less to 18 and 51 or more to 50.
71	X1HFRACE	X1 RACE/ETHNICITY – HH FATHER/MALE GUARD	These data recoded for respondent confidentiality.
76	X1HFPCIL	X1 HH FATHER/MALE GUARD-PAC ISLANDER	These data suppressed for respondent confidentiality.
78	X1HFMLRC	X1 HH FATHER/MALE GUARD-MULTIRACE UNSP	These data suppressed for respondent confidentiality.
79	X1FTHED	X1 RES FATHER HIGHEST EDUCATION LEVEL	These data recoded for respondent confidentiality.
81	X1FTHOCC	X1 HH FATHER/MALE GUARD-OCCUPATION	These data suppressed for respondent confidentiality.
83	X1NRFED	X1NON-RES FATHER HIGHEST EDUCATIONLEVEL	These data recoded for respondent confidentiality.
86	X1HPARNT	X1 CH PARENTS WHO RESIDE IN HOUSEHOLD	These data recoded for respondent confidentiality.
87	X1MARSTA	X1 MARITAL STATUS OF PARENT(S) IN HH	These data recoded for respondent confidentiality.
88	X1NUMSIB	X1 NUMB SIBLINGS IN HH-W/FSTR PRN OTH CHL	These data recoded for respondent confidentiality. Recoded 7 or more to 6.
90	X1LESS18	X1 NUMBER OF HH MEMBERS LESS THAN 18	These data recoded for respondent confidentiality. Recoded 9 or more to 8.
91	X118OVER	X1 NUM HH MEMBERS 18 AND OLDER	These data recoded for respondent confidentiality. Recoded less than 1 to 1 and 8 or more to 7.
92	X1PARED	X1 HH PARENT HIGHEST EDUCATION LEVEL	These data recoded for respondent confidentiality.
127	X1HMWHBC	X1 BIRTH CERT USAGE INDICATOR-X1HMWHT	These data suppressed for respondent confidentiality.
128	X1HMBLBC	X1 BIRTH CERT USAGE INDICATOR-X1HMBLCK	These data suppressed for respondent confidentiality.
129	X1HMHSBC	X1 BIRTH CERT USAGE INDICATOR-X1HMHISP	These data suppressed for respondent confidentiality.
130	X1HMASBC	X1 BIRTH CERT USAGE INDICATOR-X1HMASN	These data suppressed for respondent confidentiality.
131	X1HMPIBC	X1 BIRTH CERT USAGE INDICATOR-X1HMPCIL	These data suppressed for respondent confidentiality.
132	X1HMAIBC	X1 BIRTH CERT USAGE INDICATOR-X1HMAMIN	These data suppressed for respondent confidentiality.
133	X1HFWHBC	X1 BIRTH CERT USAGE INDICATOR-X1HFWHT	These data suppressed for respondent confidentiality.
134	X1HFBLBC	X1 BIRTH CERT USAGE INDICATOR-X1HFBLCK	These data suppressed for respondent confidentiality.
135	X1HFHSBC	X1 BIRTH CERT USAGE INDICATOR-X1HFHISP	These data suppressed for respondent confidentiality.
136	X1HFASBC	X1 BIRTH CERT USAGE INDICATOR-X1HFASN	These data suppressed for respondent confidentiality.
137	X1HFPIBC	X1 BIRTH CERT USAGE INDICATOR-X1HFPCIL	These data suppressed for respondent confidentiality.
138	X1HFAIBC	X1 BIRTH CERT USAGE INDICATOR-X1HFAMIN	These data suppressed for respondent confidentiality.
149	X1IFDDOC	X1 HH FATHER/MALE GUARD-OCCPTN IMPFLG	These data suppressed for respondent confidentiality.

See note at end of table.

Table 7-7. Recoded and suppressed data on the ECLS-B 9-Month Public-Use Data File: 2001–02—Continued

Field ID	Variable	Field label	Comment
166	P1AGE_01	P1 FS015 AGE IN YEARS OF PERSON 1	These data suppressed for respondent confidentiality.
167	P1SEX_01	P1 FS020 GENDER OF PERSON 1	These data suppressed for respondent confidentiality.
168	P1REL_01	P1 FS040 SPECIFY RELATIONSHIP-PERSON1	These data suppressed for respondent confidentiality.
169	P1MOM_01	P1 FS045 SPECIFIC MOTHER TYPE-PERSON1	These data suppressed for respondent confidentiality.
170	P1DAD_01	P1 FS050 SPECIFIC FATHER TYPE-PERSON1	These data suppressed for respondent confidentiality.
171	P1SIS_01	P1 FS055 SPECIFIC SISTER TYPE-PERSON1	These data suppressed for respondent confidentiality.
172	P1BRO_01	P1 FS060 SPECIFIC BROTHR TYPE-PERSON1	These data suppressed for respondent confidentiality.
173	P1NRS_01	P1 FS065 SPEC NONRELATIVE DESC-PERSON1	These data suppressed for respondent confidentiality.
174	P1HSP_01	P1 FS075 HISPANIC/LATIN ORIGIN-PERSON1	These data suppressed for respondent confidentiality.
175	P1MEX_01	P1 FS080 MEXICAN HISPANIC GROUP-PERS1	These data suppressed for respondent confidentiality.
176	P1PR_01	P1 FS080 PUERTO RICAN HISP GROUP-PERS1	These data suppressed for respondent confidentiality.
177	P1CUB_01	P1 FS080 CUBAN HISPANIC GRP-PERS1	These data suppressed for respondent confidentiality.
178	P1HGS_01	P1 FS080 ANOTH SPECIFIC HISP GRP-PERS1	These data suppressed for respondent confidentiality.
179	P1HGU_01	P1 FS080 ANOTH UNSPEC HISP GRP-PERS1	These data suppressed for respondent confidentiality.
180	P1WH_01	P1 FS085 RACE WHITE – PERSON1	These data suppressed for respondent confidentiality.
181	P1BLK_01	P1 FS085 RACE BLACK - PERSON1	These data suppressed for respondent confidentiality.
182	P1AIS_01	P1 FS085 RACE AMER INDIAN TRB-PERS1	These data suppressed for respondent confidentiality.
183	P1ASN_01	P1 FS085 RACE ASIAN INDIAN-PERS1	These data suppressed for respondent confidentiality.
184	P1CHN_01	P1 FS085 RACE CHINESE - PERS1	These data suppressed for respondent confidentiality.
185	P1FIL_01	P1 FS085 RACE FILIPINO – PERSON1	These data suppressed for respondent confidentiality.
186	P1JPN_01	P1 FS085 RACE JAPANESE – PERSON1	These data suppressed for respondent confidentiality.
187	P1KOR_01	P1 FS085 RACE KOREAN – PERSON1	These data suppressed for respondent confidentiality.
188	P1VTN_01	P1 FS085 RACE VIETNAMESE-PERSON1	These data suppressed for respondent confidentiality.
189	P1AAS_01	P1 FS085 RACE OTHER ASIAN- PERSON1	These data suppressed for respondent confidentiality.
190	P1NHI_01	P1 FS085 RACE NATIVE HAWAIIAN – PERS1	These data suppressed for respondent confidentiality.
191	P1GC_01	P1 FS085 RACE GUAM/CHAMORRO-PERSON1	These data suppressed for respondent confidentiality.
192	P1SAM_01	P1 FS085 SAMOAN – PERSON1	These data suppressed for respondent confidentiality.
193	P1PIS_01	P1 FS085 RACE SPEC OTH PACIFIC IS-PERS1	These data suppressed for respondent confidentiality.
194	P1ARS_01	P1 FS085 ANOTHER SPECIFIED RACE-PERS1	These data suppressed for respondent confidentiality.
195	P1AIU_01	P1 FS085 UNSPEC OTHER AMER INDIAN TRB-PERS1	These data suppressed for respondent confidentiality.
196	P1ASU_01	P1 FS085 UNSPEC OTHER ASIAN RACE-PERS1	These data suppressed for respondent confidentiality.
197	P1PIU_01	P1 FS085 UNSPEC OTH PAC ISLANDER-PERS1	These data suppressed for respondent confidentiality.
198	P1ARU_01	P1 FS085 ANOTH UNSPECIFIED RACE-PERS1	These data suppressed for respondent confidentiality.
796	P1CSTATE	P1 IN000ST STATE INTERVIEWED IN	These data suppressed for respondent confidentiality.
800	P1TWDODM	P1 IN015 TWINS MONTH OF DEATH	These data suppressed for respondent confidentiality.
809	P1STBMOM	P1 IN065 STATUS OF BIO-MOM	These data suppressed for respondent confidentiality.
810	P1BMDOD	P1 IN070 BIO-MOMS DATE OF DEATH	These data suppressed for respondent confidentiality.
838	P1LANG01	P1 HE010 LANG SPOKEN AT HOME-ARABIC	P1LANG01 suppressed for respondent confidentiality. These data have been combined with other data. See X1LANGME in table 7-6.
840	P1LANG03	P1 HE010 LANG SPOKEN AT HOME-FILIPINO	P1LANG03 suppressed for respondent confidentiality. These data have been combined with other data. See X1LANGOA in table 7-6.
841	P1LANG04	P1 HE010 LANG SPOKEN AT HOME-FRENCH	P1LANG04 suppressed for respondent confidentiality. These data have been combined with other data. See X1LANGEU in table 7-6.
842	P1LANG05	P1 HE010 LANG SPOKEN AT HOME-GERMAN	P1LANG05 suppressed for respondent confidentiality. These data have been combined with other data. See X1LANGEU in table 7-6.

See note at end of table.

Table 7-7. Recoded and suppressed data on the ECLS-B 9-Month Public-Use Data File: 2001–02—Continued

Field ID	Variable	Field label	Comment
843	P1LANG06	P1 HE010 LANG SPOKEN AT HOME-GREEK	P1LANG06 suppressed for respondent confidentiality. These data have been combined with other data. See X1LANGEU in table 7-6.
844	P1LANG07	P1 HE010 LANG SPOKEN AT HOME-ITALIAN	P1LANG07 suppressed for respondent confidentiality. These data have been combined with other data. See X1LANGEU in table 7-6.
845	P1LANG08	P1 HE010 LANG SPOKEN AT HOME-JAPANESE	P1LANG08 suppressed for respondent confidentiality. These data have been combined with other data. See X1LANGOA in table 7-6.
846	P1LANG09	P1 HE010 LANG SPOKEN AT HOME-KOREAN	P1LANG09 suppressed for respondent confidentiality. These data have been combined with other data. See X1LANGOA in table 7-6.
847	P1LANG10	P1 HE010 LANG SPOKEN AT HOME-POLISH	P1LANG10 suppressed for respondent confidentiality. These data have been combined with other data. See X1LANGEU in table 7-6.
848	P1LANG11	P1 HE010 LANG SPOKEN AT HOME-PORTUGUESE	P1LANG11 suppressed for respondent confidentiality. These data have been combined with other data. See X1LANGEU in table 7-6.
850	P1LANG13	P1 HE010 LANG SPOKEN AT HOME-VIETNAMESE	P1LANG13 suppressed for respondent confidentiality. These data have been combined with other data. See X1LANGOA in table 7-6.
851	P1LANG14	P1 HE010 LANG SPOKEN AT HOME-AFRICAN	P1LANG14 suppressed for respondent confidentiality. These data have been combined with other data. See X1LANGOL in table 7-6.
852	P1LANG15	P1 HE010 LANG SPOKEN AT HOME-EAST EUROPE	P1LANG15 suppressed for respondent confidentiality. These data have been combined with other data. See X1LANGEU in table 7-6.
853	P1LANG16	P1 HE010 LANG SPOKEN AT HOME-NAT AMERICA	P1LANG16 suppressed for respondent confidentiality. These data have been combined with other data. See X1LANGOL in table 7-6.
854	P1LANG17	P1 HE010 LANG SPOKEN AT HOME-SIGN LANGUAGE	P1LANG17 suppressed for respondent confidentiality. These data have been combined with other data. See X1LANGOL in table 7-6.
855	P1LANG18	P1 HE010 LANG SPOKEN AT HOME-MIDDLE EAST	P1LANG18 suppressed for respondent confidentiality. These data have been combined with other data. See X1LANGME in table 7-6.
856	P1LANG19	P1 HE010 LANG SPOKEN AT HOME-WEST EUROPE	P1LANG19 suppressed for respondent confidentiality. These data have been combined with other data. See X1LANGEU in table 7-6.
857	P1LANG20	P1 HE010 LANG SPOKEN AT HOME-INDIAN SUBC	P1LANG20 suppressed for respondent confidentiality. These data have been combined with other data. See X1LANGOA in table 7-6.
858	P1LANG21	P1 HE010 LANG SPOKEN AT HOME-STHEST ASIA	P1LANG21 suppressed for respondent confidentiality. These data have been combined with other data. See X1LANGOA in table 7-6.
859	P1LANG22	P1 HE010 LANG SPOKEN AT HOME-PAC ISLAND	P1LANG22 suppressed for respondent confidentiality. These data have been combined with other data. See X1LANGOA in table 7-6.
860	P1LANG23	P1 HE010 LANG SPOKEN AT HOME-OTHER	P1LANG23 suppressed for respondent confidentiality. These data have been combined with other data. See X1LANGOL in table 7-6.
862	P1PRMLNG	P1 HE020 PRIMARY LANGUAGE AT HOME	These data recoded for respondent confidentiality.
872	P1KYHH3	P1 HE030 SPKS OTH LANG-HH NUM-KEY ADULT3	These data suppressed for respondent confidentiality.
873	P1FRQOL3	P1 HE030 FREQ OTH LANG SPKN BY KY ADLT 3	These data suppressed for respondent confidentiality.
874	P1KYHH4	P1 HE030 SPKS OTH LANG-HH NUM-KEY ADULT4	These data suppressed for respondent confidentiality.
875	P1FRQOL4	P1 HE030 FREQ OTH LANG SPKN BY KY ADLT 4	These data suppressed for respondent confidentiality.

See note at end of table.

Table 7-7. Recoded and suppressed data on the ECLS-B 9-Month Public-Use Data File: 2001–02—Continued

Field ID	Variable	Field label	Comment
891	P1CGDIAP	P1 HE106A DAD RSP-CHANGE DIAPER	P1CGDIAP suppressed for respondent confidentiality. These data have been combined with other data. See X1CGDCMB in table 7-6.
953	P1RELTV	P1 CC040 WHICH RELATIVE GIVES MOST CARE	These data recoded for respondent confidentiality.
958	P1RCHLD	P1 CC075 # CHILDREN CARED FOR TOGETHER	These data recoded for respondent confidentiality. Recoded 10 or more to 9.
960	P1RLANG	P1 CC090 LANG USES IN CARE	These data recoded for respondent confidentiality.
962	P1RPAID	P1 CC105 AMOUNT PAID TO RELATIVE	These data recoded for respondent confidentiality.
970	P1NRNUM	P1 CC140 # CURRENT NON-REL ARRANGE	These data recoded for respondent confidentiality.
971	P1NPLACE	P1 CC145 LOCATION OF NON-REL CARE	These data recoded for respondent confidentiality.
975	P1NCHLD	P1 CC175 # CHILDREN CARED FOR	These data recoded for respondent confidentiality. Recoded 10 or more to 9.
977	P1NLANG	P1 CC190 LANG NON-REL USES IN CARE	These data recoded for respondent confidentiality.
979	P1NPAID	P1 CC205 AMOUNT PAID TO NON-REL	These data recoded for respondent confidentiality.
982	P1NAMT2	P1 CC215 # CHILDREN INCL IN NON-REL AMT	These data recoded for respondent confidentiality. Recoded 5 or more to 4.
987	P1CTRNUM	P1 CC240 # CURRENT CTR CARE ARRANGE	These data recoded for respondent confidentiality. Recoded 2 or more to 2.
988	P1CDAYS	P1 CC255 # OF DAYS/WK OF CTR CARE	These data recoded for respondent confidentiality. Recoded 6 or more to 5.
990	P1CCHLD	P1 CC265 # CHILDREN CARED FOR TOGETHER	These data recoded for respondent confidentiality. Recoded 2 or less to 3 and 10 or more to 9.
991	P1CADULT	P1 CC270 # ADULTS CARING FOR CHILDREN	These data recoded for respondent confidentiality. Recoded 7 or more to 6.
992	P1CLANG	P1 CC280 LANG PROVIDER USES IN CTR CARE	These data recoded for respondent confidentiality.
994	P1CPAID	P1 CC295 AMOUNT PAID FOR CTR CARE	These data recoded for respondent confidentiality.
998	P1CHROTH	P1 CC308 HRS AT OTHER CTR	These data suppressed for respondent confidentiality.
1011	P1SWAB	P1 CH008A IDENT TWIN TEST-CHEEK SWAB DNA	These data suppressed for respondent confidentiality.
1012	P1DNA	P1 CH008B IDENT TWIN TEST-BLOOD DNA TEST	These data suppressed for respondent confidentiality.
1017	P1QSTNR	P1 CH008G IDENT TWIN TEST-CMPLTD QUESTNR	These data suppressed for respondent confidentiality.
1018	P1OTTOLD	P1 CH008H IDENT TWIN TEST-OTH NOT DLVRY	These data suppressed for respondent confidentiality.
1021	P1DYSHSP	P1 CH015 # DAYS IN HOSPITAL	These data recoded for respondent confidentiality.
1025	P1VENTDY	P1 CH033 NUMBER OF DAYS ON VENTILATION	These data recoded for respondent confidentiality.
1027	P1APMONU	P1CH040 USE APNEA MONITOR AT HOME	These data suppressed for respondent confidentiality.
1108	P1DOWNS	P1 CH165I CHILD-DOWN SYNDROME	P1DOWNS suppressed for respondent confidentiality. These data have been combined with other data. See X1SYNDRM in table 7-6.
1109	P1TRNRS	P1 CH165J CHILD-TURNERS SYNDROME	P1TRNRS suppressed for respondent confidentiality. These data have been combined with other data. See X1SYNDRM in table 7-6.
1110	P1SPNBFD	P1 CH165K CHILD-SPINA BIFIDA	P1SPNBFD suppressed for respondent confidentiality. These data have been combined with other data. See X1SYNDRM in table 7-6.
1116	P1SRVPS	P1 CH170E PSYCHOLOGICAL SERVICES	These data suppressed for respondent confidentiality.
1120	P1PRGSD	P1 CH180A PARTICIPATE IN SCHOOL DISTRICT	P1PRGSD suppressed for respondent confidentiality. These data have been combined with other data. See X1PRGCMB in table 7-6.
1125	P1PRGOT	P1 CH180F OTHER SOURCE-EARLY INTERVENTN	P1PRGOT suppressed for respondent confidentiality. These data have been combined with other data. See X1PRGCMB in table 7-6.
1126	P1HITWIN	P1 CH190 SAME HEALTH INSURANCE	These data suppressed for respondent confidentiality.

See note at end of table.

Table 7-7. Recoded and suppressed data on the ECLS-B 9-Month Public-Use Data File: 2001–02—Continued

Field ID	Variable	Field label	Comment
1135	P1CHIPKN	P1 205A WHYNOTCHIP-WASN'T KNOWQUALIFY	P1CHIPKN suppressed for respondent confidentiality. These data have been combined with other data. See X1CHPCMB in table 7-6.
1136	P1CHIPJN	P1 205B WHYNOTCHIP-WASN'T KNW HOW JOIN	P1CHIPJN suppressed for respondent confidentiality. These data have been combined with other data. See X1CHPCMB in table 7-6.
1138	P1CHIPBR	P1 205D WHYNOTCHIP-TOO MUCH BOTHER	P1CHIPBR suppressed for respondent confidentiality. These data have been combined with other data. See X1CHPCMB in table 7-6.
1139	P1CHIPSK	P1 205E WHYNOTCHIP-CHILD WASN'T SICK	P1CHIPSK suppressed for respondent confidentiality. These data have been combined with other data. See X1CHPCMB in table 7-6.
1141	P1CHIPOT	P1 CH205G WHYNOTCHIP-OTHER REASON	P1CHIPOT suppressed for respondent confidentiality. These data have been combined with other data. See X1CHPCMB in table 7-6.
1146	P1HIPDCH	P1 CH210E HLTH INS PAID BY-CHIP	These data suppressed for respondent confidentiality.
1148	P1HIPDOT	P1 CH210G HLTH INS PAID BY-OTHER	These data suppressed for respondent confidentiality.
1151	P1NVRDV	P1 CH230A WHY NEVER-DIVORCE	These data suppressed for respondent confidentiality.
1152	P1NVRRE	P1 CH203B WHY NEVER-REMARRIED	These data suppressed for respondent confidentiality.
1154	P1NVRENO	P1 CH230D WHY NEVER-EMPLOYER NOT OFFER	These data suppressed for respondent confidentiality.
1155	P1NVRNE	P1 CH230E WHY NEVER-NOT ELIGIBLE	These data suppressed for respondent confidentiality.
1157	P1NVRRF	P1 CH230G WHY NEVER-INS COMP REF COVERG	These data suppressed for respondent confidentiality.
1158	P1NVRME	P1 CH230H WHY NEVER-LST MEDICAID NEW JOB	These data suppressed for respondent confidentiality.
1160	P1NVRCH	P1 CH230J WHY NEVER-INELIGIBLE FOR CHIP	These data suppressed for respondent confidentiality.
1164	P1PASTDV	P1 CH245A WHY PAST NO COV-DIVORCE	P1PASTDV suppressed for respondent confidentiality. These data have been combined with other data. See X1CVGCMB in table 7-6.
1165	P1PASTRE	P1 CH245B WHY PAST NO COV-REMARRIED	P1PASTRE suppressed for respondent confidentiality. These data have been combined with other data. See X1CVGCMB in table 7-6.
1167	P1PASTNO	P1 CH245D WHY PAST NO COV-EMP NOT OFFER	P1PASTNO suppressed for respondent confidentiality. These data have been combined with other data. See X1CVGCMB in table 7-6.
1169	P1PASTCT	P1 CH245F WHY PAST NO COV-HIGH COST	P1PASTCT suppressed for respondent confidentiality. These data have been combined with other data. See X1CVGCMB in table 7-6.
1170	P1PASTRF	P1 CH245G WHY PAST NO COV-INS COMP REF	P1PASTRF suppressed for respondent confidentiality. These data have been combined with other data. See X1CVGCMB in table 7-6.
1174	P1PASTO	P1 CH245K WHY PAST NO COV-OTH REASON	P1PASTO suppressed for respondent confidentiality. These data have been combined with other data. See X1CVGCMB in table 7-6.
1189	P1NMVS2T	P1 FH042 # VISITS IN 2ND TRIMESTER	These data recoded for respondent confidentiality. Recoded 12 or more to 11.
1190	P1NMVS3T	P1 FH043 # VISITS IN 3RD TRIMESTER	These data recoded for respondent confidentiality. Recoded 14 or more to 13.
1191	P1PLCVST	P1 FH044 PLACE OF PRENATAL VISIT	These data recoded for respondent confidentiality.
1198	P1NOPYMT	P1FH045G PRENATAL CARE-NO PYMT REQUIRED	P1NOPYMT suppressed for respondent confidentiality. These data have been combined with other data. See X1CRCMB in table 7-6.
1199	P1PNCOTH	P1 FH045H PRENATAL CARE PAID-OTHER	P1PNCOTH suppressed for respondent confidentiality. These data have been combined with other data. See X1CRCMB In table 7-6.
1209	P1WEIGHT	P1 FH055 WEIGHT BEFORE PREGNANCY	These data recoded for respondent confidentiality.
1211	P1GNPRG	P1 FH057 GAIN DURING PREGNANCY	These data recoded for respondent confidentiality.

See note at end of table.

Table 7-7. Recoded and suppressed data on the ECLS-B 9-Month Public-Use Data File: 2001–02—Continued

Field ID	Variable	Field label	Comment
1220	P1SMKNUM	P1 FH082 SMOKED 3 MONTHS BEFORE PREG	These data recoded for respondent confidentiality. Recoded 6-9 to 6, 10-14 to 7, and 15 or more to 8.
1222	P1SMK3MN	P1 FH086 SMOKED LAST 3 MTHS OF PREGNANCY	These data recoded for respondent confidentiality.
1225	P1NUMCIG	P1 FH092 # OF CIG SMOKED NOW	These data suppressed for respondent confidentiality.
1227	P1SMKHM	P1 FH097 # PEOPLE SMOKING AT HOME	These data recoded for respondent confidentiality. Recoded 4 or more to 3.
1229	P1NMDKP3	P1 FH101 # ALCOHOLIC DRKS 3 MTHS BEFORE	These data recoded for respondent confidentiality.
1230	P1NMDK3M	P1 FH103 # ALCOHOLIC DRKS LAST 3 MTHS	These data recoded for respondent confidentiality.
1233	P1NMDK1S	P1 FH115 DRINKS AT ONE SETTING	These data recoded for respondent confidentiality. Recoded 6 or more to 5.
1234	P1MARSTS	P1 MH005 CURRENT MARITAL STATUS	These data recoded for respondent confidentiality.
1235	P1NUMMRD	P1 MH010 NUMBER OF TIMES MARRIED	These data recoded for respondent confidentiality. Recoded 4 or more to 3.
1236	P1LVPARM	P1 MH015 MONTH BEGAN LIVING W/ PARTNER	These data suppressed for respondent confidentiality.
1237	P1LVPARY	P1 MH017 YEAR BEGAN LIVING W/ PARTNER	These data suppressed for respondent confidentiality.
1238	P1MTMRRY	P1 MH020 MONTH MARRIED SPOUSE/PARTNER	These data suppressed for respondent confidentiality.
1239	P1YRMRRY	P1 MH022 YEAR MARRIED SPOUSE/PARTNER	These data recoded for respondent confidentiality.
1241	P1MRRYMH	P1 MH030 MONTH MARRIED CHILDS BIOFATHER	These data suppressed for respondent confidentiality.
1242	P1MRRYR	P1MH032 YEAR MARRIED CHILDS BIOFATHER	These data recoded for respondent confidentiality.
1244	P1LIVMTH	P1 MH040 MONTH BEGAN LIVING W/ BIOFATHER	These data suppressed for respondent confidentiality.
1245	P1LIVYR	P1 MH042 YEAR BEGAN LIVING W/ BIOFATHER	These data recoded for respondent confidentiality.
1246	P1MRRYEN	P1 MH045 HOW DID MARRIAGE TO BIOFTHR END	These data suppressed for respondent confidentiality.
1247	P1BFLVNG	P1 MH050 BIOLOGICAL FATHER STILL LIVING	These data suppressed for respondent confidentiality.
1248	P1BFDIEM	P1 MH055 MONTH BIOFATHER PASSED AWAY	These data suppressed for respondent confidentiality.
1249	P1BFDIEY	P1 MH057 YEAR BIOFATHER PASSED AWAY	These data suppressed for respondent confidentiality.
1250	P1STPLVM	P1 MH060 MONTH STOPPED LIVING W/BIOFATHER	These data suppressed for respondent confidentiality.
1251	P1STPLVY	P1 MH062 YEAR STOPPED LIVING W/BIOFATHER	These data recoded for respondent confidentiality.
1252	P1MRENDM	P1 MH065 MONTH MARRIAGE ENDED W/BIOFATHER	These data suppressed for respondent confidentiality.
1253	P1MRENDY	P1 MH066 YEAR MARRIAGE ENDED W/BIOFATHER	These data suppressed for respondent confidentiality.
1255	P1ASKHLP	P1 SS005 ASK FOR HELP IF DEPRESS/CONFUSE	These data recoded for respondent confidentiality.
1256	P1BRWMNY	P1 SS010 ASK TO BORROW MONEY FROM	These data recoded for respondent confidentiality.
1257	P1EMERG	P1 SS015 CALL IN AN EMERGENCY	These data recoded for respondent confidentiality.
1283	P1EDUASS	P1 CS020B RECEIVED EDUCATIONAL ASST	These data suppressed for respondent confidentiality.
1290	P1AGEBM	P1 RI020 AGE OF RSP BIO MOM	These data recoded for respondent confidentiality. Recoded 37 or less to 38 and 84 or more to 83.
1291	P1BMYRS	P1 RI025 NUM YRS RSP BIOMOM PASSED AWAY	These data recoded for respondent confidentiality.
1295	P1AGEBF	P1 RI045 AGE OF RSP BIO DAD	These data recoded for respondent confidentiality. Recoded 37 or less to 38 and 84 or more to 83.
1296	P1BFYRS	P1 RI050 NUM YRS RSP BIODAD PASSED AWAY	These data recoded for respondent confidentiality.
1299	P1MEDU	P1 RI065 RSP MOTHERS HIGHEST ED LEVEL	These data suppressed for respondent confidentiality.
1300	P1FEDU	P1 RI070 RSP FATHERS HIGHEST ED LEVEL	These data suppressed for respondent confidentiality.
1301	P1REDU	P1 RI075 RSP HIGHEST ED LEVEL	These data suppressed for respondent confidentiality.
1319	P1NUMJOB	P1 RI140 NUMBER OF JOBS RSP WORKING	These data recoded for respondent confidentiality. Recoded 4 or more to 3.
1320	P1PAIDHR	P1 RI145 RSP NUM PAID HRS PER WEEK	These data recoded for respondent confidentiality.
1321	P1PAYBTX	P1 RI150 RSP PAY BEFORE TAXES	These data suppressed for respondent confidentiality.
1322	P1PAYUNT	P1 RI152 UNIT OF RSP PAY	These data suppressed for respondent confidentiality.
1342	P1WRK4WK	P1 RI250 # HRS RSP WORK 4 WK PRIOR CH BORN	These data recoded for respondent confidentiality.
1345	P1NMMTWK	P1 RI270 # WKS RSP MATERNITY LEAVE	These data recoded for respondent confidentiality.
1346	P1PDMTLV	P1 RI275 # WKS PAID RSP MATERNITY LEAVE	These data recoded for respondent confidentiality.

See note at end of table.

Table 7-7. Recoded and suppressed data on the ECLS-B 9-Month Public-Use Data File: 2001–02—Continued

Field ID	Variable	Field label	Comment
1357	P1NMLVSP	P1 SI020 # MTHS LIV WITH SP SINCE CB	These data recoded for respondent confidentiality.
1358	P1SPDOBD	P1 SI030 SP DAY OF BIRTH	These data suppressed for respondent confidentiality.
1359	P1SPDOBM	P1 SI031 SP MONTH OF BIRTH	These data suppressed for respondent confidentiality.
1360	P1SPDOBY	P1 SI033 SP YEAR OF BIRTH	These data suppressed for respondent confidentiality.
1361	P1SPEDUC	P1 SI050 SP HIGHEST ED LEVEL	These data recoded for respondent confidentiality.
1369	P1SPNMJB	P1 SI075 NUMBER JOBS SP WORKING	These data recoded for respondent confidentiality. Recoded 4 or more to 3.
1370	P1SPJBHR	P1 SI080 SP NUMBER OF PD HRS PER WEEK	These data recoded for respondent confidentiality.
1371	P1SPINCM	PI SI085 SP PAY BEFORE TAXES	These data suppressed for respondent confidentiality.
1372	P1SPINCU	P1 SI086 UNIT OF SP PAY	These data suppressed for respondent confidentiality.
1388	P1SPLSWK	P1 SI132 WHAT SP DO LAST WK	These data recoded for respondent confidentiality.
1389	P1SPABWK	P1 SI135 SP ABLE TO WORK IF OFFERED JOB	These data suppressed for respondent confidentiality.
1411	P1BFLVMO	P1 BF020 MTH BIOFATHER IN HH AFTER BIRTH	These data recoded for respondent confidentiality. Recoded 14 or more to 13.
1412	P1BFDOBM	P1 BF030 BIOLOGICAL FATHER DOB-MONTH	These data suppressed for respondent confidentiality.
1413	P1BFDOBD	P1 BF031 BIOLOGICAL FATHER DOB-DAY	These data suppressed for respondent confidentiality.
1414	P1BFDOBY	P1 BF032 BIOLOGICAL FATHER DOB-YEAR	These data recoded for respondent confidentiality. Recoded 1966 or earlier to 1967 and 1985 or later to 1984.
1415	P1BFFAGE	P1 BF033 WHAT IS BIOLOGICAL FATHER'S AGE	These data suppressed for respondent confidentiality.
1416	P1BFEDUC	P1 BF050 BIOLOGICAL FATHERS EDUCATION	These data suppressed for respondent confidentiality.
1425	P1BFSEEU	P1 BF180 WHY UNABLE TO SEE CHILD	These data recoded for respondent confidentiality.
1435	P1BFOWE	P1 BF205 HOW MUCH OWED FOR CHILD SUPPORT	These data recoded for respondent confidentiality.
1436	P1BFGIVE	P1 BF210 HOW MUCH GET FOR CHILD SUPPORT	These data recoded for respondent confidentiality.
1451	P1FDSTMO	P1 WP050 HOW MANY MTHS-HH RCVD FOODSTMP	These data recoded for respondent confidentiality. Recoded 14 or more to 13.
1452	P1MEDIMO	P1 WP055 HOW MANY MTHS-HH RCVD MEDICAID	These data recoded for respondent confidentiality. Recoded 14 or more to 13.
1453	P1WELMO	P1 WP060 HOW MANY MTHS-HH RCVD WELFARE	These data recoded for respondent confidentiality. Recoded 14 or more to 13.
1456	P1NMADER	P1 HI005 NUM ADULTS IN HH EARNING MONEY	These data recoded for respondent confidentiality. Recoded 6 or more to 5.
1459	P1HHINCY	P1 HI020 TOTAL HH INCOME LAST YEAR-1000S	These data recoded for respondent confidentiality.
1460	P1HSTYPE	P1 HI025 TYPE OF HOUSING LIVING IN	These data recoded for respondent confidentiality.
1461	P1HSSIT	P1 HI030 HOUSING SITUATION-OWN RENT	These data recoded for respondent confidentiality.
1463	P1HMVALU	P1 HI040 PRESENT HOME VALUE	These data recoded for respondent confidentiality.
1478	P1HOSKML	P1 HF030 HOW OFTEN CUT/SKIP MEALS	These data suppressed for respondent confidentiality.
1485	P1CHSKML	P1 HF065 CHILD SKIP MEAL – LACK OF MONEY	These data suppressed for respondent confidentiality.
1486	P1HOCHSK	P1 HF070 HOW OFTEN DID CHILD SKIP MEALS	These data suppressed for respondent confidentiality.
1488	P1CHSKDY	P1 HF080 CHILD SKIP FOOD WHOLE DAY-MONEY	These data suppressed for respondent confidentiality.
1523	P1NUMCH	P1 SAQ7 NUMBER OF BIOLOGICAL CHILDREN	These data recoded for respondent confidentiality. Recoded 8 or more to 7.
1524	P1AGECH1	P1 SAQ8 AGE WHEN FIRST CHILD WAS BORN	These data recoded for respondent confidentiality. Recoded 14 or less to 15 and 41 or more to 40.
1526	P1NUMCHO	P1 SAQ10 NUM BIO CHLDRN OUTSIDE HOUSEHLD	These data recoded for respondent confidentiality. Recoded 5 or more to 4.
1543	P1WNTNCH	P1 SAQ19 NUMBER OF CHILDREN WOULD CHOOSE	These data recoded for respondent confidentiality. Recoded 12 or more to 11.
1556	BCDOBMM	BC MONTH OF BIRTH	These data suppressed for respondent confidentiality.
1561	BCSTOCFP	BC STATE OF OCCURRENCE (FIPS)	These data suppressed for respondent confidentiality.
1562	BCCNOCFP	BC COUNTY OF OCCURRENCE (FIPS)	These data suppressed for respondent confidentiality.
1564	BCSTRSFP	BC STATE OF RESIDENCE (FIPS)	These data suppressed for respondent confidentiality.

See note at end of table.

Table 7-7. Recoded and suppressed data on the ECLS-B 9-Month Public-Use Data File: 2001–02—Continued

Field ID	Variable	Field label	Comment
1565	BCCNRSFP	BC COUNTY OF RESIDENCE (FIPS)	These data suppressed for respondent confidentiality.
1568	BCMOMHSP	BC HISPANIC ORIGIN OF MOTHER	These data suppressed for respondent confidentiality.
1569	BCMOMRCI	BC RACE OF MOTHER IMPUTATION FLAG	These data suppressed for respondent confidentiality.
1570	BCMOMRC	BC RACE OF MOTHER	These data suppressed for respondent confidentiality.
1575	BCADQCAR	BC CARE ADEQUACY RECODE-KESSNER INDEX	These data suppressed for respondent confidentiality.
1576	BCNLBNL	BC NUMBER OF LIVE BIRTHS, NOW LIVING	These data suppressed for respondent confidentiality.
1577	BCNLBND	BC NUMBER OF LIVE BIRTHS, NOW DEAD	These data suppressed for respondent confidentiality.
1578	BCNOTERM	BC NUMBER OF OTHER TERMINATIONS	These data suppressed for respondent confidentiality.
1581	BCLMPMM	BC MONTH LAST NORMAL MENSES BEGAN	These data suppressed for respondent confidentiality.
1582	BCLMPDD	BC DAY LAST NORMAL MENSES BEGAN	These data suppressed for respondent confidentiality.
1583	BCLMPYY	BC YEAR LAST NORMAL MENSES BEGAN	These data suppressed for respondent confidentiality.
1584	BCFTHAGE	BC AGE OF FATHER	These data suppressed for respondent confidentiality.
1585	BCFTHHSP	BC FATHER OF HISPANIC ORIGIN	These data suppressed for respondent confidentiality.
1586	BCFTHRC	BC RACE OF FATHER	These data suppressed for respondent confidentiality.
1589	BCGESTWK	BC GESTATION-DETAIL IN WEEKS	These data suppressed for respondent confidentiality.
1591	BCSEX	BC SEX	These data suppressed for respondent confidentiality.
1592	BCPLURI	BC PLURALITY IMPUTATION FLAG	These data suppressed for respondent confidentiality.
1593	BCPLURAL	BC PLURALITY	These data suppressed for respondent confidentiality.
1602	BCCARDIC	BC MED RISK FACTOR-CARDIAC DISEASE	These data suppressed for respondent confidentiality.
1605	BCHERPES	BC MED RISK FACTOR-GENITAL HERPES	These data suppressed for respondent confidentiality.
1607	BCHEMO	BC MED RISK FACTOR-HEMOGLOBINOPATHY	These data suppressed for respondent confidentiality.
1610	BCECLAMP	BC MED RISK FACTOR-ECLAMPSIA	These data suppressed for respondent confidentiality.
1611	BCINCRVX	BC MED RISK FACTOR-INCOMPETENT CERVIX	These data suppressed for respondent confidentiality.
1612	BCPRV4KG	BC MED RISK FACTOR-PREVIOUS 4000+ GRAM	These data suppressed for respondent confidentiality.
1613	BCPRVPTM	BC MED RISK FACTOR-PREV PRETERM OR SMALL	These data suppressed for respondent confidentiality.
1614	BCRENAL	BC MED RISK FACTOR-RENAL DISEASE	These data suppressed for respondent confidentiality.
1615	BCRH_SNS	BC MED RISK FACTOR-RH SENSITIZATION	These data suppressed for respondent confidentiality.
1616	BCUTERIN	BC MED RISK FACTOR-UTERINE BLEEDING	These data suppressed for respondent confidentiality.
1617	BCOTHMR	BC MED RISK FACTOR-OTHER	These data suppressed for respondent confidentiality.
1618	BCTOBACC	BC TOBACCO USE DURING PRENANCY	These data suppressed for respondent confidentiality.
1619	BCNUMCIG	BC AVG NUMBER CIGS PER DAY DURING PREG	These data suppressed for respondent confidentiality.
1620	BCALCOHL	BC ALCOHOL USE DURING PREGNANCY	These data suppressed for respondent confidentiality.
1621	BCNUMDRK	BC NUMBER OF DRKNS PER WEEK DURING PREG	These data suppressed for respondent confidentiality.
1626	BCSTIMUL	BC OBSTETRIC PROC-LABOR STIMULATED	These data suppressed for respondent confidentiality.
1630	BCFEBRL	BC LABOR COMPLC-FEBRILE(>100 DEGREE F)	These data suppressed for respondent confidentiality.
1633	BCABRPT	BC LABOR COMPL-ABRUPTIO PLACENTA	These data suppressed for respondent confidentiality.
1634	BCPREVIA	BC LABOR COMPLC-PLACENTA PREVIA	These data suppressed for respondent confidentiality.
1635	BCEXCBLD	BC LABOR COMPLC-OTHER EXCESSIVE BLEEDING	These data suppressed for respondent confidentiality.
1636	BCSEIZR	BC LABOR COMPLC-SEIZURES DURING LABOR	These data suppressed for respondent confidentiality.
1637	BCPRECIP	BC LABOR COMPLC-PRECIP LBR (<3 HOURS)	These data suppressed for respondent confidentiality.
1638	BCPROLNG	BC LABOR COMPLC-PROLONGED LBR > 20 HOURS	These data suppressed for respondent confidentiality.
1639	BCDYSFNC	BC LABOR COMPLC-DYSFUNCTIONAL LABOR	These data suppressed for respondent confidentiality.
1641	BCCEPHLO	BC LABOR COMPLC-CEPHALOPELVIC DISPRPRTN	These data suppressed for respondent confidentiality.
1642	BCCORD	BC LABOR COMPLC-CORD PROLAPSE	These data suppressed for respondent confidentiality.
1643	BCANESTH	BC LABOR COMPLC-ANESTHETIC COMPLIC	These data suppressed for respondent confidentiality.
1645	BCOTHLBR	BC LABOR COMPLC-OTHER	These data suppressed for respondent confidentiality.
1646	BCCHANEM	BC NEWBORN COND-ANEMIA HCT<39/HGB<13	These data suppressed for respondent confidentiality.
1647	BCINJURY	BC NEWBORN COND-BIRTH INJURY	These data suppressed for respondent confidentiality.

See note at end of table.

Table 7-7. Recoded and suppressed data on the ECLS-B 9-Month Public-Use Data File: 2001–02—Continued

Field ID	Variable	Field label	Comment
1648	BCCHFAS	BC NEWBORN COND-FETAL ALCOHOL SYNDROME	These data suppressed for respondent confidentiality.
1649	BCHYALIN	BC NEWBORN COND-HYALINE MEMBRANE DISEASE	These data suppressed for respondent confidentiality.
1650	BCMECSYN	BC NEWBORN COND-MECONIUM ASPIRATION SYN	These data suppressed for respondent confidentiality.
1653	BCCHSEIZ	BC NEWBORN COND-SEIZURES	These data suppressed for respondent confidentiality.
1654	BCOTHNC	BC NEWBORN COND-OTHER SPECIFY	These data suppressed for respondent confidentiality.
1655	BCANECPH	BC CONGEN ANOMALY-ANENCEPHALUS	These data suppressed for respondent confidentiality.
1656	BCSPINA	BC CONGEN ANOMALY-SPINA BIFIDA	These data suppressed for respondent confidentiality.
1657	BCHYDCPH	BC CONGEN ANOMALY-HYDROCEPHALUS	These data suppressed for respondent confidentiality.
1658	BCMCRCPH	BC CONGEN ANOMALY-MICROCEPHALUS	These data suppressed for respondent confidentiality.
1659	BCNRVS	BC CONGEN ANOMALY-OTHER CENTRL NERV.	These data suppressed for respondent confidentiality.
1660	BCHEART	BC CONGEN ANOMALY-HEART MALFORMATIONS	These data suppressed for respondent confidentiality.
1661	BCCIRCUL	BC CONGEN ANOMALY-OTHER CIRCL/RESP	These data suppressed for respondent confidentiality.
1662	BCRECTAL	BC CONGEN ANOMALY-RECTAL ATRESIA/STNOSIS	These data suppressed for respondent confidentiality.
1663	BCTRACHE	BC CONGEN ANOMALY-TRACHEO/ESOPHAGEAL	These data suppressed for respondent confidentiality.
1664	BCOMPHL	BC CONGEN ANOMAL-OMPHALOCELE-GASTROSCH	These data suppressed for respondent confidentiality.
1665	BCGASTRO	BC CONGEN ANOMALY-OTHER GASTROINTEST	These data suppressed for respondent confidentiality.
1666	BCGENITL	BC CONGEN ANOMAL-MALFORMED GENITALIA	These data suppressed for respondent confidentiality.
1667	BCRNLAGE	BC CONGEN ANOMALY-RENAL AGENESIS	These data suppressed for respondent confidentiality.
1668	BCUROGEN	BC CONGEN ANOMALY-OTHER UROGENITAL	These data suppressed for respondent confidentiality.
1669	BCCLFTLP	BC CONGEN ANOMALY-CLEFT LIP/PALATE	These data suppressed for respondent confidentiality.
1670	BCACTYLY	BC CONGEN ANOMALY-(POLY/SYN/A)DACTYLY	These data suppressed for respondent confidentiality.
1671	BCCLUBFT	BC CONGEN ANOMALY-CLUB FOOT	These data suppressed for respondent confidentiality.
1672	BCHERNIA	BC CONGEN ANOMALY-DIAPHRAGMATIC HERNIA	These data suppressed for respondent confidentiality.
1673	BCMUSCLO	BC CONGEN ANOMAL-OTHER MUSCULOSKELETAL	These data suppressed for respondent confidentiality.
1674	BCDOWNS	BC CONGEN ANOMALY-DOWNS SYNDROME	These data suppressed for respondent confidentiality.
1675	BCCHROM	BC CONGEN ANOMALY-OTHER CHROMOSOMAL	These data suppressed for respondent confidentiality.
1677	F1RELCH	F1 Q1 RELATIONSHIP TO THE CHILD	These data recoded for respondent confidentiality.
1682	F1CGDIAP	F1 Q3A CHANGE CHILDS DIAPER	F1CGDIAP suppressed for respondent confidentiality. These data have been combined with other data. See X1CGDCMBin table 7-6.
1757	F1NUMMRY	F1 Q27 HOW MANY TIMES MARRIED	These data recoded for respondent confidentiality. Recoded 4 or more to 3.
1758	F1NUMCH	F1 Q28 NUMBER OF BIOLOGICAL CHILDREN	These data recoded for respondent confidentiality. Recoded 8 or more to 7.
1759	F1AGECH1	F1 Q29 AGE WHEN FIRST CHILD WAS BORN	These data recoded for respondent confidentiality. Recoded 14 or less to 15 and 41 or more to 40.
1761	F1NUMCHO	F1 Q31 NUM BIO CHLDRN OUTSIDE HOUSEHLD	These data recoded for respondent confidentiality. Recoded 5 or more to 4.
1763	F1FDOBMM	F1 Q33 YOUR BIRTH DAY (MONTH)	These data suppressed for respondent confidentiality.
1764	F1FDOBDD	F1 Q33 YOUR BIRTH DAY (DAY)	These data suppressed for respondent confidentiality.
1765	F1FDOBYY	F1 Q33 YOUR BIRTH DAY (YEAR)	These data suppressed for respondent confidentiality.
1766	F1CTRYBN	F1 Q34 IN WHAT COUNTRY WERE YOU BORN	These data recoded for respondent confidentiality.
1767	F1AGEUS	F1 Q35 AGE WHEN MOVED TO US	These data recoded for respondent confidentiality.
1774	F1HIGHGR	F1 Q39 HIGHEST GRADE OF SCHOOL COMPLETED	These data suppressed for respondent confidentiality.
1789	F1NUMJBS	F1 Q48 HOW MANY JOBS DO YOU HAVE NOW	These data recoded for respondent confidentiality. Recoded 4 or more to 3.
1790	F1HRSWK	F1 Q49 TOTAL HOURS PER WEEK WORK FOR PAY	These data recoded for respondent confidentiality.
1791	F1SALARY	F1 Q50 TOTAL SALARY BEFORE TAXES	These data suppressed for respondent confidentiality.
1792	F1SRYUNT	F1 Q50 TOTAL SALARY – PAY UNIT	These data suppressed for respondent confidentiality.
1807	F1LASTWK	F1 Q56 DOING MOST OF LAST WEEK	These data recoded for respondent confidentiality.

See note at end of table.

Table 7-7. Recoded and suppressed data on the ECLS-B 9-Month Public-Use Data File: 2001–02—Continued

Field ID	Variable	Field label	Comment
1848	F1LVMAGE	F1 Q76 AGE STOPPED LIVING WITH BIOMTH	These data recoded for respondent confidentiality.
1851	F1LVFAGE	F1 Q79 AGE STOPPED LIVING WITH BIOFTH	These data recoded for respondent confidentiality.
1905	N1CHSPMO	N1 Q21 HOW MUCH CH SUP SHOULD PAY/MONTH	These data recoded for respondent confidentiality.
1906	N1SPLTMO	N1 Q22 HOW MUCH CH SUPP PAID LAST MONTH	These data recoded for respondent confidentiality.
1907	N1FDOBMM	N1 Q23 YOUR BIRTH DATE (MONTH)	These data suppressed for respondent confidentiality.
1908	N1FDOBDD	N1 Q23 YOUR BIRTH DATE (DAY)	These data suppressed for respondent confidentiality.
1909	N1FDOBYY	N1 Q23 YOUR BIRTH DATE (YEAR)	These data recoded for respondent confidentiality.
1910	N1CTRYBN	N1 Q24 IN WHAT COUNTRY WERE YOU BORN	These data recoded for respondent confidentiality.
1911	N1USCTZN	N1 Q25 CITIZEN OF THE UNITED STATES	These data suppressed for respondent confidentiality.
1912	N1HIGHGR	N1 Q26 HIGHEST GRADE OF SCHOOL COMPLETED	These data suppressed for respondent confidentiality.
1925	N1FNWKOT	N1 Q33 SOMETHING ELSE	These data suppressed for respondent confidentiality.
1940	N1OTHPPL	N1 Q37 # PEOPLE LIVING W/ YOU LAST MONTH	These data recoded for respondent confidentiality. Recoded 7 or more to 6.
1941	N1HHINCM	N1 Q38 TOTAL HOUSEHOLD INCOME	These data suppressed for respondent confidentiality.
1942	N1HHINSP	N1 Q39 TOTAL HOUSEHOLD INCOME, SPECIFIC	These data suppressed for respondent confidentiality.
1950	R1PRNLNG	R1 IR120 LANGUAGE USED FOR PRNT INTRVW	These data recoded for respondent confidentiality.
1986	BCBRTHWT	C1 CHILD BIRTH WEIGHT (GRAMS-BIRTH CERT)	These data recoded for respondent confidentiality.

SOURCE: U.S. Department of Education, National Center for Education Statistics, Early Childhood Longitudinal Study, Nine-Month Data Collection, 2001–02.

REFERENCES

Duncan, O. D. (1961). Properties and Characteristics of the Socioeconomic Index. In A. J. Reiss, Jr. (Ed.), *Occupations and Social Status* (pp. 139-161). New York: Free Press of Glencoe.

Executive Office of the President, Office of Management and Budget. (2000). *Standard Occupational Classification Manual.* Springfield, VA: National Technical Information Service (NTIS).

Fischer, G. H., and Molenaar, I. W. eds. (1995). *Rasch Models: Foundations, Recent Developments, and Applications.* New York: Springer-Verlag.

Hamilton, W.L., Cook, J.T., Thompson, W.W., Buron, L.F., Frongillo, E.A., Jr., Olson, C.M., and Wehler. C.A. (1997a). *Household Food Security in the United States in 1995: Executive Summary*. Alexandria, VA: U.S. Department of Agriculture, Food and Consumer Service. (Available: http://www.fns.usda.gov/oane/MENU/Published/FoodSecurity/SUMMARY.PDF)

Hamilton, W.L., Cook, J.T., Thompson, W.W., Buron, L.F., Frongillo, E.A., Jr., Olson, C.M., and Wehler, C.A. (1997b). *Household Food Security in the United States in 1995: Technical Report of the Food Security Measurement Project*. Alexandria, VA: U.S. Department of Agriculture, Food and Consumer Service. (Available: http://www.fns.usda.gov/oane/MENU/Published/FoodSecurity/TECH_RPT.PDF)

Little, R.J.A., and Rubin, D.B. (2002). Statistical Analysis with Missing Data, New York; Chichester: John Wiley & Sons.

Nord, M., and Bickel, G. (2001). Estimating the Prevalence of Children's Hunger from the Current Population Survey Food Security Supplement. In M. Andrews and M. Prell (Eds.), *Second Food Security Measurement and Research Conference, Volume II: Papers—Food Assistance and Nutrition Research Report 11-2* (pp. 31-49). Washington, DC: Economic Research Service, U.S. Department of Agriculture.

Radimer, K.L. (1990). *Understanding Hunger and Developing Indicators to Assess It.* Ph.D. dissertation, Cornell University, Ithaca, New York.

Radimer, K.L., Olson, C.M. and Campbell, C.C. (1990). Development of Indicators to Assess Hunger. *Journal of Nutrition* 120 (Supplement 11): 1544-1548.

Radimer, K.L., Olson, C.M., Greene, J.C., Campbell, C.C., and Habicht, J.P. (1992). Understanding Hunger and Developing Indicators to Assess It in Women and Children. *Journal of Nutrition Education* 24 (1): 36S-44S.

U.S. Department of Agriculture, Food and Nutrition Service. (2000). *Guide to Measuring Household Food Security, Revised 2000*. Alexandria, VA.: U.S. Department of Agriculture.

U.S. Department of Agriculture, Economic Research Service. (2002). *Measuring Children's Food Security in U.S. Households, 1995-99.* Washington, DC: U.S. Department of Agriculture.

U.S. Department of Commerce, U.S. Census Bureau. (2002). *Current Population Survey,* "Poverty 2001." Washington, DC: U.S. Department of Commerce. (Available: http://www.census.gov/hhes/poverty/threshld/thresh01.html)

U.S. Department of Education, National Center for Education Statistics. (Forthcoming). *Early Childhood Longitudinal Study, Birth Cohort (ECLS-B), Methodology Report for the Nine-Month Data Collection, Volume 1: Psychometric Characteristics.* Washington, D.C.: U.S. Department of Education.

U.S. Department of Health and Human Services, National Center for Health Statistics (2002). *Births, Final Data, 2001*. Hyattsville, MD: U.S. Department of Health and Human Services. (Available: http://www.cdc.gov/nchs/data/nvsr/nvsr51/nvsr51_02.pdf)

U.S. Department of Health and Human Services, National Center for Health Statistics (Revised February 2003). *Vital Statistics of the United States, 2001 Natality Technical Appendix*. Hyattsville, MD: U.S. Department of Health and Human Services. (Available: http://www.cdc.gov/nchs/data/techap01.pdf)

Wehler, C.A., Scott, R.I., and Anderson, J.J. (1992). The Community Childhood Hunger Identification Project: A Model of Domestic Hunger—Demonstration Project in Seattle, Washington. *Journal of Nutrition Education* 24 (1): 29S-35S.

Wright, B.D. (1977). *Solving Measurement Problems with the Rasch Model.* Mesa Psychometric Laboratory, The University of Chicago, College of Education, Chicago, IL. (Available: http://www.rasch.org/memos.htm)

Wright, B.D. (1983). *Fundamental Measurement in Social Science and Education.* Mesa Psychometric Laboratory, The University of Chicago, College of Education, Chicago, IL. (Available: http://www.rasch.org/memos.htm)

8. ELECTRONIC CODE BOOK

8.1 Introduction

The purpose of this chapter is to provide users of the Early Childhood Longitudinal Study, both Kindergarten Class of 1998–99 (ECLS-K) and Birth Cohort (ECLS-B), with specific directions for using the Electronic Code Book (ECB) CD-ROM. The information in this chapter provides a comprehensive tour through the ECB that addresses all of the functions and capabilities of the program. These functions allow users to access the accompanying catalog and “view” the data in various ways by performing customized searches, queries, and extractions. The organization of this document provides a “start to finish” approach through the system, beginning with the installation of the ECB, utilizing the ECB’s functions, navigating through the catalog, and performing user-specified data extractions.

Sections 8.1 through 8.6 contain general instructions on using the ECB and apply to both the ECLS-K ECB and the ECLS-B ECB, including descriptions of the menu bars (exhibit 8-57). The exhibits and examples given in these sections are generic and will not match exactly what the users see on their own screens. The last section, 8.7, is file-specific and reflects the actual data on the file.

The ECB CD-ROM contains an ECB that allows users to easily examine the variables in the ECB data set. The data user can create SAS, SPSS for Windows, and Stata programs that will generate an extract data file from the text (ASCII) data file on the CD-ROM. For more information about the data file, see section 8.7.

Additionally, the CD-ROM contains Portable Document Format (PDF) files of the associated questionnaires in appendix A and the record layout for the data file in appendix B. When needed, additional user’s guides and additional supplementary files may also be included in additional appendices.

8.1.1 Hardware/Software Requirements

The ECB program is designed to run under Windows 95®, Windows 98®, Windows 2000®, Windows XP®, or Windows NT® 4.0 on a Pentium-class or higher PC. The PC should also have a minimum of 20 megabytes (MB) of available disk space. The program will visually fit best on screens set to a desktop area of 800 x 600 pixels. It will still work on other screen settings, but it may not make the best use of the available screen space. You can check/set your desktop area as follows:

1. Click on the Windows Start button.
2. Select the Settings menu and then the Control Panel folder icon.
3. In the Control Panel window, click on the Display icon.
4. Select the Settings tab.
5. Set the Desktop Area to 800 x 600 pixels with the Desktop Area slidebar.

The ECB requires approximately 20 MB of available disk space on your hard drive. If 20 MB of space is not available, you may wish to delete unnecessary files from the drive to make space for the ECB.

8.1.2 ECB Features

The ECB allows a user to do the following:

- Search the names and labels of variables in the database (called the catalog) to select variables for analysis (see section 8.3, Variable List).
- Examine the question wording, response categories, and response frequencies for variables the user selects (see section 8.4.9, Viewing Code Book and Variable Information).
- Create a list of variables to be extracted from the catalog, save the list for later use, print the list as a code book, or use a predefined list on the ECB (see section 8.4, Working Taglist).

- Automatically generate SAS, SPSS for Windows, or Stata programs to extract selected variables from the whole data set or for a subset of the cases that are defined by the user (see section 8.5, Extracting Data from the ECB).

The ECB does not create a SAS, SPSS for Windows, or Stata data file. It will prepare the statements that you can use with your own SAS, SPSS for Windows, or Stata software to create your file. As noted earlier, the CD-ROM contains an ASCII data set that the ECB uses to extract specific subdata files. The CD-ROM must be in the drive for the data to be extracted.

8.2 Installing, Starting, and Exiting the ECB

The ECB is provided on a CD-ROM and is intended to be installed and run from within the Windows 95 (or any later version of Windows) environment. The sections in this chapter provide you with step-by-step instructions for installing the program on your personal computer (PC), starting the program, and exiting the program once you have completed your tasks.

If you installed the ECB on your PC before, you have to uninstall the old program before installing this version. For details on how to remove the ECB program from your PC, please see section 8.2.4.

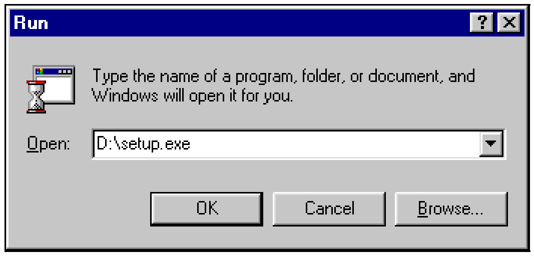
8.2.1 Installing the ECB Program on Your Personal Computer

Program installation is initiated by running the Setup.exe file found within the CD-ROM's root directory.

How To Install the Program:

1. Close all applications on your computer.
2. Insert the installation CD-ROM into your PC's CD-ROM drive.
3. From the desktop Start menu, select Run.
4. Type "D:\Setup.exe" into the "Open" field of the Run screen, shown in exhibit 8-1. If your CD-ROM drive is assigned a different drive letter, substitute it for the "D."

Exhibit 8-1. Windows Run screen



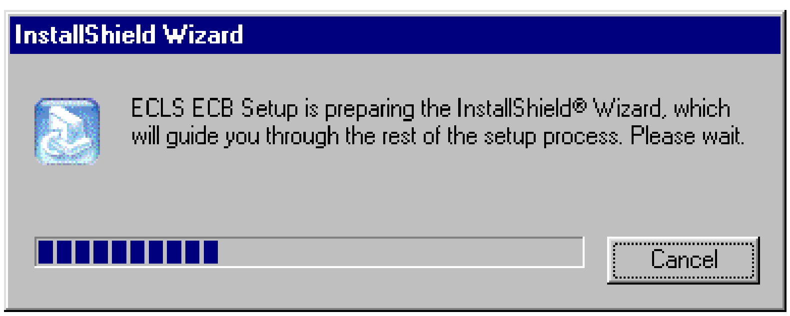
5. Click on the OK button to start the installation. You will now see several installation screens, some of which will prompt you for a response.

   Depending on your PC's configuration, you may encounter warning messages during installation. To respond, always keep the newer version of a file being copied and ignore any access violations that occur during file copying.

   If you are installing multiple ECBs (not different versions of the same ECB) on your PC, you may receive a message warning that setup is about to replace pre-existing files. To respond, always opt to continue the installation although the default is to cancel the setup. When you get a followup message to confirm whether the installation should be continued, press "Yes" to continue although the default is "No."

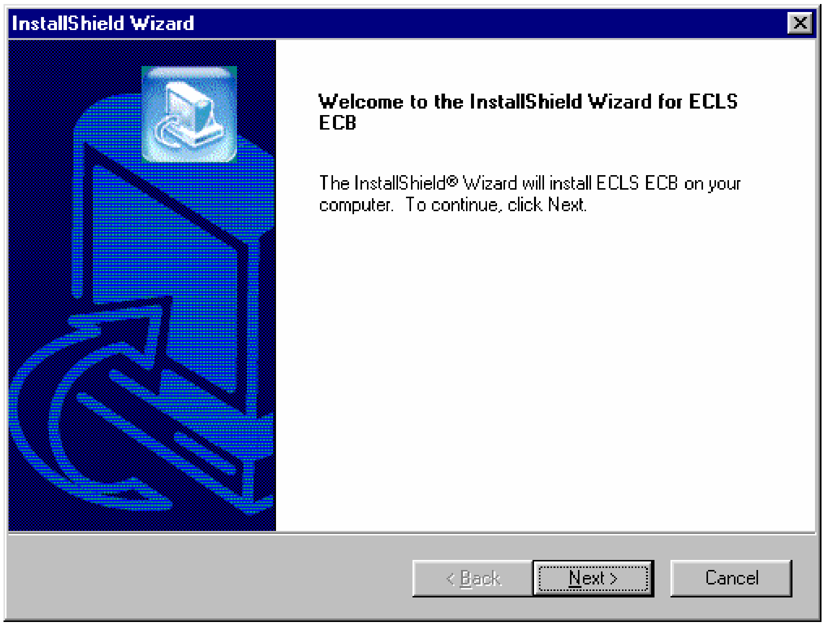
6. The screen shown in exhibit 8-2 indicates that the setup is being prepared.

Exhibit 8-2. InstallShield Wizard



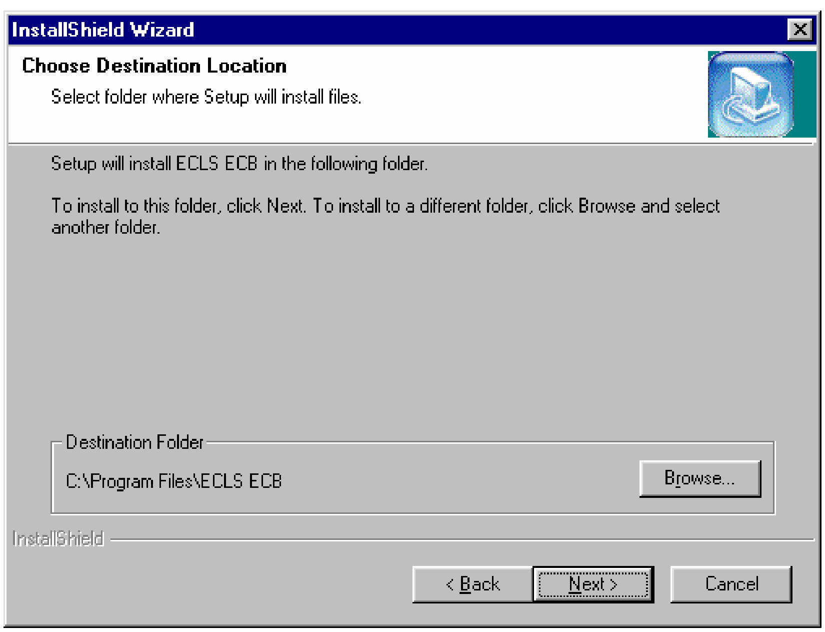
7. You will be prompted to continue with the installation in the Welcome window shown in exhibit 8-3. Click on the Next button to continue.

Exhibit 8-3. Welcome window



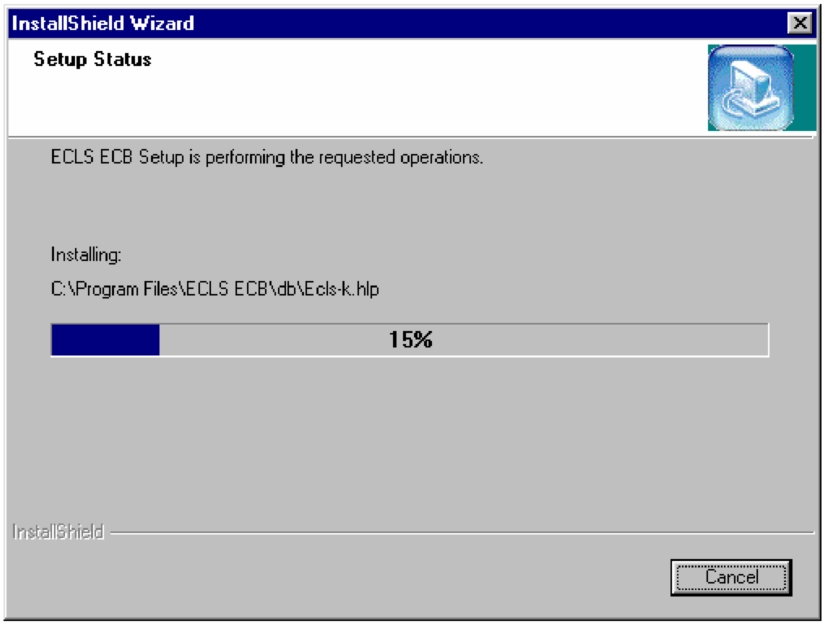
8. When you continue, you will be prompted to choose a destination location for the installation in the window shown in exhibit 8-4. If you wish to change the destination location, click on the Browse button to change the directory. Click on the Next button when the desirable destination folder is shown.

Exhibit 8-4. Choose Destination Location



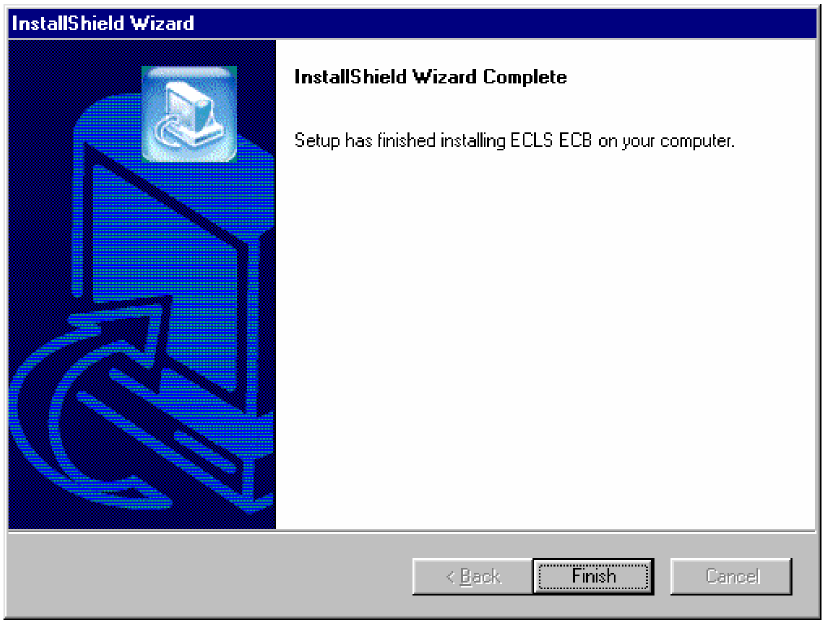
9. Setup will then start installing files. Exhibit 8-5 shows the setup status.

Exhibit 8-5. Setup Status



10. Once the installation is completed, the InstallShield Wizard Complete window shown in exhibit 8-6 will appear. Click on the Finish button to finish the process and return to your PC's desktop.

Exhibit 8-6. InstallShield Wizard Complete



11. The installation process should take about a minute, depending on the speed of the computer on which the ECB is being installed.

Another option for installing the ECB software is to go to the Start menu and go to Settings. Select Control Panel and select Add/Remove Programs from the options. Click on the Install button and follow the directions. Make sure the ECB CD-ROM is in the CD-ROM drive before starting. The program will automatically find the file Setup.exe in the CD-ROM and begin installation. The process will begin at point 5 in the section above.

8.2.2 Starting the ECB

Now that you have installed the ECB on your PC, you can start the program by simply selecting it from the Windows Start, Programs Menu, ECB.

How to Start the ECB:

1. On the desktop screen (exhibit 8-7b), click on the Start button and then point to Programs. Click on the ECB title to invoke the program. Alternatively, click on the ECB desktop icon (exhibit 8-7a) shown below to invoke the program.

Exhibit 8-7a. Desktop icon

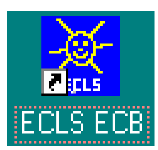
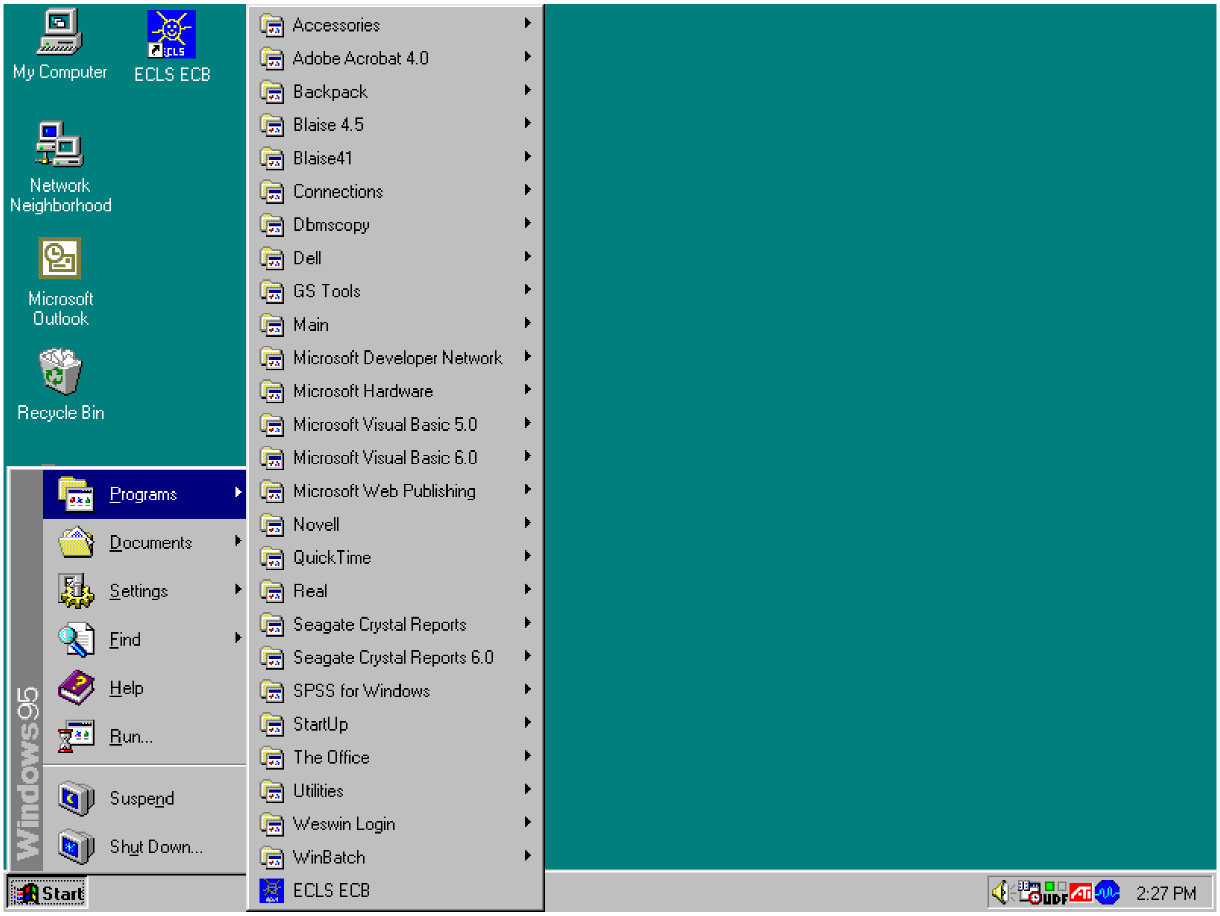


Exhibit 8-7b. Desktop screen—click start

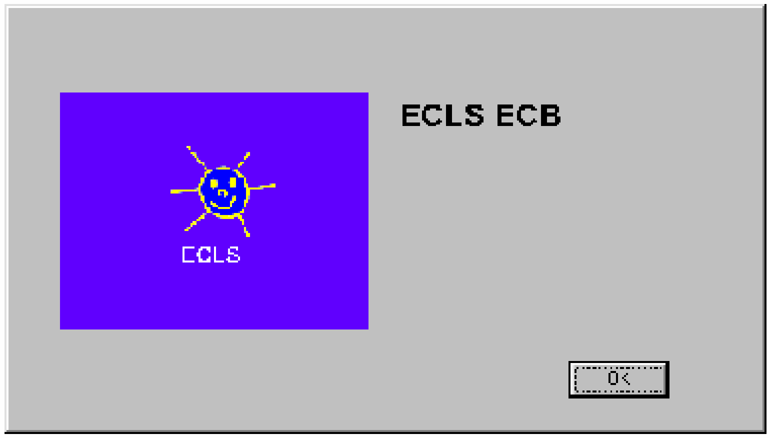


2. If you are a first-time user of the ECB, exhibit 8-8 will display and ask if you are a new ECB user.

Exhibit 8-8. First-time user dialog box

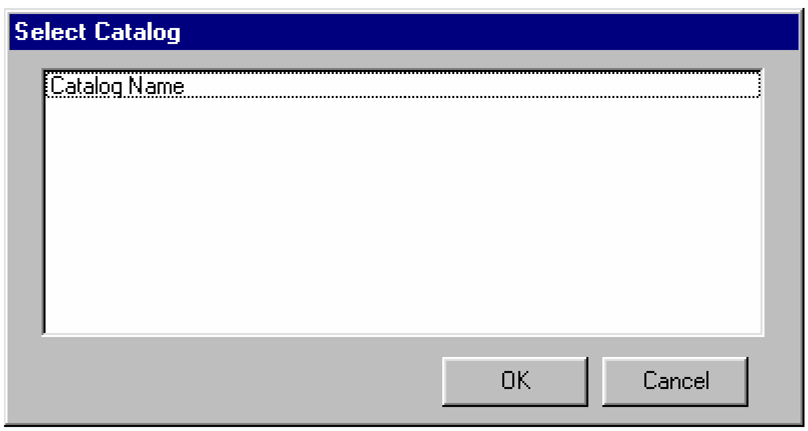
3. Click "Yes" if you are a first-time user. The ECB splash-screen shown in exhibit 8-9 will display.

Exhibit 8-9. ECB splash screen



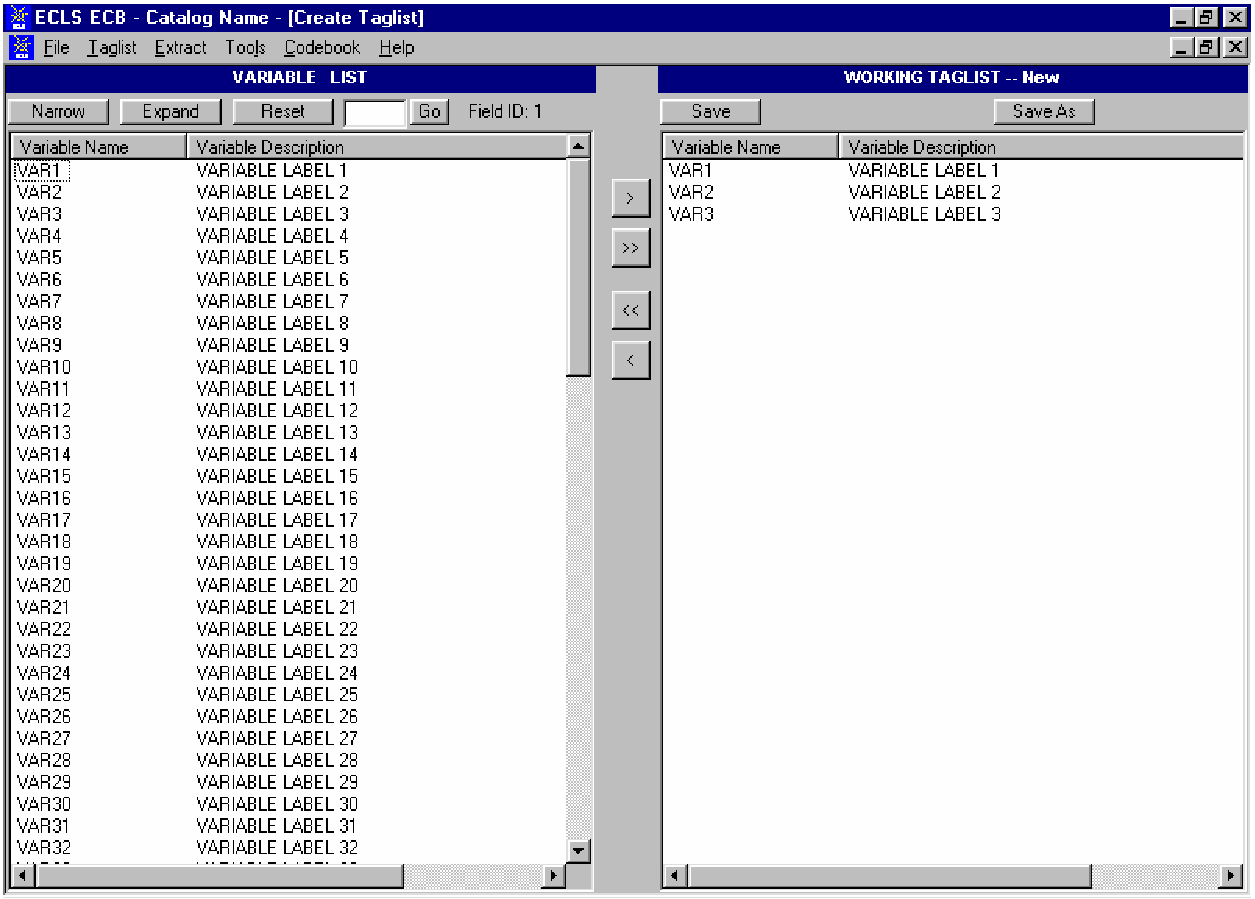
4. On the Select Catalog screen (exhibit 8-10), highlight the name of the catalog. (The ECB has only one catalog.)

Exhibit 8-10. Select Catalog screen



5. Click OK to open the Main ECB screen, shown in exhibit 8-11.

Exhibit 8-11. Main ECB screen



6. You are now ready to use the functions of the ECB as described in the following sections.

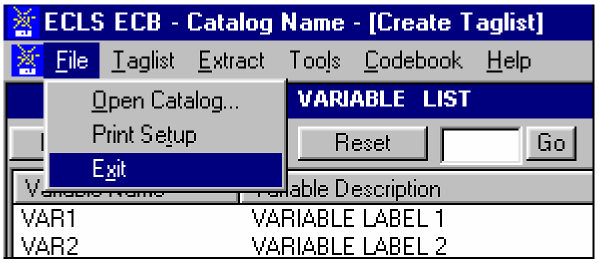
8.2.3 Exiting the ECB

The ECB can be shut down at any time; however, you will be prompted to save any unsaved information.

How To Shut Down the ECB:

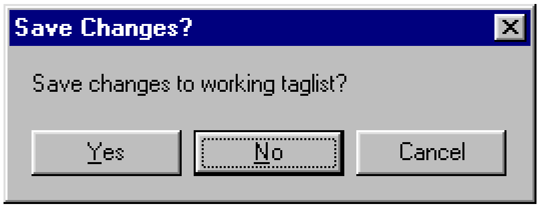
1. From the File menu, click on the Exit option as shown in exhibit 8-12.

Exhibit 8-12. Exit screen



2. If you have not saved your Working Taglist, you will be prompted with the dialog box shown in exhibit 8-13.

Exhibit 8-13. Save working taglist dialog box



3. If you DO NOT wish to save your Working Taglist, click on the "No" button. If you DO wish to save your Working Taglist, click the "Yes" button. For more information, refer to section 8.4.4, Saving Taglists.

8.2.4 Removing the ECB Program From Your Personal Computer

How to Uninstall the ECB:

1. Click on the Windows Start button.

2. Select the Settings menu.

3. In the Control Panel window, click on the Add/Remove Programs.

4. Select "ECB" and click on the Add/Remove button.

5. Follow any prompts. You will be prompted by the InstallShield Wizard to confirm the uninstallation and finish the process.

6. The program is designed so that the uninstallation will keep the taglists when the ECB program is uninstalled in order that all the saved taglists will be retained when the ECB is reinstalled. As a result, the uninstallation will not remove the directory where the ECB was located.

8.2.5 Title Bar

The Title Bar, shown below in exhibit 8-14, is the horizontal bar located at the top of the main screen. It will list the name of the program and the catalog that you have opened, and it will indicate that you are in the "Create Taglist" mode.

Exhibit 8-14. Title Bar

8.2.6 Menu Bar

Selecting items from the pulldown menus listed on the Menu Bar (exhibit 8-15) provides access to the available action commands. Section 8.6 shows the choices and functions available within each menu.

Exhibit 8-15. Menu Bar

How to Access the Menu Bar Items:

1. Point to an item on the Menu Bar and click.

2. Click on a command from the dropdown list.

The Menu Bar may also be activated and its options selected using the shortcut keys described in section 8.2.7.

8.2.7 Using Shortcut Keys to Navigate

The shortcut keys provide a means for selecting menu options and screen buttons without the use of a mouse. These shortcut keys are identified by an <u>underscore</u> under the shortcut letter within the option or button label. The menus that appear on the windows are activated by simultaneously selecting the <ALT> key and the <u>underscored</u> letter. An example of this is the activation of the Taglist Menu by selecting the key combination of <ALT>**-**<T>. Once the menu is activated and all options are displayed, the options can be selected by then pressing the underscored letter for the desired option or by pressing the arrow keys to move between the options.

Not all screens have shortcut keys. They may, however, be used without mouse capability by pressing the <TAB> key. The <TAB> key moves the cursor or highlight through the options and buttons within the windows. When the desired option or button is highlighted, it can be selected by pressing the <ENTER> key.

8.3 Variable List

The ECB main screen, shown in exhibit 8-16, comprises two primary lists that each provide functions for reviewing, grouping, and extracting variable data from the opened catalog. These lists include the Variable List and the Working Taglist.

The Variable List, shown in exhibit 8-17, is a list of all variables associated with the current catalog. When you first open a catalog, all variables contained in the catalog are displayed in the Variable List. Once the catalog is open and the Variable List is displayed, you can scroll through the list using the scrollbar controls at the right side of the Variable List screen. Additionally, you can press <PgUp> and <PgDn> to scroll the list one screen at a time. <Ctrl><Home> and <Ctrl><End> will move to the first and last variable in the list, respectively. Also, the arrow keys can be used to move through the list of variable names.

The “Field ID” at the upper right corner of the Variable List shows the field ID of the selected variable on the Variable List.

Exhibit 8-16. ECB main screen

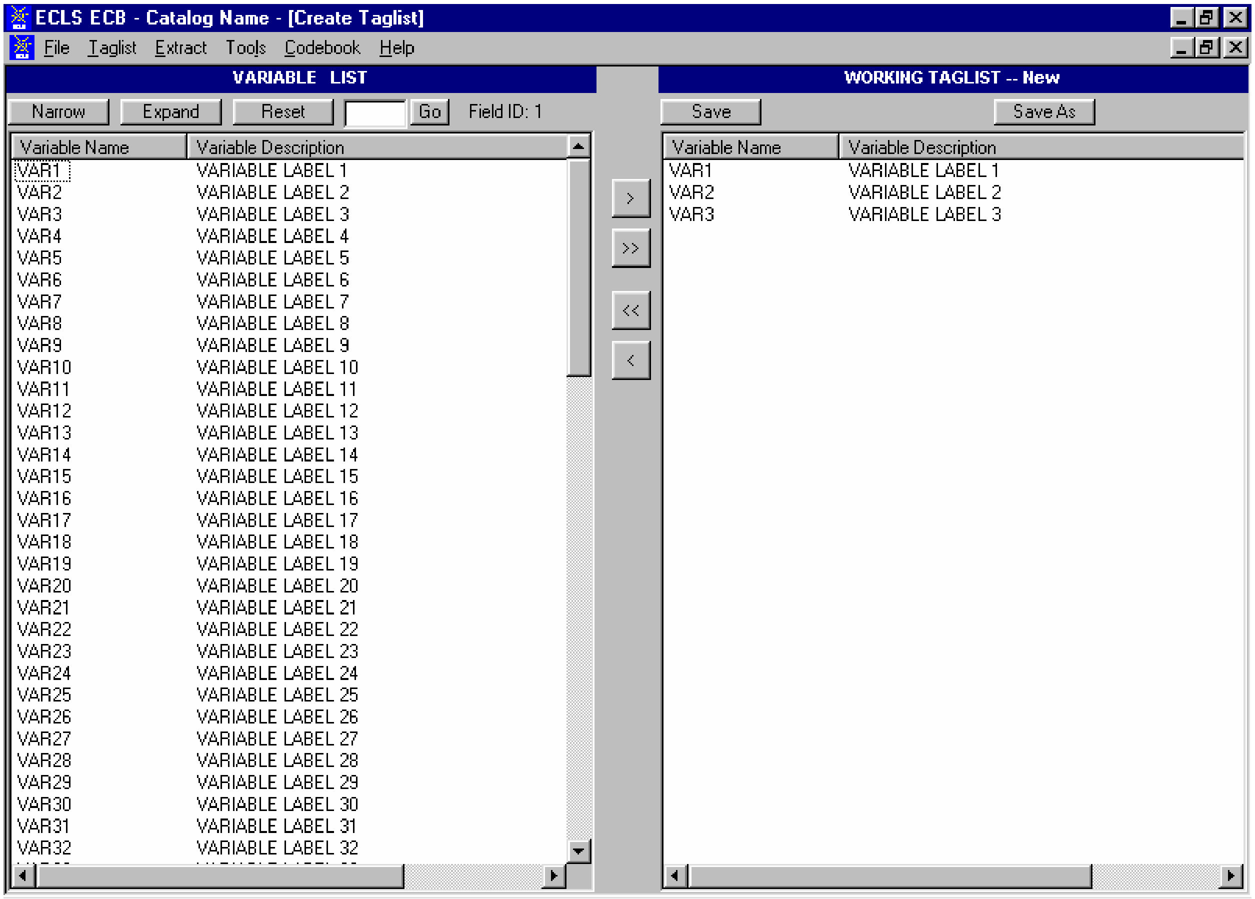
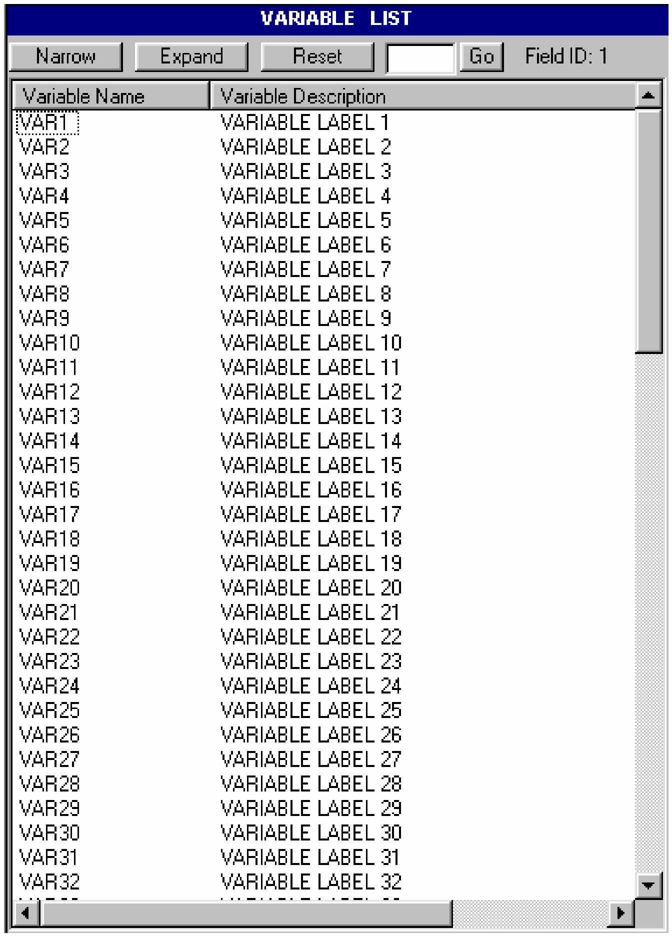


Exhibit 8-17. Variable List



The Variable List provides you with a comprehensive means of reviewing and identifying the variables that you want to use. To help you select the desired variables, the ECB provides you with the following capabilities:

- Perform searches of variable names and descriptions (see section 8.3.1);
- View code book information for each variable (see section 8.4.9); and
- Move selected variables to a Working Taglist (see section 8.4.2).

8.3.1 Searching the Code Book for Variables

The ECB allows you to search a catalog's Variable List for variables meeting criteria you specify. The Narrow Search and Expand Search functions are used to develop and refine the variables listed in your Variable List before adding them to your Working Taglist. Help screens with topical variable groupings were designed for each catalog to expedite searching. The catalog-specific topical variable groupings can be found in exhibit 8-60 at the end of this chapter.

8.3.1.1 Using the "Go" Button

Using the Go button, located at the top of the Variable List column, allows you to quickly move to a particular variable in the Variable List. You use the field ID presented in the help screens described earlier.

How To Use the Go Button:

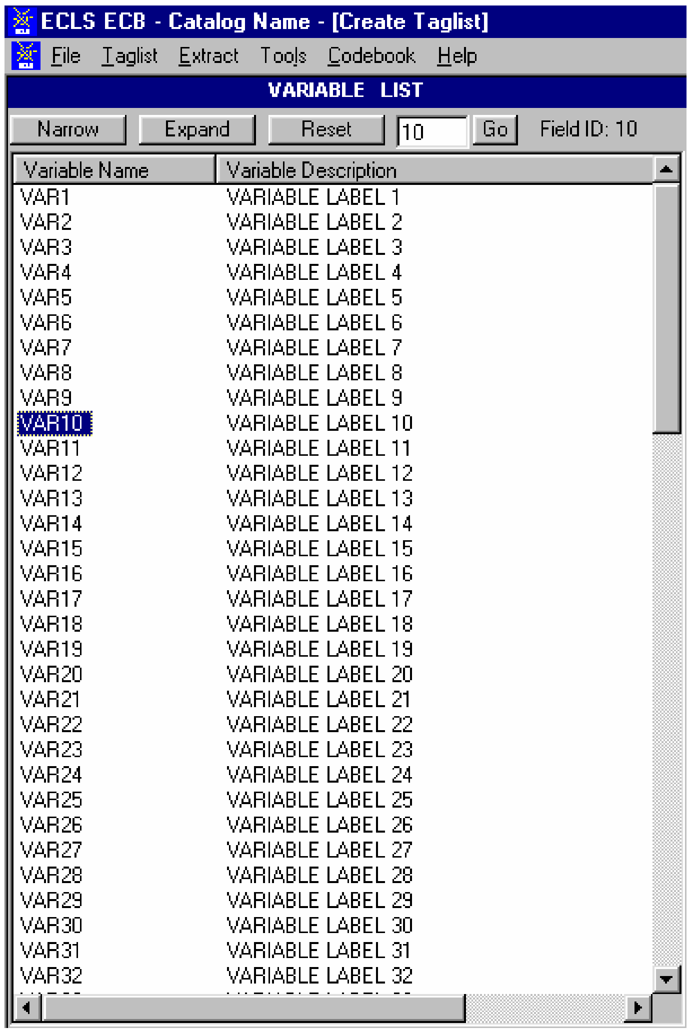
1. Type the field ID in the input box on the left of the Go button.
2. Click on the Go button.
3. The Variable List will then scroll down automatically to show the selected variable.
4. The selected variable is highlighted.
5. The field ID of the current variable selected is shown on the right of the Go button (exhibit 8-18).
6. Click the Reset button to return to the top of the original Variable List (Field ID 1) or enter another field ID to scroll to another variable.

For field IDs that identify different groups of variables, please refer to exhibit 8-60 for the catalog-specific topical variable groupings.

The Go button will not be available in a narrowed or expanded list. After a Narrow Search or an Expand Search, you must reset the Variable List (see section 8.3.1.4) before you can use the Go button.

The "Field ID" remains active in a narrowed or expanded list. However, the field IDs indicate the order of the variables in the catalog rather than that in the Variable List. As a result, the field IDs would not change in a narrowed or expanded list.

Exhibit 8-18. "Go" button



8.3.1.2 Narrowing Your Variable Search

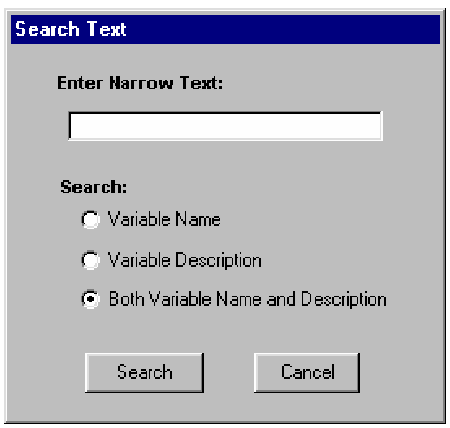
The Narrow Search function can be used to narrow the list of variables displayed in the Variable List. Since some catalogs have several thousand variables, this feature helps eliminate the variables that do not apply to your analysis. In performing the Narrow Search, you can enter key characters, words, or phrases as your criteria for searching the variable names, variable descriptions, or both. Use the keywords in the "Variable Identifier" column in exhibit 8-60 to search the catalog variables by variable description. If "Variable Name" appears under the variable in the Variable Identifier column, you must search by variable name. Also, the Narrow Search can be performed multiple times allowing you to repeatedly refine the list of variables displayed in the Variable List column.

Performing the Narrow Search function will only narrow down the variables listed in the Variable List window and will not affect those in the Working Taglist window.

How To Conduct a Narrow Search:

1. Click on the Narrow button located above the Variable List window.
2. The Narrow Search dialog box appears as shown in exhibit 8-19.

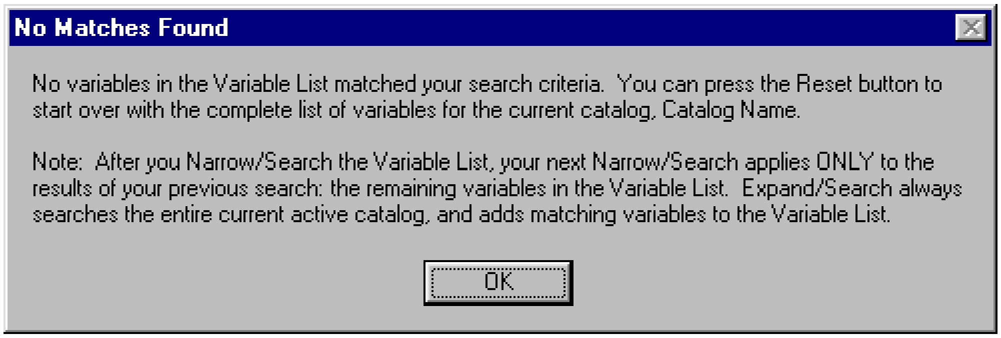
Exhibit 8-19. Narrow Search Text dialog box



3. Enter a key character string, word, or phrase in the Enter Narrow Text field. Character strings can include a single alphanumeric character or a sequence of several characters. The search is not case sensitive. The results returned will be all entries that contain that exact sequence of letters, numbers, spaces, and words.
4. Click in the Variable Name, Variable Description, or Both Variable Name and Description radio button to specify where to search.
5. Click on the Search button to initiate the search.
6. The variables meeting the specified criteria will be displayed in the Variables List column.

If no variable names or descriptions in the catalog contain the specified search text, then the message shown in exhibit 8-20 will appear.

Exhibit 8-20. No Matches Found message



7. Repeat the Narrow Search procedure if necessary.

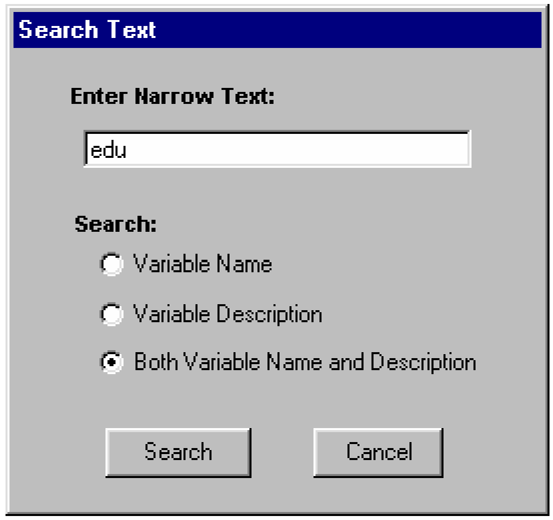
Please note that the field ID at the upper right corner of the Variable List reflects the order of the variables in the catalog rather than that in the narrowed Variable List.

Example of Narrowing a Search

The following example shows you how to narrow the Variable List. In this example, you want to include all the variables from the catalog that measure education. Do the following:

1. In the Variable List, click on the Narrow button.
2. In the Search Text Box (shown in exhibit 8-21), type in “edu” and then click on the Search button.

Exhibit 8-21. Example of narrowing a search



3. The new Variable List will include only the variables that have the text "edu" in the variable name or the variable description.

Exhibit 8-60 (at the end of this chapter) displays the appropriate keywords from the associated catalog. Simply find the topic of interest in the Topic column first. And then enter in the Search Text Box the matching keywords in the Variable Identifier to narrow the search.

8.3.1.3 Expanding Your Variable Search

The Expand Search function can be used to expand a previously narrowed list of variables displayed in the Variable List. After performing a Narrow Search operation, you can add variables to your current Variable List that meet your specified criteria. In performing the Expand Search, you can enter key characters, words, or phrases as your criteria for searching the variable names, variable descriptions, or both. Also, the Expand Search can be performed multiple times, allowing you to repeatedly expand the list of variables displayed in the Variable List column.

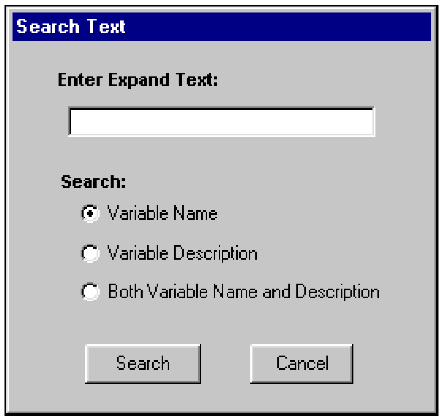
Performing the Expand Search function will only expand the variables listed in the Variable List window and will not affect those in the Working Taglist window.

How To Conduct an Expand Search:

1. Click on the Expand button located above the Variable List window.

2. The Expand Search dialog box will appear as shown in exhibit 8-22.

Exhibit 8-22. Expand Search Text dialog box



3. Enter a key character string, word, or phrase in the Enter Expand Text field. Character strings can include a single alphanumeric character or a sequence of several characters. The search is not case sensitive. The results returned will be all entries that contain that exact sequence of letters, numbers, spaces, and words.

4. Click in the Variable Name, Variable Description, or Both Variable Name and Description radio button to specify where to search.

5. Click on the Search button to initiate the search.

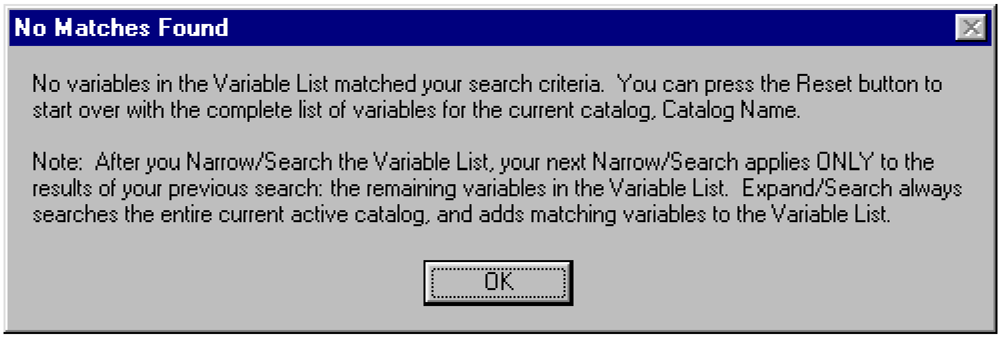
6. The variables meeting the specified criteria will be added to the variables already displayed in the Variables List column.

7. Repeat the Expand Search procedure if necessary.

If no variable names or descriptions in the catalog contain the specified search text, then the message shown in exhibit 8-23 will appear.

Please note that the field ID at the upper right corner of the Variable List reflects the order of the variables in the catalog rather than that in the expanded Variables List.

Exhibit 8-23. No Matches Found message



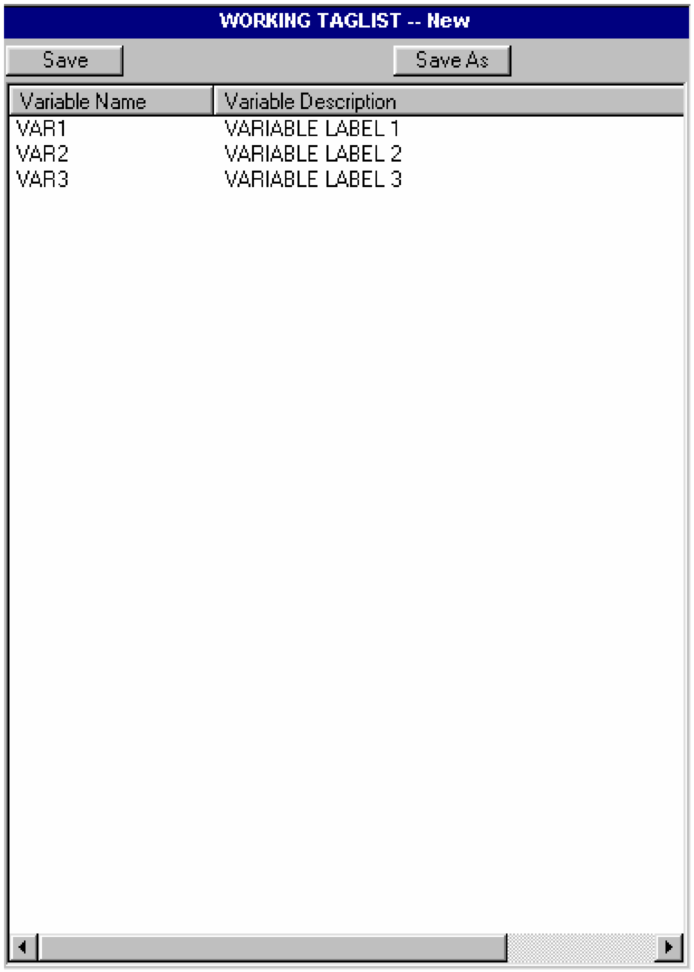
8.3.1.4 Resetting Your Variable List

Following a narrowing or expanding of the Variable List as described earlier, it is possible to reset the list to display ALL of the variables available in the catalog. The Variable List is reset by clicking on the Reset button located at the top of the Variable List column. Resetting the Variable List does not affect the variables listed in the Working Taglist.

8.4 Working Taglist

The Working Taglist, shown in exhibit 8-24, displays a list of variables that are currently selected or tagged for extraction. All Working Taglists contain a set of variables, called required variables that will be automatically included in all data files that the user creates. The required variables provide a foundational data set upon which other variables rely. These required variables cannot be untagged or deleted from the Working Taglist by the user. When a catalog is first opened, the default Working Taglist consists of only the required variables for that catalog. (See exhibit 8-58 for the catalog-specific required variables.) To create a taglist, add the variables you have selected to the required variables.

Exhibit 8-24. ECB Working Taglist



8.4.1 Opening a Taglist

The ECB allows you to open a predefined or previously saved taglist and display it in the Working Taglist column. Taglists, however, are saved as part of a particular catalog and can only be opened as part of the associated catalog.

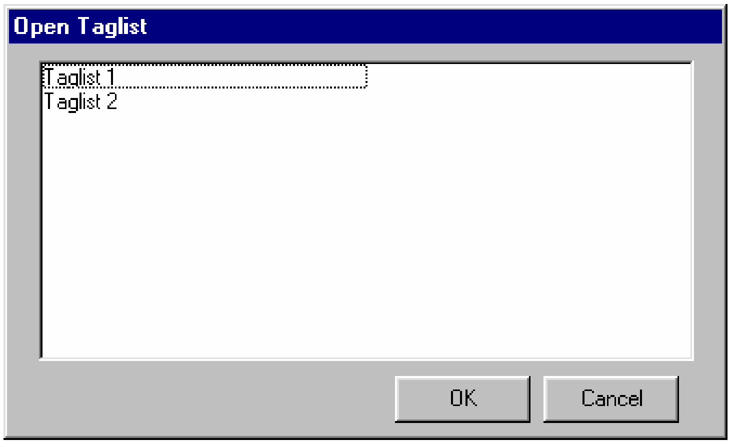
How To Open a Taglist:

1. Open a catalog.

2. Select Open from the Taglist pulldown menu.

3. The Open Taglist dialog box, shown in exhibit 8-25, appears.

Exhibit 8-25. Open Taglist dialog box



4. Highlight the taglist that you wish to open.

5. Click on the OK button.

If you have made modifications to the taglist currently open in the Working Taglist column, you will be prompted to save your changes.

8.4.2 Adding Variables to the Working Taglist

Variables can be added to your Working Taglist after you have identified the variables in the ECB's catalog that you want to extract. The user-selected variables can be added to the Working Taglist by selecting one of the two command buttons described in exhibit 8-26. The Working Taglist may also have variables added to it from a previously saved taglist. When moving or adding variables to the Working Taglist, the ECB will not permit variables to be listed multiple times. This is an automatic feature of the ECB.

Exhibit 8-26. Add variables buttons

Command Button	Description
>	The Tag button moves variables that are selected in the Variable List to the Working Taglist for extraction.
>>	The Tag All button moves all variables in the Variable List to the Working Taglist for extraction.

Multiple variables can be selected by using the following Microsoft Windows© techniques:

- Simultaneously pressing the <SHIFT> + Up/Down arrow keys or
- Pressing <CTRL> + left-mouse clicking on the items to be selected (or deselected). Also, <SHIFT> + left-mouse clicking extends the selection to include all list items between the current selection and the location of the click.

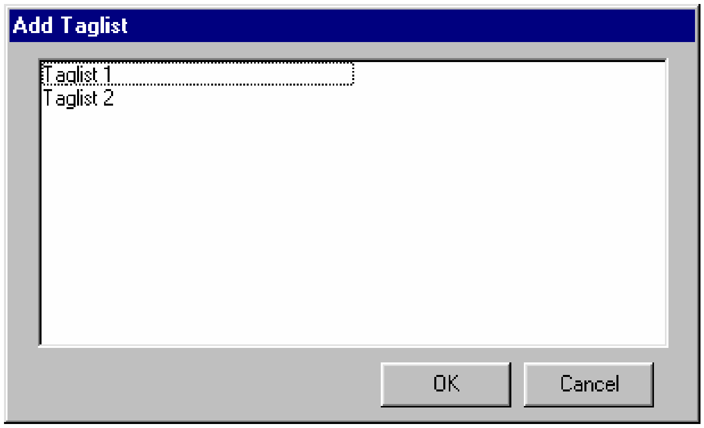
How To Add Variables to a Working Taglist:

1. Highlight the variable(s) in the Variables List that you wish to add. (See Microsoft Windows© techniques discussed earlier.)
2. Click on the Tag button, and the selected variables are added to your Working Taglist. To add all variables from the catalog displayed in the Variable List window to your Working Taglist, click on the Tag All button.

How To Add Variables From Another Taglist:

1. Click on the Taglist pulldown menu to display the menu options.
2. Select the Add option to display a list of previously saved taglists, shown in exhibit 8-27.
3. Highlight the saved taglist whose variables you wish to add to your Working Taglist.
4. Click on the OK button.
5. The new variables are added to your Working Taglist.

Exhibit 8-27.   Add Taglist dialog box



8.4.3 Removing Variables From the Working Taglist

Variables are removed from your Working Taglist by selecting one or more of the nonrequired variables and clicking one of the two command buttons described in exhibit 8-28. All variables can be removed by clicking on the Untag All button. All but the required variables will be deleted from your Working Taglist. Required variables are variables that are automatically extracted for all user-created files and cannot be removed from the taglist by the user.

Exhibit 8-28.   Remove variables buttons

Command Button	Description
<	The Untag button removes variables that are selected from the Working Taglist.
<<	The Untag All button removes all non-required variables from the Working Taglist.

Attempting to remove or untag required variables from the Working Taglist is not permitted by the ECB. A message will be displayed indicating that the required variable cannot be untagged.

How To Untag Variables From the Working Taglist:

1. Highlight the variable(s) in the Working Taglist that you wish to remove. (See Microsoft Windows© techniques discussed in previous page.)

2. Click on the Untag button, and the selected variables are removed from your Working Taglist. To remove all nonrequired variables from the Working Taglist, click on the Untag All button.

8.4.4 Saving Taglists

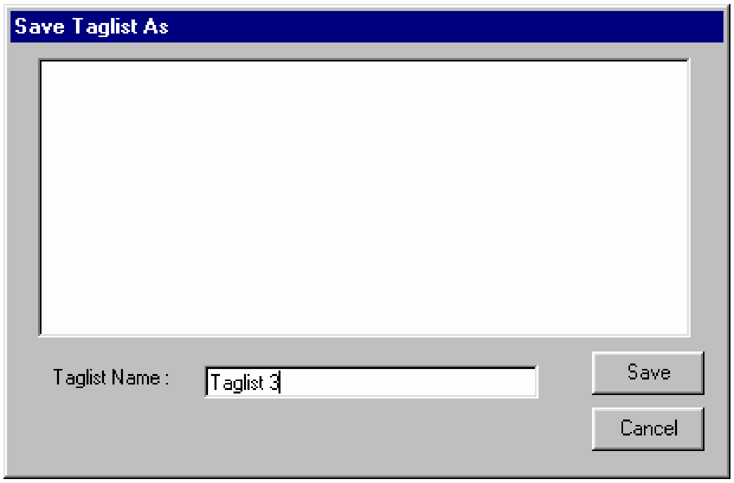
The ECB has the ability to save the newly created or modified taglist displayed in the Working Taglist column. Taglists can be saved either under the name already assigned or under a new name. If you have opened a new taglist and have not yet assigned it a name, you will be presented with the Save As dialog box. If you have opened a predefined taglist and have made modifications to it, you must save the modified taglist to a new name. You will also be prompted to save your Working Taglist changes if you attempt to close the catalog or if you open or import another taglist.

How To Save a <u>New</u> Taglist:

1. Complete any changes you wish to make to the new taglist.
2. Click on the Save or Save As button above the Working Taglist column. You can also select the Save or Save As options from the Taglist pulldown menu.
3. The Save Taglist As dialog box appears as shown in exhibit 8-29.
4. Enter the new name for the taglist in the Taglist Name field.
5. Click on the Save button.
6. The newly assigned taglist name now appears in the Working Taglist header bar.

If a name that already exists is entered, you will be prompted to replace the old taglist with the new taglist. Click “Yes” only if you wish to replace the old taglist with the new taglist.

Exhibit 8-29. Save Taglist As dialog box



How To Save an Existing Taglist Under a New Name:

1. Complete any changes you wish to make to the existing taglist.

2. Click on the Save As button above the Working Taglist column. You can also click on the Taglist pulldown menu and select the Save As option.

3. The Save Taglist As dialog box appears, shown in exhibit 8-30, with the current taglist name in the Taglist Name field.

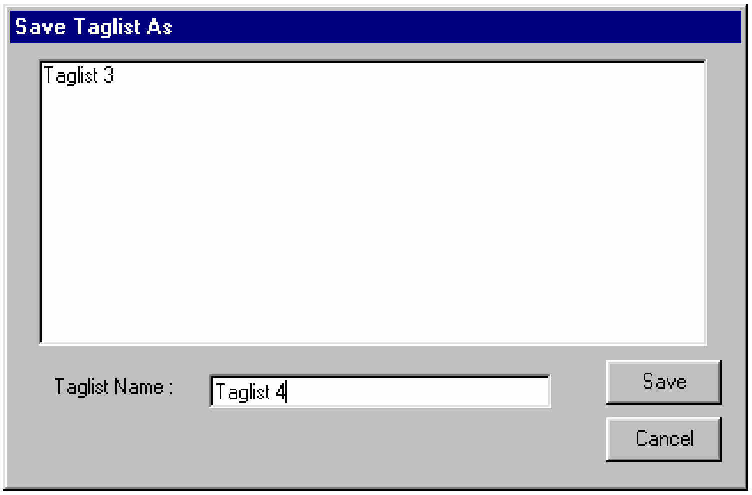
4. Enter the new name of the taglist in the Taglist Name field.

5. Click on the Save button.

6. The newly assigned taglist name now appears in the Working Taglist header bar.

If a name that already exists is entered, you will be prompted to replace the old taglist with the new taglist. Click "Yes" <u>only</u> if you wish to replace the old taglist with the new taglist or enter a unique name.

Exhibit 8-30. Save Taglist As dialog box (#2)



8.4.5 Exporting Taglists

Taglists can be saved as external files (*.tlt) for distribution. However, the exported files should be accessed only through the ECBs. Manually modifying the files outside of the ECB software is not recommended.

How To Export a Taglist:

1. Add to the Working Taglist all the variables that you would like to export.
2. Click on the Taglist pulldown menu (exhibit 8-31) and select the Export option.
3. The Export Working Taglist To dialog box appears.
4. Enter the file name for your taglist.
5. Click on the Save button.
6. You will be prompted to replace the file if the file name you entered already exists. Do so or click on "No" to enter a new file name.

The Working Taglist will be saved under the filename you enter.

Exhibit 8-31. Pulldown menu to select Taglist Export

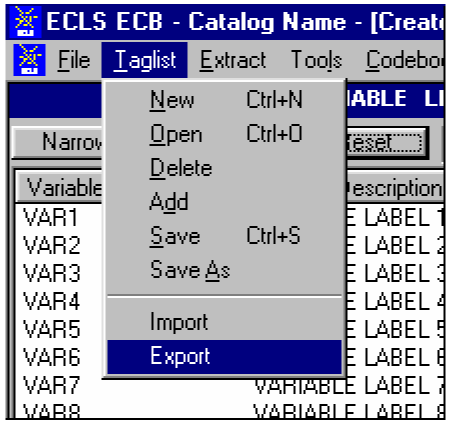
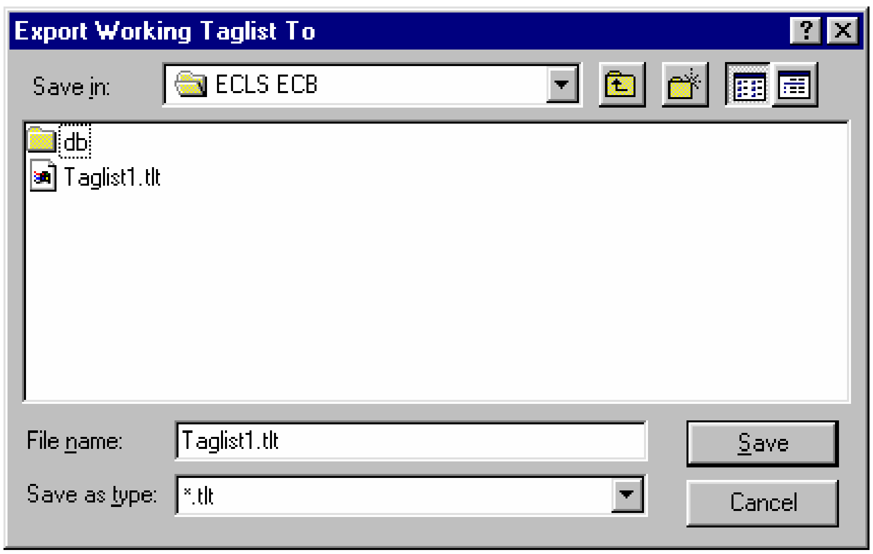


Exhibit 8-32. Export Taglist dialog box



8.4.6 Importing Taglists

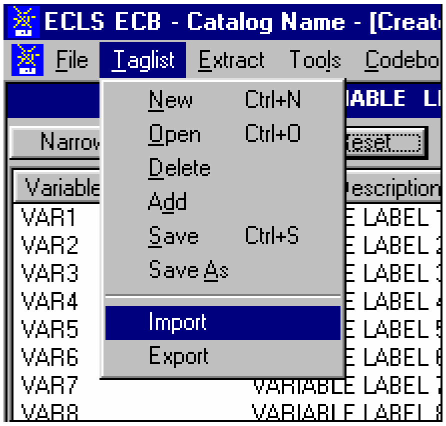
Taglists can be imported to the Working Taglist from external *.tlt files that are created by the ECB Taglist/Export function. Please note that only taglists exported from the same catalog of the same version ECB should be imported.

How To Import a Taglist:

1. Save the current Working Taglist before importing new taglist if desired.

2. Click on the Taglist pulldown menu (exhibit 8-33) and select the Import option.

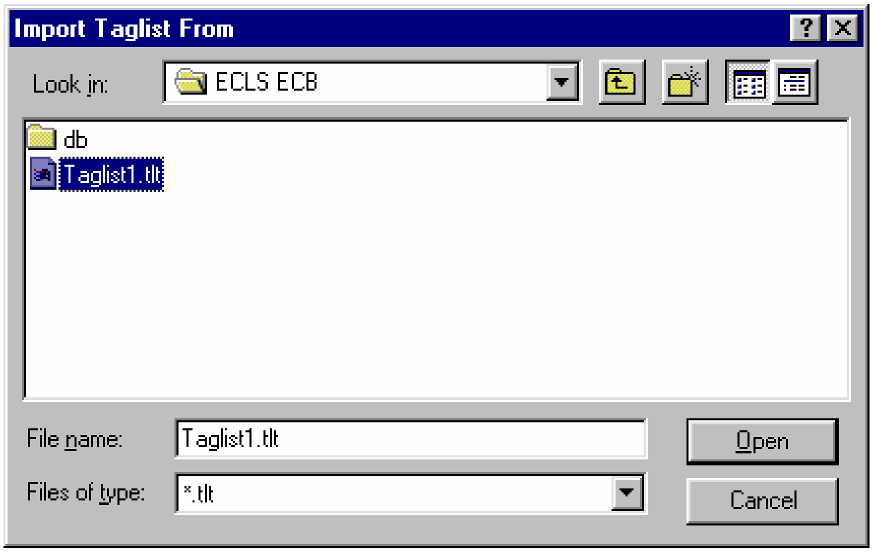
Exhibit 8-33. Pulldown menu to select Taglist Import



3. You will be prompted to save the current Working Taglist if unsaved changes have been made. Save the taglist if desired.

4. The Import Taglist From dialog box (exhibit 8-34) appears.

Exhibit 8-34. Import Taglist From dialog box



5. Enter the file name for the taglist you want to import.

6. Click on the Open button.

The Working Taglist will be replaced by the new imported taglist.

8.4.7 Using Predefined Taglists

The ECB provides predefined taglists that address specific topics. These predefined taglists can be added to your Working Taglist or can be opened as a new Working Taglist. Opening these predefined taglists is performed using the same steps as opening a user-saved taglist presented in section 8.4.1. Users can add as many of the predefined taglists as desired to the open Working Taglist. See section 8.7.2 for listings and descriptions of the catalog-specific predefined taglists.

8.4.8 Deleting Taglists

The ECB provides the capability to permanently delete previously saved taglists. Predefined taglists provided with the ECB, however, cannot be deleted through this function.

How To Delete a Taglist:

1. Close the taglist currently displayed in the Working Taglist column by selecting the New option from the Taglist pulldown menu.

2. The Working Taglist will be replaced by a New taglist.

3. Click on the Taglist pulldown menu and select the Delete option.

4. The Delete Taglist selection screen, shown in exhibit 8-35, appears with the taglists listed that may be deleted.

5. Highlight the taglist that is to be deleted and click on the OK button.

6. A confirmation screen, shown in exhibit 8-36, verifies your intention to delete the taglist.

7. Click on the “Yes” button to permanently delete the saved taglist.

Please note that you cannot delete the taglist that is currently open as the Working Taglist.

Exhibit 8-35. Delete Taglist selection

Exhibit 8-36. Delete Taglist confirmation window

8.4.9 Viewing Code Book and Variable Information

The code book for a taglist displayed in the Working Taglist column can be created, viewed, and printed from the ECB main screen. The code book displays several pieces of information about each variable that are described in exhibit 8-37.

Exhibit 8-37. Code book information

Field	Description
Question Text	The question that was asked of the respondent by the interviewer or that was on the self-administered instruments.
Variable Name/ Description	The name of the variable as it appears in the catalog and a brief description of its content.
Record Number	The row number of the variable within the catalog data file.
Format	The format of the variable. The first character is either “A” or “N” for alphabetical or numeric. Most variables are numeric except the identifiers—which begin with an “A.” The number following the “A” or “N” is the length of the variable. For numeric variables, the number after the decimal point is the number of decimal places.
Comment	Information to clarify specific information about a variable.
Position	The column number (position) of the variable within the catalog data file.
Response	A brief statement of each response code’s meaning.
Codes	The numeric codes specifying each response.
Frequency	The numeric count of respondents providing the corresponding response code. The frequency counts are unweighted.
Percent	The percentage of respondents providing the corresponding response code. The percents are unweighted.

How To View the Code Book for Tagged Variables:

1. Complete any changes you wish to make to the displayed taglist.

2. Click on the Code Book pulldown menu and select the View option.

3. The code book for the current taglist opens in a new window as shown in exhibit 8-38.

4. Use the buttons described in exhibit 8-39 to navigate through the displayed code book.

Exhibit 8-38. Code book view

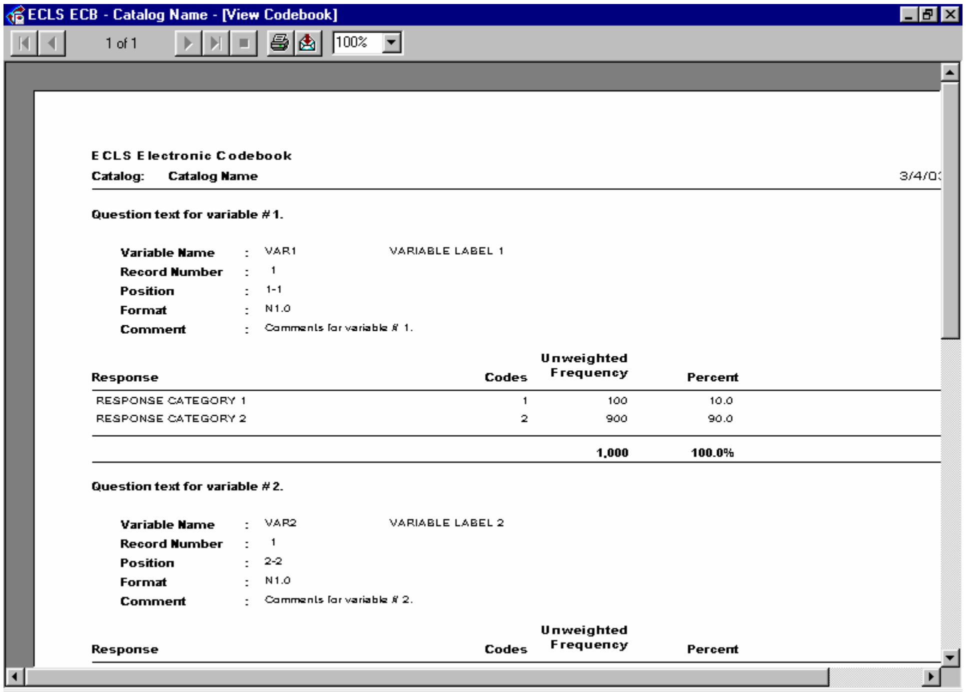


Exhibit 8-39. Navigation buttons

Command button	Description
	Click this button to change the displayed page to the first page.
	Click this button to change to the previous page.
	Click this button to advance to the next page.
	Click this button to change the displayed page to the last page.
	Click this button to discontinue a page change.
	Click this button to print the code book. Refer to the procedure below for steps on printing the code book.
	Click this button to export the code book to a different destination and save it as a different file format. Refer to the procedure below for steps on exporting the code book.
100%	Click the dropdown arrow to select a display magnification of the code book.

NOTE: The counter "1 of 1+" on the tool bar on top of the screen indicates the current page number and the last page number of the report. Users must navigate to the last page of the report to load the entire report. Once the user has viewed the last page of the report, the "+" sign will disappear and the correct last page number will show.

5. Once you have finished viewing the code book, close the screen by clicking on the Windows "X" control located in the top right corner of the window. You may also close the window using the other standard Windows defaults: by clicking on the windows icon in the upper left corner and selecting Close, or by pressing Alt-F4.

How To Print the Code Book:

1. Complete any changes you wish to make to the displayed taglist.

2. Click on the Code Book pulldown menu and select the Print option.

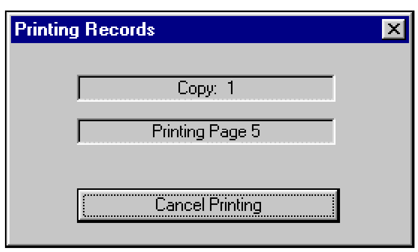
3. The Printing Status screen, shown in exhibit 8-40, appears, and the code book prints on your PC's default printer.

How To Export the Code Book:

1. Complete any changes you wish to make to the displayed taglist.

2. Click on the Code Book pulldown menu and select the View option.

Exhibit 8-40. Printing status screen

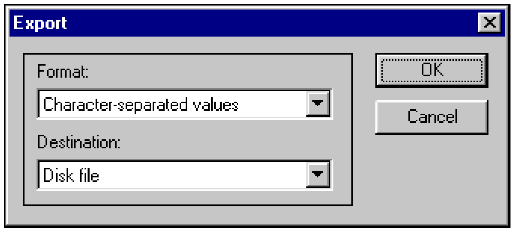


3. The code book for the current taglist opens in a new window, similar to the one shown in exhibit 8-38.

4. Click on the Export code book button: 

5. The Export code book selection screen, shown in exhibit 8-41, appears.

Exhibit 8-41. Export code book selection screen



6. Select the desired options from the “Format” pulldown menu and the “Destination” pulldown menu.

7. Click on the OK button and complete any subsequent screens required for exporting the file.

Please note that exporting a catalog in its entirety will take a long time due to the large size.

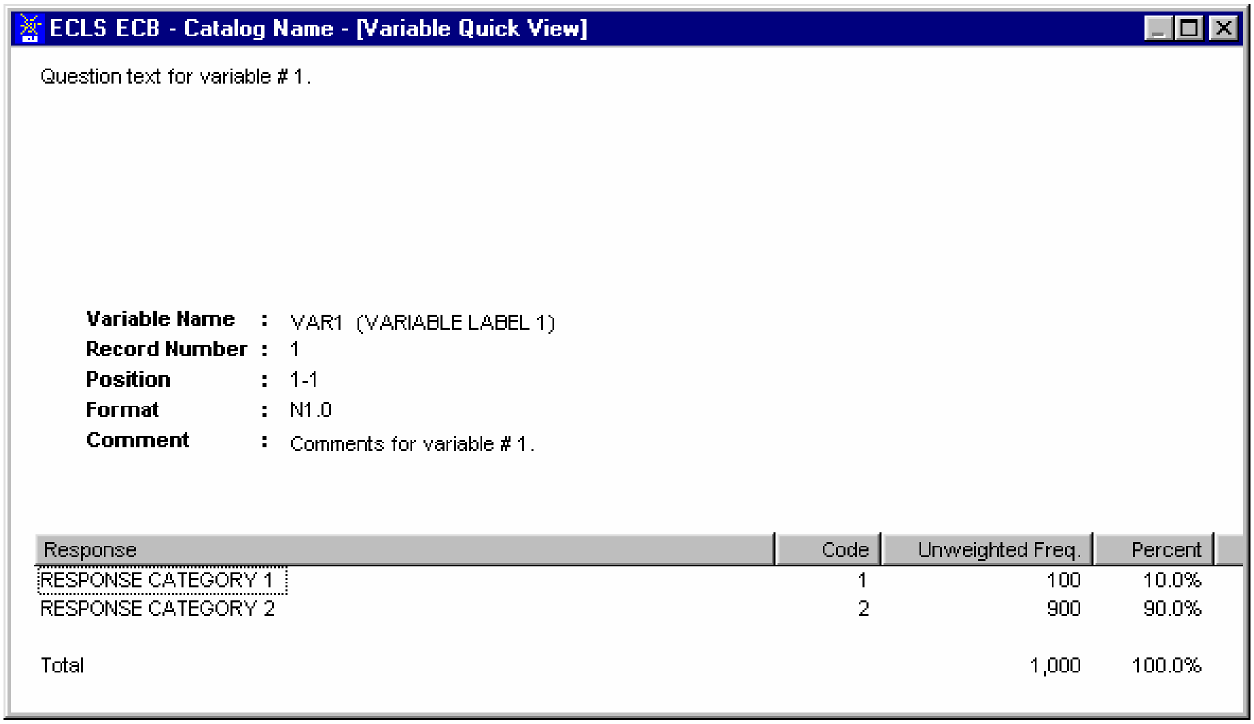
The code book and its variables can be selected to display their information from either the Variable List or the Working Taglist. The information that can be displayed for a variable includes the variable name and label, the question wording associated with the variable, the position and format of the variable on the data file, each response value and its label, unweighted frequencies, and the unweighted percentage distributions as listed on exhibit 8-37. The entire code book can also be viewed after moving all of the catalog’s variables to the Working Taglist. The following procedures describe how to view some or all code book variables:

How To Display Information for a Single Code Book Variable:

1. Locate the desired variable from either the Variable List or the Working Taglist.

2. Click on the variable name to highlight it and press <ENTER> -or- double-click on the variable name to view the variable information as shown in exhibit 8-42.

Exhibit 8-42. Variable Quick View



The Variable Name is the only field that can be highlighted for displaying the variable's code book information. Clicking on the variable description field will not activate the Variable Quick View.

3. When you are done reviewing the variable information, close the window by clicking on the Windows control "X" in the upper right corner of the screen. You'll return to the main screen.

How to Print Information for a Single Code Book Variable:

The ECB currently does not support printing the information for a single variable directly to the printer. If you must print the information for a single variable, follow these steps:

1. Double-click on the variable to activate the Variable Quick View (see the previous "How To" section for details).

2. With the Variable Quick View being the active window on top, press <Alt> + <Print Screen> to save the image of the Variable Quick View window.

3. In any application that supports bitmap images (e.g., Microsoft Paint, Microsoft Word, etc.), paste the saved image.

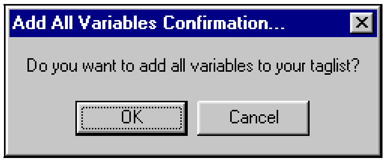
4. Print the image to the printer using the print function of the application that you are using.

How to Display and Print the Entire Code Book or Selected Pages:

1. Move all of the catalog's variables displayed in the Variable List to the Working Taglist by clicking on the Tag All button.

2. Click on the OK button of the Add All Variables Confirmation dialog box, shown in exhibit 8-43.

Exhibit 8-43. Add All Variables dialog box

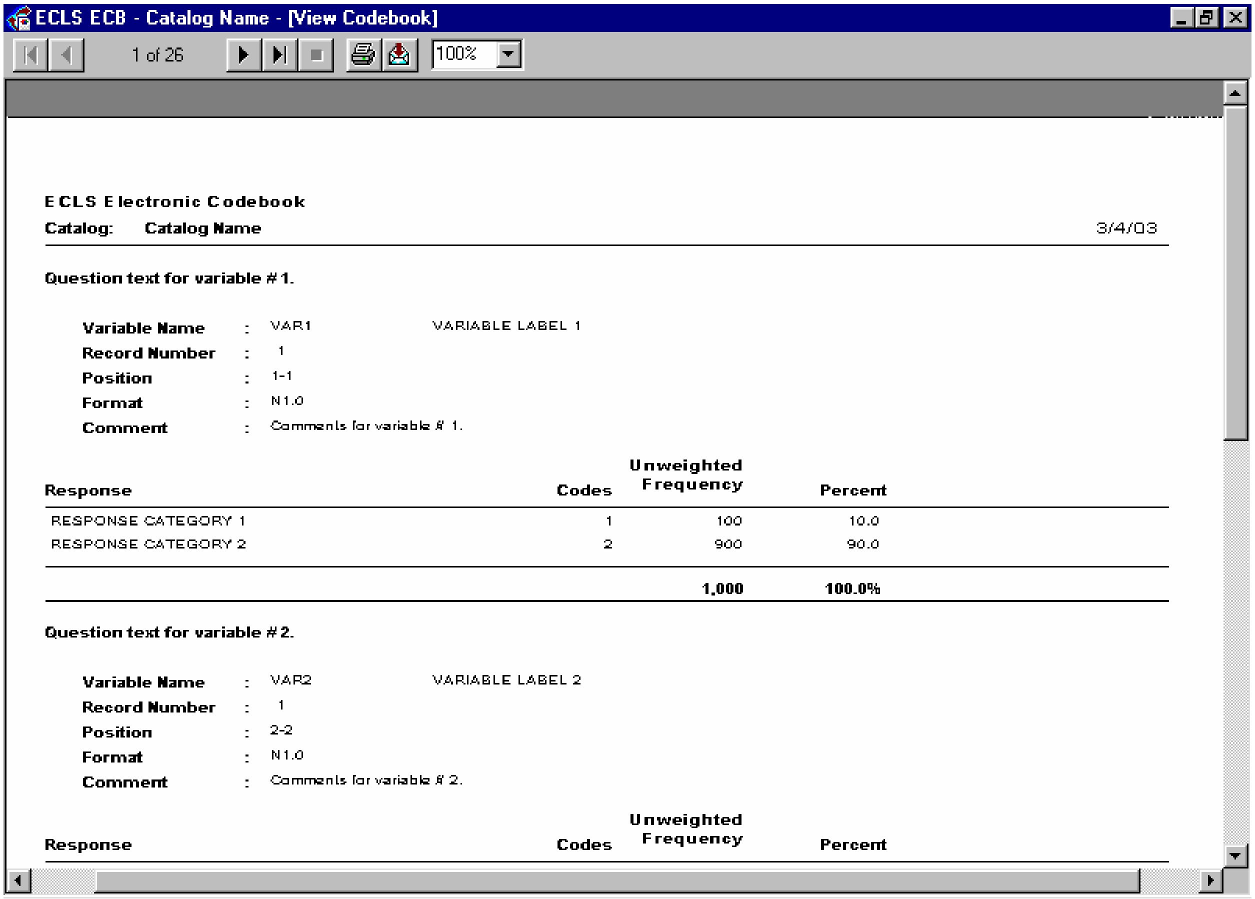


3. All of the variables listed in the Variable List are now displayed in the Working Taglist.

4. Select View from the Code Book pulldown menu.

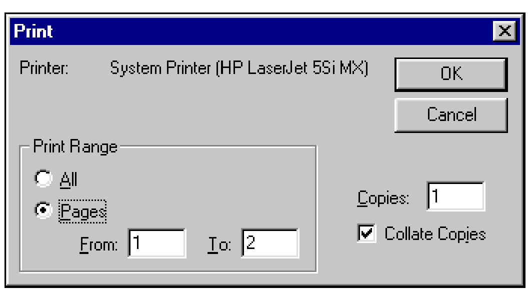
5. The entire code book displays as shown in exhibit 8-44. Note that this view includes ALL variables in the catalog and can span more than 1000 pages depending on the size of the ECB. The page number is in the upper left corner of the window.

Exhibit 8-44. View of the entire code book



6. To print the entire code book, click on the printer icon displayed at the top of the code book screen. Select ALL from the Printer Dialog box (exhibit 8-45). Enter the number of copies you want and click on the OK button.

Exhibit 8-45. Printer dialog box



7. To print selected pages of the code book, select Pages from the Printer Dialog box. Enter the pages you want to print and the number of copies you want. Click on the OK button.

8. When you are done viewing the entire code book, close the window by clicking on the Windows control “X” in the upper right corner of the screen. You will return to the main screen.

8.5 Extracting Data from the ECB

Once the variables have been selected (tagged) for extraction and reside in the Working Taglist, the next step is to generate the code through which the statistical analysis software can retrieve and display the results. The ECB provides options for generating the code for analyzing data with the SAS, SPSS for Windows, or Stata statistical analysis programs.

To run these programs, you will need the appropriate statistical software and the ECB CD-ROM from which the program can extract data.

SPSS users should note that an entire catalog can produce a Frequencies command statement with more than 500 variables. This may produce a warning of “too many variables,” and the Frequencies command will not execute. Users may work around this limitation by dividing the Variable List into two or more Frequencies commands.

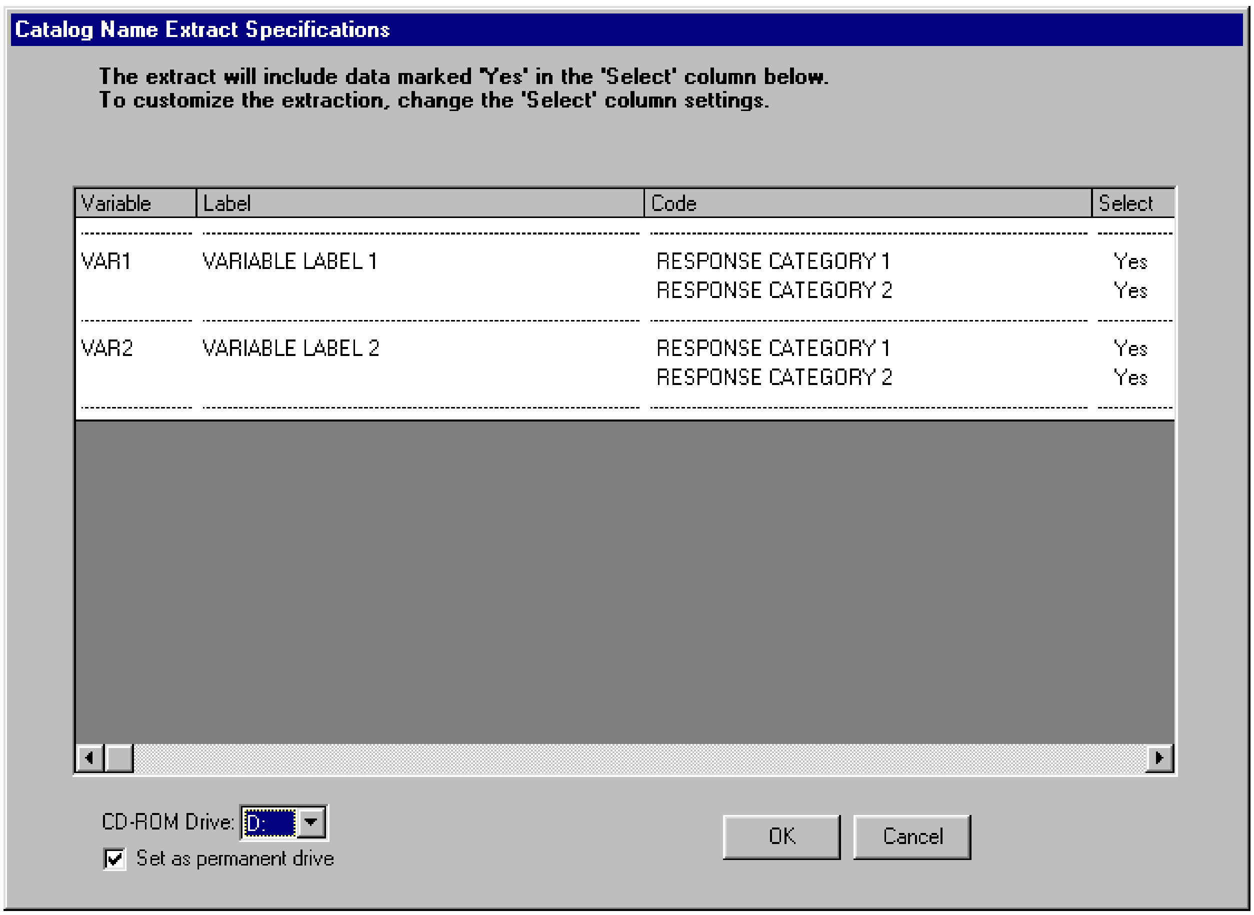
When extracting data to be used with either the SAS, SPSS for Windows, or Stata programs, a dialog box will be presented that allows the user to define the extract population through the Limiting Fields. See exhibit 8-46. The Limiting Fields include various subgroups of respondents that are typically of interest to analysts. These subgroups can be selected or deselected to narrow the data field that is extracted.

Also, please note that the ECB extract function allows the user to specify the drive letter of the CD-ROM drive. If you attempt to run the resulting SAS, SPSS, and Stata programs on a workstation with a different CD-ROM drive letter, you must alter the program code accordingly or regenerate the program code using the ECB.

The SAS, SPSS, or Stata source code generated by the ECB to read in the data may contain code statements that are "commented" out (e.g., with * in SAS). These code statements either run descriptive statistics (e.g., frequencies, means, etc.), or associate formats with variables. They are commented out because not all analysts will want them included in the source code.

SAS users (prior to SAS, Version 8) should note that, although the ECB will allow data set names larger than eight characters, the SAS system will reject these names at run-time.

Exhibit 8-46. Limiting fields dialog box



Refer to section 8.7.3 for instructions on using and modifying the catalog-specific limiting variables.

How To Extract a File to SAS Format:

1. Complete any changes you wish to make to the displayed taglist.

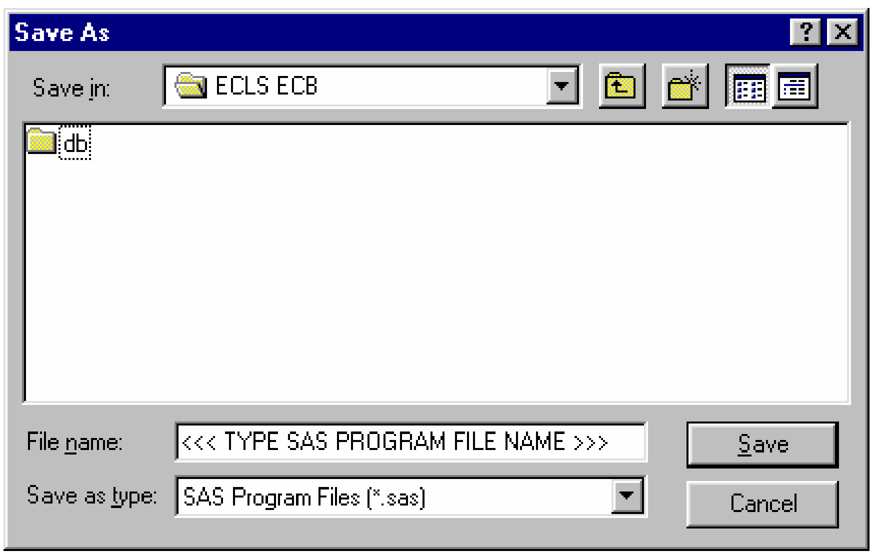
2. Click on the Extract pulldown menu and select the SAS option.

3. The Limiting Fields screen for the open catalog appears. Make your selections for each limiting variable indicator.

4. Verify that the ECB CD-ROM is loaded in your PC's default CD-ROM drive and then click on the OK button.

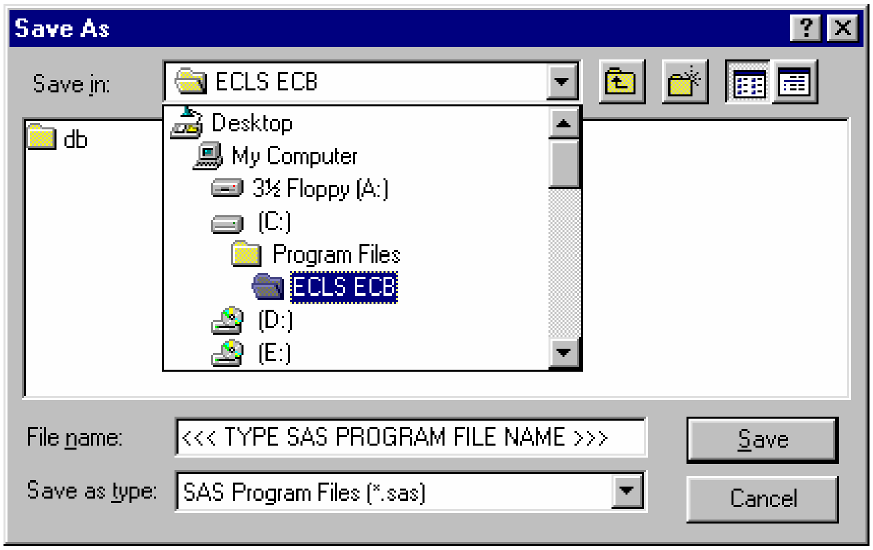
5. Type the desired name of the extract program file in the file name field of the screen shown in exhibit 8-47.

Exhibit 8-47. Save SAS program file dialog box



6. To save the file to another directory, click on the "Save in" dropdown menu button to browse to the new location, as shown in exhibit 8-48.

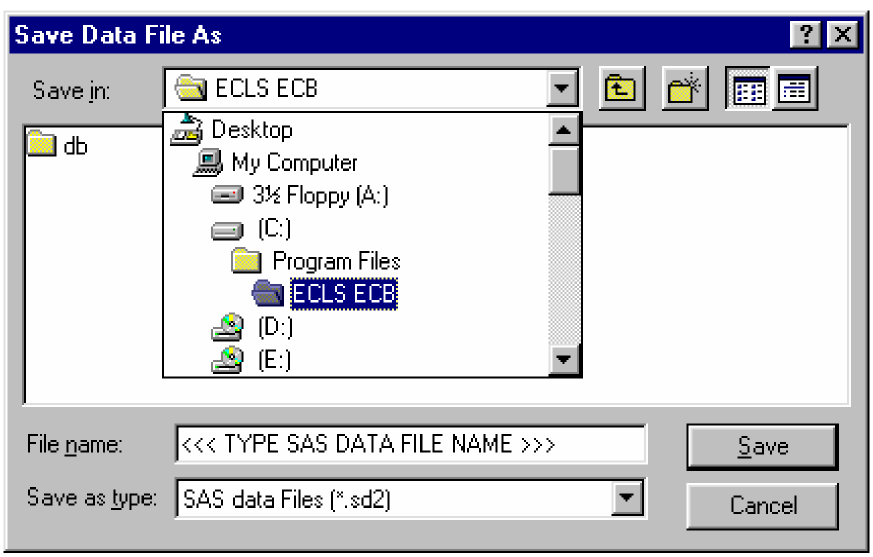
Exhibit 8-48. Save SAS program file location browse screen



7. Click on the Save button to store the file.

8. In the Save Data File As window (exhibit 8-49) type in the file name you want the data file to save to and then click on Save.

Exhibit 8-49. Save SAS data file dialog box



9. Run the saved extract program in SAS to extract the data.

How To Extract a File to SPSS Format:

1. Complete any changes you wish to make to the displayed taglist.

2. Click on the Extract pulldown menu and select the SPSS option.

3. The Limiting Fields screen for the open catalog appears. Make your selections for each limiting variable indicator.

4. Verify that the ECB CD-ROM is loaded in your PC's default CD-ROM drive and then click on the OK button.

5. Type the desired name of the extract program file in the file name field of the screen shown in exhibit 8-50.

6. To save the file to another directory, click on the "Save in" dropdown menu button to browse to the new location, as shown in exhibit 8-51.

7. Click on the Save button to store the file.

Exhibit 8-50. Save SPSS program file dialog box

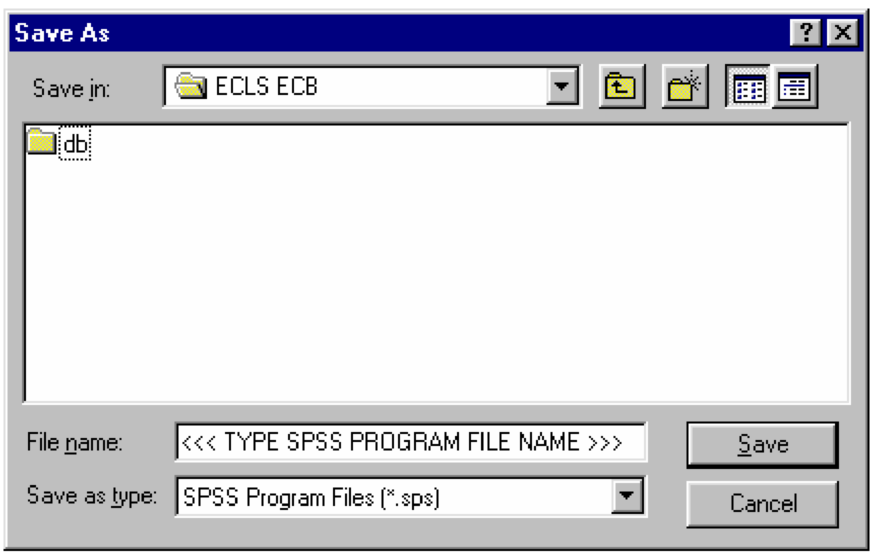
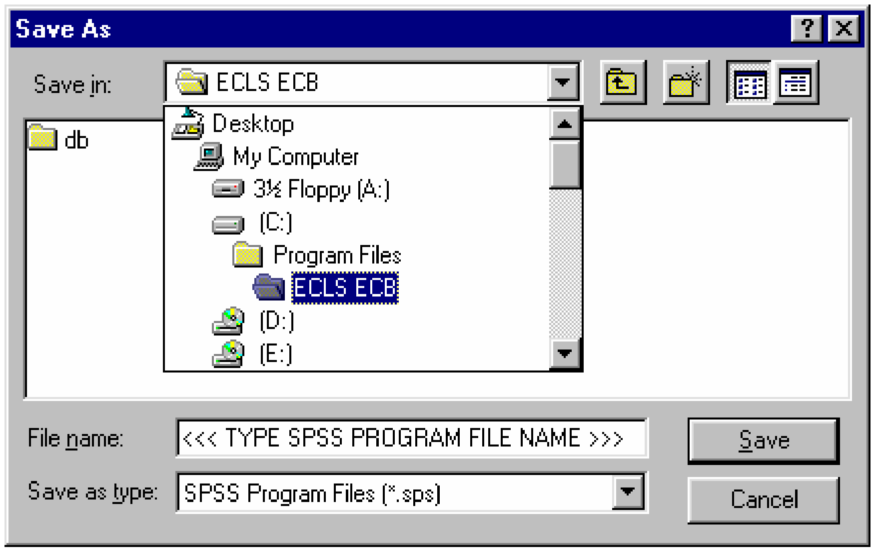


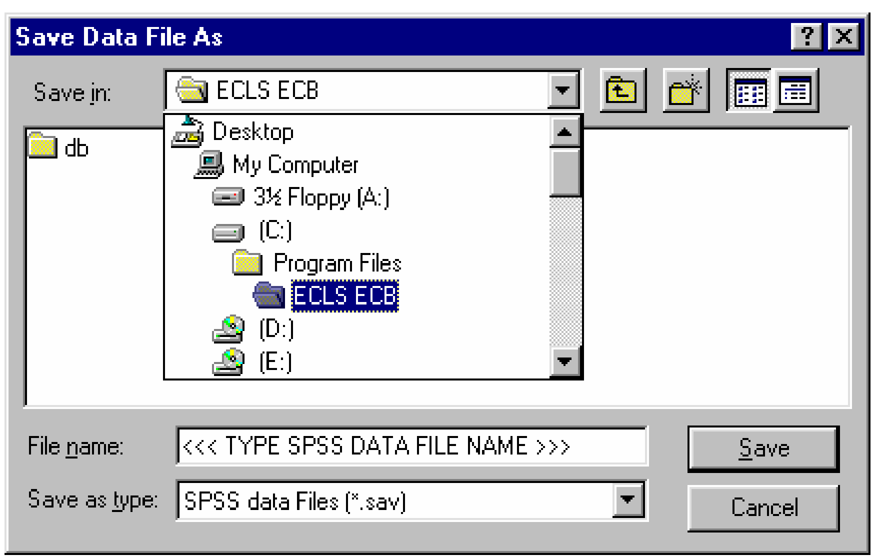
Exhibit 8-51. Save SPSS program file location browse screen



8. In the Save Data File As window (exhibit 8-52), type in the file name you want the data file to save to and then click on Save.

9. Run the saved extract program in SPSS to extract the data.

Exhibit 8-52. Save SPSS data file dialog box



How To Extract a File to Stata Format:

1. Complete any changes you wish to make to the displayed taglist.

2. Click on the Extract pulldown menu and select the Stata option.

3. The Limiting Fields screen for the open catalog appears. Make your selections for each limiting variable indicator.

4. Verify that the ECB CD-ROM is loaded in your PC's default CD-ROM drive and then click on the OK button.

5. Type the desired name of the extract program file in the file name field of the screen shown in exhibit 8-53.

6. To save the file to another directory, click on the "Save in" dropdown menu button to browse to the new location, as shown in exhibit 8-54.

7. Click on the Save button to store the file.

8. In the Save Data File As window (exhibit 8-55), type in the file name you want the data file to save to and then click on Save.

9. Run the saved extract program in Stata to extract the data.

Exhibit 8-53. Save Stata program file dialog box

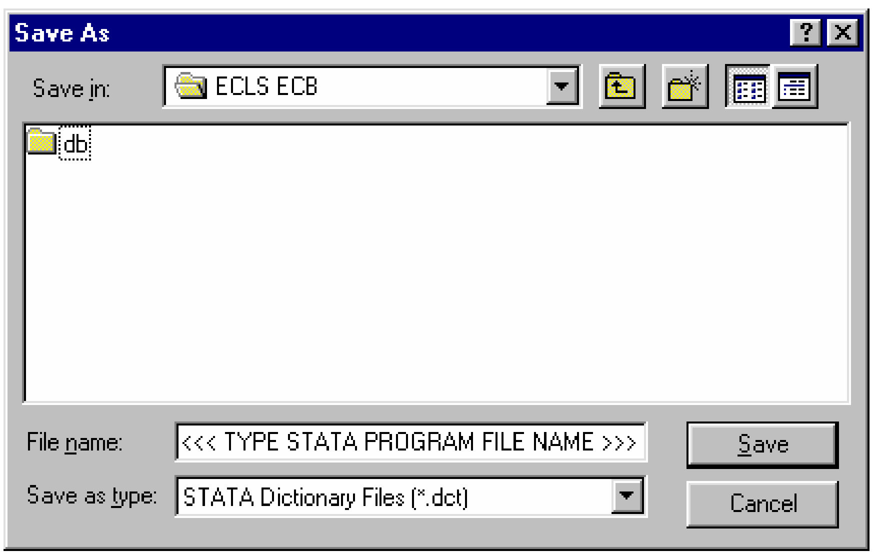


Exhibit 8-54. Save Stata program file location browse screen

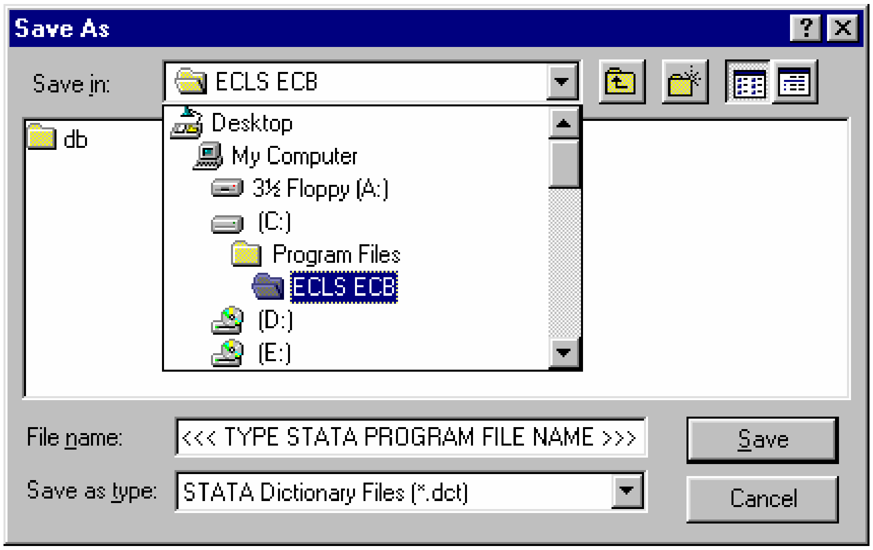
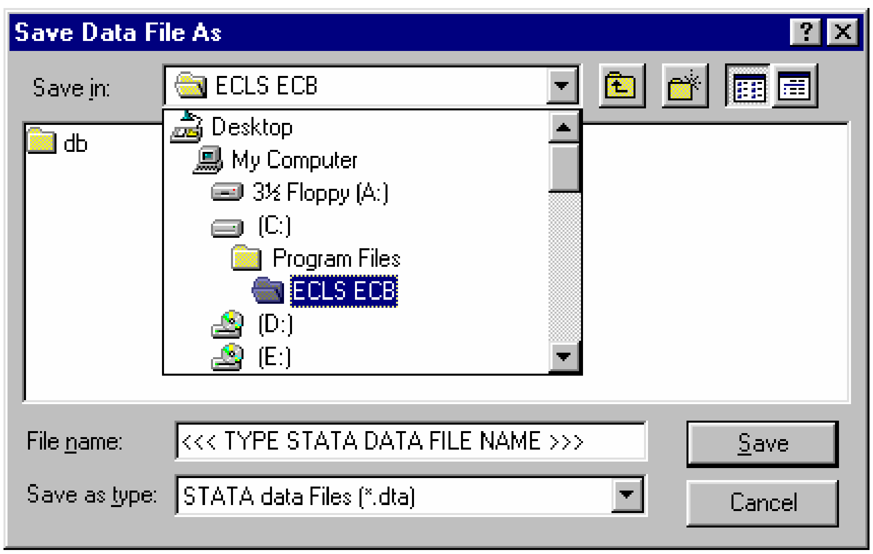


Exhibit 8-55. Save Stata data file dialog box



8.5.1 Reviewing the Extract Specifications

Users should review the SAS, SPSS, or Stata program code that is generated before running it to check that any statements subsetting the data are correct. Note that the ECB sometimes outputs superfluous code for selecting cases; this code is consistent with extract specifications, but users may wish to delete it.

If a mistake in defining the criteria is made, and it is not discovered until after writing out or running the extract program, it is very easy to correct if the taglist was saved before exiting the ECB program. Simply restart the ECB and select the appropriate catalog, open the taglist that you saved, define the extract criteria correctly by modifying the saved taglist as desired and saving it, and write out the extract program again. The program should be reviewed before running it because it may need to be customized.

8.5.2 Repairing and Compacting the Database

Periodically users may wish to repair and compact the database that contains the data of the ECB program. If many taglists are created and deleted on a regular basis, the database will contain lingering references to old taglists that are no longer needed. When the database is repaired and compacted, the ECB program "cleans house" and makes the database more efficient. It also decreases the size of the database, so space is conserved.

How To Repair and Compact the ECB Database:

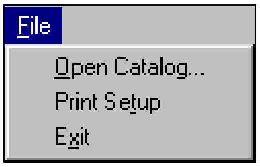
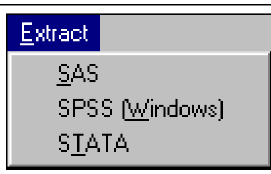
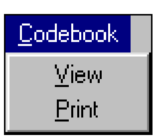
1. Select the Tools pulldown menu and select the Repair and Compact Database option.
2. After a few seconds, the screen shown in exhibit 8-56 appears indicating that the repair and compact of the database was successfully completed.

Exhibit 8-56. Repair database completed screen

3. Click on the OK button.

8.6 Menu Bar Descriptions

Exhibit 8-57. Menu Bar Descriptions

	
	The File menu contains the commands needed to: ■ Select and open a catalog; ■ Set up your software for printing; and ■ Exit the ECB.
	The Taglist menu contains the commands required to manipulate the variable lists once a catalog has been selected: ■ Create a new taglist; ■ Open a previously saved or predefined taglist; ■ Delete a previously saved taglist; ■ Add a previously saved or predefined taglist to the working taglist; ■ Save the working taglist; ■ Save a taglist with another name; ■ Import a previously exported taglist as working taglist and; ■ Export the working taglist for distribution.
	The Extract menu contains options to create a syntax file for: ■ SAS; ■ SPSS for Windows; or ■ Stata.
	The Tools menu contains: ■ The command for repairing and compacting the database.
	The Code Book menu contains the command for: ■ Viewing the entire code book based on the working taglist; and ■ Printing the entire code book based on the working taglist.
	The Help menu provides access to the detailed online help system.

8.7 Child Catalog

This section discusses catalog-specific information. For general information about how the ECB functions work, please refer to previous sections of chapter 8.

The 9-month child catalog contains information for children who participated in the 9-month ECLS-B survey and includes data at the child level from the child and from the child's parents. Exhibit 8-58 presents the pretagged required variables that will be automatically extracted for all user-created data files. The user cannot remove the required variables from the taglist. Five of the variables on the list, X1CHRACE, X1BTHWGT, X1PRIMNW, X1FTHTYP, and X1HPARNT, are also limiting fields; they may be used to subset the data being extracted. See section 8.7.3 for an explanation and description of limiting fields.

Exhibit 8-58. Child catalog required variables, 9-month data collection: 2001–02

Field Name	Field Label
I_ID	ECLS-B CASE ID
I_TWINID	ECLS-B TWINS CASEID
X1ASAGE	X1 CHILDS ASSESSMENT AGE IN MONTHS
X1CHSEX	X1 CHILD SEX
X1CHRACE	X1 RACE/ETHNICITY - CHILD
X1BTHWGT	X1 CHILD BIRTH WEIGHT STATUS
X1PRIMNW	X1 PRIM CARE ARRNGMNT WHERE MOST HRS/WK
X1FTHTYP	X1 TYPE RES FATHER-BIRTH/ADOPT/STEP/FOST
X1HPARNT	X1 CH PARENTS WHO RESIDE IN HOUSEHOLD
X1SESQ5	X1 QUINTILE INDICATOR FOR SOCIOECON SCAL

SOURCE: U.S. Department of Education, National Center for Education Statistics, Early Childhood Longitudinal Study, Birth Cohort, Nine-Month Data Collection, 2001–02.

The child catalog is organized into blocks of data corresponding to specific content areas:

- **Identifiers** are the child identification number (and twin's ID if applicable).
- **Full sample weights** are described in detail in chapter 4.
- **Composite variables**, including the child demographic variables, were created to facilitate data analysis. See chapter 7, section 7.5 for a detailed description of the

composite variables and table 7-6 for a list of all composite variables. Composite variables include the assessment scale scores, sociodemographic characteristics and child care information.

- **Observation data** are data from the interviewer's responses to questions about the child's behavior and affect during the child assessment and to questions about the household environment and the interaction between the parent and the child.

- **Parent interview data** are the data items from the parent interview.

- **Birth certificate data** are the data items from the birth certificate.

- **Resident father data** are the data items from the resident father questionnaire.

- **Nonresident father data** are the data items from the nonresident father questionnaire.

- **Methodological/Supplemental data** are the data items from the Interviewer Remarks Questionnaire, the Child Activity Booklet, the NCATS Score sheet and coder and quality control checklists, and the CAPI child assessment.

- **Weight replicates** are used to estimate standard errors using the paired jackknife replication method (JK2) with WesVar. Each replicate is a subsample of the full sample. The variability of the replicate estimates about the full sample estimate is used to estimate the variance of the full sample estimate.

- **Taylor Series Sampling Strata and primary sampling units (PSUs)** are the stratum and first-stage unit identifiers that are required for estimating standard errors with the Taylor Series linearization method using software such as SUDAAN or Stata.

Exhibit 8-59 presents the order of the data blocks on the child data file.

Exhibit 8-59. Child catalog blocks, 9-month data collection: 2001–02

Order	Block
1	Identifiers
2	Weights (full sample only)
3	Composites
4	Observation
5	Parent
6	Birth Certificate
7	Resident Father
8	Nonresident Father
9	Methodological/Supplemental
10	Replicate weights
11	Taylor Series variables

SOURCE: U.S. Department of Education, National Center for Education Statistics, Early Childhood Longitudinal Study, Birth Cohort, Nine-Month Data Collection, 2001–02.

8.7.1 Child Catalog Topical Variable Groupings

The variables within the child catalog are organized into topical categories to help locate and identify desired variables in the ECB. These categories are identified in exhibit 8-60. The first column of the exhibit describes the topic, the second column lists the variable identifiers, which generally indicate the source of the data (e.g., the parent interview, the resident father questionnaire). The third column is a description of the topic. The last column is the Field ID, which is used to search the data file for the topics. Use the keywords in the Variable Identifier column to search for variables while using the ECB Narrow or Expand function (see sections 8.3.1.2 and 8.3.1.3). If (Variable Name) appears under the variable in the Variable Identifier column, you must search using the variable name. The field ID can be used by the Go function in the Variable List (see section 8.3.1.1).

Exhibit 8-60 is only for your convenience in searching for variables and includes only a general description of variables that belong to certain topics. It does not reflect the contents of the entire child catalog. It is highly recommended that you search the full child catalog and the instruments to identify every item of interest.

Exhibit 8-60. Selected child catalog topical variable groupings, 9-month data collection: 2001–02

Topic	Variable name	Description	Field ID
Identifiers	I_ID, I_TWINID	Child identification numbers	1,2
Composite/ Derived Variables	X1ASAGE – X1CHPREM X1MTLTSC – X1NCATTC X1CHLENG – X1CHCRFM	Child characteristics (e.g., child sociodemographic characteristics and child care, child assessment scale scores, physical measurements, etc.)	8 – 19 26 – 45 46 – 49
	X1HFAMIL – X1HHURBN	Household characteristics (e.g., household composition, poverty status, income, household food sufficiency, primary language, etc.)	85 – 108
	X1FTHID – X1FTHSCR	Resident father characteristics (e.g., father type, age, race/ethnicity, education, employment status, etc.)	68 – 82

See note at end of exhibit.

Exhibit 8-60. Selected child catalog topical variable groupings, 9-month data collection: 2001–02—Continued

Topic	Variable name	Description	Field ID
Composite/ Derived Variables—Continued	X1MOMID – X1MOMSCR	Resident mother (e.g., mother type, age, race/ethnicity, education, employment status, etc.)	52 – 66
	X1NRFED – X1NRFEMP	Nonresident father characteristics (e.g., education, employment status)	83 – 84
	X1RESPID – X1RSPREL	Respondent characteristics (e.g., respondent's relationship to child)	50 – 51
	X1STATPI – X1STATPM	Component status (e.g., status of the parent CAPI instrument, status of the BSF-R mental and motor assessments, status of the NCATS assessment, status of the resident father questionnaire, status of the nonresident father questionnaire, etc.)	111 – 122
	X1TWSAMP – X1AISAMP	Related to sample (e.g., child sampled as part of twin, child sampled as part of American Indian population)	123 – 126
Observation	R1POSAFF – R1IOENVS	Includes directly observed information about the home environment, the respondent's behavior towards the child, and the child's behavior observed by the interviewer during the home visit	155 – 169
Parent CAPI instrument: Introduction	P1CSTATE – P1BMDOD	Includes respondent's date of birth, age, gender, relationship to child, and child date of birth	800 – 814
Parent CAPI instrument: Family Structure	P1 Roster P1AGE_01 – P1NFTHHH	Includes information about household members: relationship to child, age, gender, and race/ethnicity	170 – 798
Parent CAPI instrument: Child Development	P1BRSTFD – P1RAISE	Includes breast feeding, developmental milestones, and child temperament	815 – 840

See note at end of exhibit.

Exhibit 8-60. Selected child catalog topical variable groupings, 9-month data collection: 2001–02—Continued

Topic	Variable name	Description	Field ID
Parent CAPI instrument: Home Environment	P1ANYLNG – P1STTR	Includes primary language spoken in home, mealtime and bedtime routines, mother activities with child (e.g., read books, tell stories, play peek-a-boo, etc.), father activities with child (e.g., change diapers, put child to sleep, etc.), and father caretaking activities (e.g., get up with child during the night, take child to the doctor, etc.)	8411 – 909
Parent Interview: Parenting Attitudes and Expectations	P1PCKCRY – P1FTHGFT	Includes parenting expectations, expectations for role of fathers (father respondents only), and delight in child (father respondents only)	910 – 951
Parent CAPI instrument: Child Care Arrangements	P1SAMETW – P1CAREEG	Includes child care arrangements (i.e., relative care, nonrelative care, center care), characteristics of child care arrangements (e.g., primary language of provider, number of hours and days in care, amount household pays for care, etc.)	952 – 1012
Parent CAPI instrument: Child Health	P1CHEALT – P1HTHCAR	Includes child's health and hospitalization, child disabilities/conditions, services received, participation in early intervention programs, and information about health insurance	1013 – 1179
Parent CAPI instrument: Family Health	P1HEALTH – P1NMDK1S	Includes respondent's health, fertility services received, prenatal care, and smoking and drinking behavior	1180 – 1237
Parent CAPI instrument: Marital History and Partner Relationships	P1MARSTS – P1RELTBF	Includes current marital status and marital history, mother's relationship history with child's biological father, and status of child's biological father	1238 – 1258
Parent CAPI instrument: Social Support	P1ASKHLP – P1OTHERS	Includes social support from people outside of the family	1259 –1282

See note at end of exhibit.

Exhibit 8-60. Selected child catalog topical variable groupings, 9-month data collection: 2001–02—Continued

Topic	Variable name	Description	Field ID
Parent CAPI instrument: Community Support	P1NEIGHB – P1MNTSRV	Includes socializing with friends, attendance of religious services, and receipt of services from community agencies	1283 – 1290
Parent CAPI instrument: Respondent Information	P1BM16 – P1FT2WKU	Includes respondents' closeness to parents (if not a parent), school performance, education level, job training, currently attending courses, employment status, benefits through current job, maternity leave (if respondent is biological mother)	1291 – 1359
Parent CAPI instrument: Spouse/Partner Information	P1LIVSP – P1RTNPSQ	Includes spouse/partner's education level, job training, currently attending courses, employment status, benefits through current job, maternity leave (if spouse is biological mother)	1360 – 1413
Parent CAPI instrument: Nonresident Biological Father	P1BFDEVR – P1BFLEGL	Includes nonresident father's education, contact with child, and child support agreements	1414 – 1443
Parent CAPI instrument: Welfare and Other Public Assistance	P1WICBFT – P1WELFAS	Includes family's utilization of WIC vouchers, TANF, or food stamps	1444 – 1459
Parent CAPI instrument: Household Income and Assets	P1NMADER – P1HHCKSV	Includes household income and assets	1460 – 1474
Parent CAPI instrument: Household Food Sufficiency	P1WRRYFD – P1CHSKDY	Includes adult and child food sufficiency in the household	1475 – 1492
Parent CAPI instrument: Closing	P1FATHCO – P1LEAVPK	Includes permission to contact non-resident biological father, distribution of nonresident father questionnaire	1493 – 1495

See note at end of exhibit.

Exhibit 8-60. Selected child catalog topical variable groupings, 9-month data collection: 2001–02—Continued

Topic	Variable name	Description	Field ID
Parent Self-Administered Questionnaire	P1MARRIG – P1HLD1CH	Includes quality of the marital relationship, pregnancy wantedness	1496 – 1558
Birth Certificate Variables	BCDOBYY – BCOTHCON	Includes family/parent background, maternal/ neonatal history, and labor and delivery information	1559 – 1681
Resident Father Self-Administered Questionnaire	F1RELCH – F1SOCLZE	Includes father's relationship to child, activities with child, delight in child, involvement and influence in decisions made about child, expectations about the role of fathers, quality of the marital relationship, education and employment status, and health	1682 – 1865
Nonresident Father Self-Administered Questionnaire	N1SAWDY – N1HHINSP	Includes contact with child, involvement and influence in decisions made about child, quality of relationship with child's mother, child support agreements, education and employment status, and income	1866 – 1947
Methodological/ Supplemental Variables	R1CMPPRN – R1TELCNT	Includes interviewer remarks related to the home visit	1948 –1972
	C1INVRID – C1DIFOTH	Child Activity Booklet (CADE)	1973 – 1990
	C1SUPPRT – C1CHSTDS	NCATS score sheet	2039 – 2112
	C1COIDCC – C1CHSTTE	NCATS coders and quality checklists	2026 – 2038
	C1BSFWMS – C1NLTRL	Child assessment CAPI (BA, NC, PM)	2113 – 2116
Replicate Weights	W1R1 – W1F90	Specialized weights	2129 – 2488
Taylor Series Variables	W1RSTR – W1RRPSU		2489 – 2490

SOURCE: U.S. Department of Education, National Center for Education Statistics, Early Childhood Longitudinal Study, Birth Cohort, Nine-Month Data Collection, 2001–02.

8.7.2 Child Catalog Predefined Taglists

There are no predefined taglists provided with the child catalog in the ECLS-B.

8.7.3 Child Catalog Limiting Fields

The limiting fields for the child catalog appear in exhibit 8-61. These limiting fields allow codes within each variable to be included or excluded from the extraction depending on the selection indicator. For example, the user can select “Yes” for Twin status if he or she would like to create a data file that only includes twin births. The selection indicator will be either a “Yes” or “No” to specify whether the variable code should be included or excluded, respectively.

Exhibit 8-61. Child catalog—9-month limiting field variables: 2001–02

Field Name	Field Label	Selection Criteria
X1CHRACE	X1 RACE/ETHNICITY - CHILD	Child’s race—White, Black, Hispanic, etc.
X1BTHWGT	X1 CHILD BIRTH WEIGHT STATUS	Type of birth weight category—normal, moderately low or very low.
X1PRIMNW	X1 PRIM CARE ARRNGMNT WHERE MOST HRS/WK	Type of primary care arrangement where the child spends the most hours—center based, relative in-home care, nonrelative-in someone else’s home, etc.
X1FTHTYP	X1 TYPE RES FATHER-BIRTH/ADOPT/STEP/FOST	Type of resident father—biological father, step, foster, adoptive, etc.
X1HPARNT	X1 CH PARENTS WHO RESIDE IN HOUSEHOLD	Type of parent(s) that reside in the household—mother only, mother/biological father, mother/non-biological father, etc.

SOURCE: U.S. Department of Education, National Center for Education Statistics, Early Childhood Longitudinal Study, Birth Cohort, Nine-Month Data Collection, 2001–02.

The limiting fields feature for the catalog allows the user to create a subset of cases based on the settings of the Select column in the Extract Specifications window. The default setting is all “Yes” in the Select column meaning that all records will be present in the extract file. To exclude records in a particular category of a variable, change the “Yes” associated with that code in the Select column to “No” by double-clicking on it. At least one of the codes for each limiting variable must be selected as "Yes" or no records will be extracted for analysis.

Exhibit 8-62 is the child catalog Extract Specifications window.

Exhibit 8-62. Child catalog Extract Specifications window

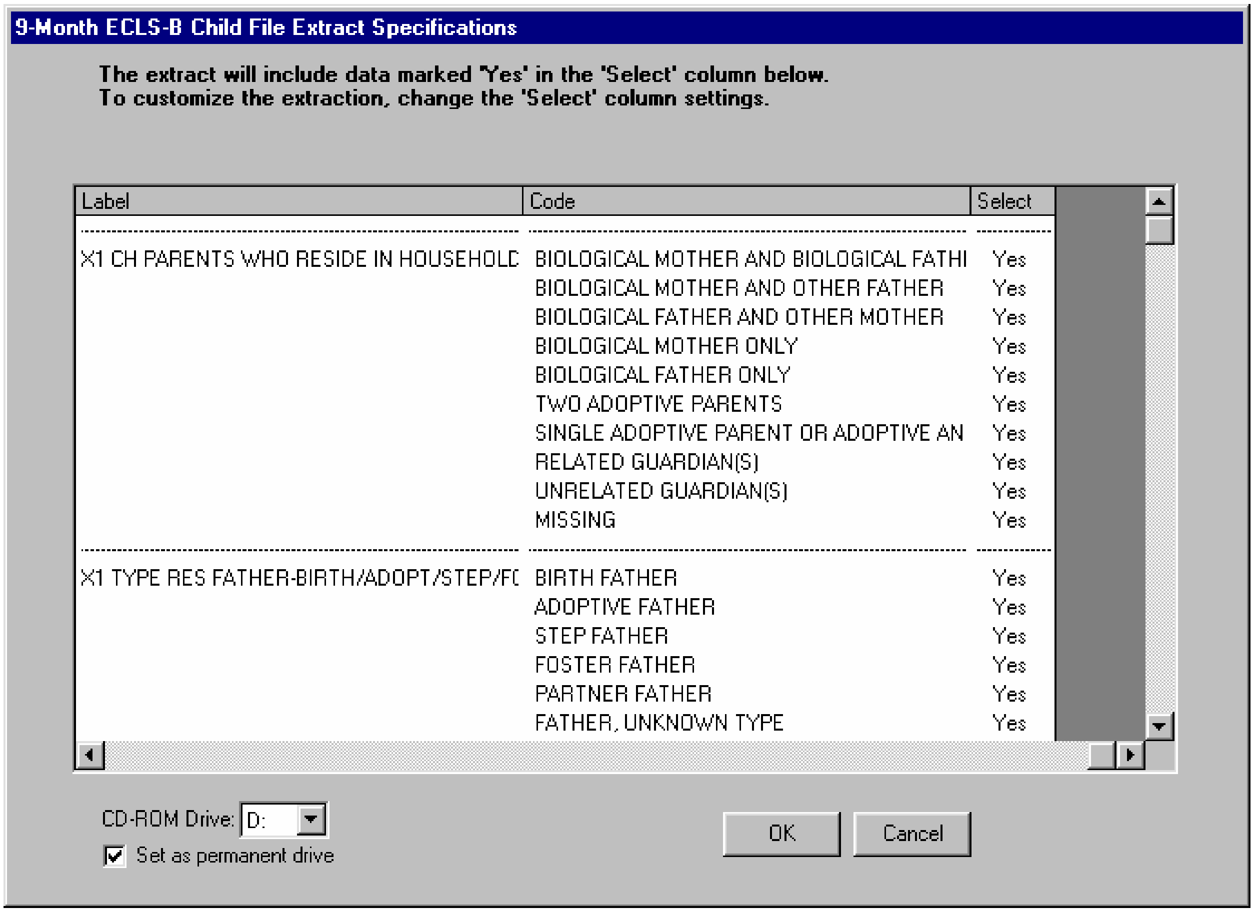


Exhibit 8-63 is the main screen of the 9-month ECLS-B ECB.

Exhibit 8-63. 9-month ECLS-B ECB main screen