NCES
National Center for
Education Statistics

Early Childhood Longitudinal Study
Kindergarten Class of 1998-99

U.S. Department of Education
Institute of Education Sciences
NCES 2006-032

Early Childhood Longitudinal Study, Kindergarten Class of 1998–99 (ECLS-K)

Combined User's Manual for the ECLS-K Fifth-Grade Data Files and Electronic Codebooks

February 2006

Karen Tourangeau
Christine Nord
Thanh Lê
Westat

Judith M. Pollack
Educational Testing Service

Sally Atkins-Burnett
University of Toledo

Elvira Germino Hausken
Project Officer
National Center for
Education Statistics

U.S. Department of Education
Margaret Spellings
Secretary

Institute of Education Sciences
Grover J. Whitehurst
Director

National Center for Education Statistics
Mark Schneider
Commissioner

The National Center for Education Statistics (NCES) is the primary federal entity for collecting, analyzing, and reporting data related to education in the United States and other nations. It fulfills a congressional mandate to collect, collate, analyze, and report full and complete statistics on the condition of education in the United States; conduct and publish reports and specialized analyses of the meaning and significance of such statistics; assist state and local education agencies in improving their statistical systems; and review and report on education activities in foreign countries.

NCES activities are designed to address high-priority education data needs; provide consistent, reliable, complete, and accurate indicators of education status and trends; and report timely, useful, and high-quality data to the U.S. Department of Education, the Congress, the states, other education policymakers, practitioners, data users, and the general public. Unless specifically noted, all information contained herein is in the public domain.

We strive to make our products available in a variety of formats and in language that is appropriate to a variety of audiences. You, as our customer, are the best judge of our success in communicating information effectively. If you have any comments or suggestions about this or any other NCES product or report, we would like to hear from you. Please direct your comments to

National Center for Education Statistics
Institute of Education Sciences
U.S. Department of Education
1990 K Street NW
Washington, DC 20006-5651

February 2006

The NCES World Wide Web Home Page address is http://nces.ed.gov.
The NCES World Wide Web Electronic Catalog is http://nces.ed.gov/pubsearch.

Suggested Citation
Tourangeau, K., Nord, C., Lê T., Pollack, J.M., and Atkins-Burnett, S. (2006). *Early Childhood Longitudinal Study, Kindergarten Class of 1998–99 (ECLS-K), Combined User's Manual for the ECLS-K Fifth-Grade Data Files and Electronic Codebooks* (NCES 2006–032). U.S. Department of Education. Washington, DC: National Center for Education Statistics.

Ordering Information
This publication is only available online at http://www.nces.ed.gov/ecls.

Content Contact:
Elvira Germino Hausken
(202) 502-7352
Elvira.Hausken@ed.gov

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)

TABLE OF CONTENTS (continued)

List of Tables

TABLE OF CONTENTS (continued)

List of Tables (continued)

TABLE OF CONTENTS (continued)

List of Tables (continued)

TABLE OF CONTENTS (continued)

List of Tables (continued)

List of Tables (continued)

TABLE OF CONTENTS (continued)

List of Tables (continued)

List of Tables (continued)

TABLE OF CONTENTS (continued)

List of Tables (continued)

TABLE OF CONTENTS (continued)

List of Tables (continued)

TABLE OF CONTENTS (continued)

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)

GETTING STARTED

This chapter highlights key information needed to work with the Early Childhood Longitudinal Study, Kindergarten Class of 1998–99 (ECLS-K) data and points users to the appropriate sections of this manual to get started quickly. For additional information about any particular topic, users should go to the indicated section of this manual, hereafter referred to as the User's Manual. In this chapter, major differences between the fifth-grade data collection and previous rounds are summarized; cautions and caveats about using the data are provided; and basic information about using the Electronic Codebook is summarized.

As described in section 1.4 of chapter 1, there are three data files available for analyzing the fifth-grade data: a restricted-use file, a public-use file, and a kindergarten–fifth grade longitudinal file. This manual serves as a guide for users of all three of these files. Most of the chapters apply to all three files, but there are a few chapters that apply to only one or two of them. Exhibit 1 summarizes the sections that do not apply to all three files and indicates the data set or sets to which they apply. The user should watch for notices (► *Please note…)* at the beginning of sections that indicate if a section does not apply to all three data sets.

Exhibit 1. Sections of User's Manual that do not apply to all three data files

Section	Description	Data set to which section applies
► *Please note …*		
7.8: table 7-15	Composite table	The last two columns of table 7-15 contain information that is file specific. The second-to-last column in table 7-15 contains information for the restricted-use file. Information for the public-use and the K–5 longitudinal files is contained in the last column of table 7-15.
7.9	Masked variables	Fifth-grade public-use and kindergarten–fifth grade longitudinal files
9.4	Merging base year, first-, third- and fifth-grade data	Fifth-grade public-use and restricted-use files
10	Longitudinal kindergarten–fifth-grade public-use file	This chapter applies to users of the K–5 longitudinal file that NCES releases.

In preparing public-use files, the National Center for Education Statistics (NCES) takes steps to minimize the likelihood that an individual school, teacher, parent, or child participating in the study can be identified. Every effort is made to protect the identity of individual respondents. Some modifications to the data contained in the restricted-use file have been made to the public-use file to ensure confidentiality. These modifications do not affect the overall data quality and most researchers should be able to find all data needed for analysis in the public-use files. Chapter 1, Section 1.4.1, provides a general description of the differences between the public-use and restricted-use files. Table 7-16 in Chapter 7 contains a list of the variables that have been modified. Section 7.9 contains additional information about the "masking" process.

Major Differences in the Fifth-Grade Data Collection

Although the fifth-grade data collection shares many similarities with earlier rounds, some modifications were made to capture important information relevant to fifth-grade students. For example, to capture information about students with learning and other disabilities which are often diagnosed in elementary school, questions were added about when diagnoses for specific disabilities were made. In addition, because by fifth grade there is more specialization in subject matter taught by teachers, the approach to collecting information from teachers was modified. Below the major differences between the fifth-grade data collection and the earlier rounds are summarized:

- **New construct areas were added to the parent interview for fifth grade. The new areas include the following:**
  - A series of questions about when a diagnosis had been made for specific disabilities, such as those related to learning or paying attention (e.g., learning disability, dyslexia, attention deficit disorder [ADD], attention deficit hyperactivity disorder [ADHD], developmental delay, autism or pervasive developmental delay, mental retardation) or those related to vision, hearing, or emotional problems;
  - For children with cochlear implants, a question about when cochlear implants were implanted;
  - A question on the use of cochlear implants in school;
  - A question that identified when a child's use of therapy services or a program for children with disabilities ended;

- A question about why the child no longer participated in services for children with special needs or special education;

- Questions about medications taken for ADD, ADHD, or hyperactivity. Parents of children taking medication were asked about the type of medication taken and the length of time that medication was taken;

- A series of questions on the receipt of family therapy, reasons for family therapy, type of therapist seen, and number of times the family saw a therapist; and

- Questions about discussions the parent has had with the child about school and friends and about smoking, sexual activity, and the use of alcohol and drugs.

■ **Science was a separate assessment domain.** In the base year and in spring-first grade, the direct cognitive assessment included a general knowledge assessment that measured children's knowledge of the social and physical worlds. In third grade and in fifth grade, children's knowledge of the world is more finely categorized into science and social studies domains. With limited time available for the direct assessment, the third- and fifth-grade assessments included only the science domain. Sections 2.1 and 3.1 provide information on the direct cognitive assessments.

■ **Information about children's food consumption was collected from the children and also from the school administrators.** Assessors administered a food consumption questionnaire (FCQ) to children to determine the kinds of food they could buy at school and the food they had eaten in the past week. See section 2.1.4 for information on the FCQ. In addition, a new content area on food consumption was added to the school administrator questionnaire. School administrators were asked a series of questions, including the types of food that children could buy at the school; where the foods could be obtained (e.g., a school store or vending machine); and how full the cafeteria was at peak meal times. See section 2.5 for information about the food consumption questions in the school administrator questionnaire.

■ **Children rated their perceptions of social skills and interest in school subjects.** In the kindergarten and first-grade rounds of the ECLS-K, parents and teachers reported about children's social skills. In the fifth grade, as in third grade of the ECLS-K, the children provided information about themselves by completing a short self-description questionnaire (SDQ). See sections 2.1.1 and 3.4 for additional information on the SDQ.

■ **Social Rating Scale (SRS) was collected from teachers.** In the base year and in spring-first grade, parents and teachers completed the Social Rating Scale, which measures children's approaches to learning, self-control, interpersonal skills, and peer relations. In spring-third and in spring-fifth grade, only teachers completed this scale. Sections 2.3.2 and 3.3 provide information about the SRS.

■ **In fifth grade, a different approach from previous rounds was used to collect information from teachers.** The approach for administering teacher questionnaires differed from that of previous rounds because many fifth-grade children were expected to have different teachers for different subject areas. In earlier rounds, all questions pertaining to the core academic subjects were asked in a single questionnaire and given to teachers who had sample children in their homeroom class. In the fifth grade, however, separate questionnaires were given to sample children's reading/language arts, mathematics, and science teachers. Teacher questionnaire content changes are described in section 2.3. Information about how to use the variables in the teacher questionnaire is presented in section 7.2.

- **Questions that were in the school fact sheet during the third-grade data collection were included in the school administrator questionnaire.** In fifth grade, questions previously asked in the third-grade school fact sheet were moved to the school administrator questionnaire to reduce the number of forms left with the schools. Items previously asked in the third-grade school fact sheet (e.g., basic information about the school including grade level, school type [public or private], length of school year, and attendance recordkeeping practices) were incorporated into the school administrator questionnaire for fifth-grade data collection (section 2.5). Prior to the third-grade data collection, the questions were part of the school administrator questionnaire. Only during the third-grade data collection were they in a separate form.

Cautions and Caveats

Users of previous rounds of the ECLS-K data have frequently asked certain questions. For example, can school-level and teacher-level estimates be made with the ECLS-K data? Or, did the ECLS-K sample whole classrooms? 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:

- **Sample is not representative of fifth-grade students, classrooms, or schools.** The ECLS-K base year sample is a representative sample of children attending kindergarten during the 1998–99 school year, of schools with kindergartens, and of kindergarten teachers. Because the first-grade sample was freshened with students who had not attended kindergarten in the United States in the previous year, the first-grade sample is representative of children attending first grades in the United States during the 1999–2000 school year. However, it is not representative of schools with first grades or of first-grade teachers. The fifth-grade sample is not representative of fifth-grade students, fifth-grade teachers, or schools with fifth grades. Children who started their schooling in the U.S. in second, third, fourth, or fifth grade are not represented in the sample. The data should not be used to make statements about fifth-grade students, schools with fifth grades, or fifth-grade teachers.

- **Not all sample children are in fifth grade.** The fifth-grade data file includes children who were in fifth grade in spring 2004, and others who were either held back (e.g., fourth-graders) or promoted ahead an extra year or more (e.g., sixth-graders). Users need to be aware of this fact when using the data and interpreting the findings. Most children in the sample have been in school for at least 6 years (K-5) and some more (those who were repeating K in the base year). A very small number may have been in school less than 6 years (some part of the freshened sample added in first grade).

- **Student mobility and its consequences.** A random subsample of students who transferred from their base year schools was flagged to be followed in fall-first grade and in subsequent rounds of data collection. Sections 4.3.1, 4.4.1, and 4.5 describe the subsampling of movers. There are a number of variables on the file that can be used to determine if a child moved to a different school between rounds or moved to a different school during the fifth-grade data collection. Section 7.7 describes these variables. Student mobility has a number of consequences for the ECLS-K. It results in a reduction in sample size, fewer children per school, and more missing school and teacher questionnaire data for movers. See section 5.7, tables 5-12 and 5-13 for more information on the response rates for movers and nonmovers.

- **Missing data.** Users should be certain to recode any missing data properly before conducting analyses. If the user is analyzing data over time, it is especially important to check that all skip patterns are the same across years because some changed between rounds of data collection. There are 5 different possible missing data codes on the file. See section 7.3 for a discussion of the different missing values codes and the circumstances when they are used.

- **Rescaled scores.** The longitudinal scales necessary for measuring gain over time were developed by pooling all rounds of item response data, from fall-kindergarten through spring-fifth grade. Scale scores reported in each successive round were based on all test items present in the assessments up to and including that round. Each time the item pool was expanded, scores were recalibrated for *all* rounds to make longitudinal comparisons possible. Each recalibration of the scale score represents the estimated number right on a larger and larger set of items. As a result, the scale score for the *same* child in the *same* grade changes each time a new set of test items is incorporated and the scale on which the score is based is expanded. Estimates of gains in scale score points should be made using the recalibrated versions for all rounds. It would be inappropriate to compare previously reported scale score means with means based on recalibrated scores in the fifth-grade data file because the set of items on which the score is based has changed. This caveat applies primarily to analyses that report gains in scale score points. The effect of rescaling on previously-reported T-scores and proficiency probability scores should be relatively small. However, to the extent that the pooling of test items across rounds represents a redefinition of the construct being measured, slight differences in these statistics may be observed as well. See the *ECLS-K Psychometric Report for the Fifth Grade* (NCES 2006–036) (Pollack et al. 2005) for more information.

- **Use of weights.** The fifth-grade data file contains 5 sets of cross-sectional weights and 10 longitudinal (panel) weights. Although there are a variety of weights on 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 sections 4.7 and 9.3.1.

- **Defining special populations.** The ECLS-K includes a number of analytic groups of interest that can be identified and studied separately. For example, the fifth-grade file contains variables that identify children who have a disability diagnosed by a professional (P6DISABL), children receiving nonparental child care (P6CARNOW), and those who live in households with incomes below the federal poverty threshold (W5POVRTY). With variables from earlier rounds of data collection, it is possible to identify children who participated in Head Start in the year prior to kindergarten (HSATTEND from the base year and P4HSBEFK asked of new respondents in spring-first grade) and language minority children (WKLANGST), as well as other subgroups. While these variables are not contained on the fifth-grade cross-sectional data file, they are available on the K–3 and K–5 longitudinal files. Users who desire to study a specific subpopulation should search the Electronic Codebook using the "NARROW" feature of the Electronic Codebook to list variables that might help them identify their population of interest. See section 8.3.1 for a description of this feature.

- **Examining school and classroom effects.** When studying the effects of school and classrooms, it is important to restrict the analytic sample to children in the same classroom and/or same schools. Each type of respondent (child, parent, regular teacher, special education teacher, and school) has a unique ID number. These ID numbers can be used to identify children in the same classrooms and schools. Section 7.1 describes the available identification variables.

- **Date of assessments and elapsed times between assessments are not the same for all children.** The Electronic Codebook contains variables that indicate the month, day, and year in which the direct assessment was administered. The Electronic Codebook also contains composite variables for children's age at assessment for each sampled child. See the NCES web site http://nces.ed.gov/ecls for information on how to calculate the elapsed time period between two assessments.

- **Measuring achievement gains.** One of the major strengths of the ECLS-K is the ability to measure children's achievement gains as they progress from kindergarten through the early elementary grades. There are several different approaches to measuring gains. See section 3.1.6 for a discussion of measuring gains with the ECLS-K.

Electronic Codebook Reference Guide

- **Electronic Codebook (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, SPSS for Windows, or Stata programs for extracting selected variables (see section 8.1.2 for a description of the ECB features). Section 8.2 of the User's Manual contains detailed instructions on how to install and open the ECB. The ECB allows users to easily examine the variables in the ECLS-K 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. This text data file is referred to as the "child catalog" and is named child5p.dat in the CD-ROM root directory. For more information about the data file, see Appendix E on the CD-ROM. The ECB CD-ROM also contains Portable Document Format (PDF) files of the associated questionnaires and of the User's Manual. Users of prior versions of the ECLS-K ECB should note that some minor changes have been made to the online Help feature in the fifth grade ECB; see section 8.6 for more information.

- **Data file.** The fifth-grade child catalog contains one record for each of 11,820 responding students in spring-fifth grade. Data collected from teachers and schools are stored in the child catalog. The file, named child5p.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. Appendix B on the CD-ROM contains the record layout for the child catalog.

- **Identification variables.** The fifth-grade data file contains a child identification variable (CHILDID) that uniquely identifies each record. The same ID is used in each round of the survey. Teachers on the child records are identified with ID variables J61T_ID (reading teacher ID) and J62T_ID (mathematics or science teacher ID); schools are identified by the ID variables S6_ID. See sections 7.1 and 7.7 in the User's Manual for further information on these identification variables.

- **Instruments.** For the ECLS-K fifth-grade data collection, data were collected using computer-assisted interviewing for parent interviews and child assessments. Self-administered questionnaires in paper/pencil format were used to collect information from teachers and school administrators or their designees. Chapter 2 of the User's Manual provides an overview of the instruments. To help decide what variables to use in analyses, the user should always review the actual instruments. Seeing the specific wording of the questions and the context in which they are asked is useful in understanding the results of the user's analyses and can help minimize errors. Appendix A on the ECLS-K ECB CD-ROM contains, with some exceptions, the fifth-grade instruments. The exceptions are measures that contain copyright-protected materials and instruments covered by agreements with the test publishers that restrict distribution.

- **Composite variables.** Numerous composites have been constructed for the ECLS-K data to make it easier for users to use the data set. Most composite variables were created using two or more variables that are on the data file or using information from other sources. Others are recodes of single variables. Composites based on the child assessment include height, weight, and body mass index (BMI). Composites based on the teacher data include class size, percentage of limited-English-proficient children in the class, and student grade level. Composites based on the school data include the percentage of minority students, school type, and school instructional level.

Composites based on the parent data include parent education, poverty status, and socio-economic status. See section 7.5 and table 7-15 of the User's Manual for details on all the composites contained on the fifth-grade public-use data file. It is strongly recommended that users consider using the composite variables in their analysis, as appropriate. These variables represent the compilation of study data, including data from sources not otherwise available on the data file.

- **Assessment scales.** A key feature of the ECLS-K data is the set of assessments administered to each child. These assessments included direct and indirect cognitive assessments and measures of children's social development. Chapter 2 provides a general description of the survey instruments, including the direct and indirect assessments. The fifth-grade direct cognitive assessment contained items in reading, mathematics, and science. See section 3.1 of the User's Manual for details on the direct cognitive assessment and the scores that are available for analysis. Section 3.1.5 of the User's Manual discusses choosing the appropriate score for analysis. Section 3.1.6 discusses approaches to measuring student gains in achievement. The indirect cognitive assessment consisted of the Academic Rating Scale (ARS), which was developed for the ECLS-K to measure teachers' evaluations of students' academic achievement in three domains: language and literacy (reading and writing), science, and mathematical thinking. See section 3.2 of the User's Manual for more information on the ARS.

  The measures of children's social development consisted of the Teacher Social Rating Scale (SRS), which asked fifth-grade teachers to report how often students exhibited certain social skills and behaviors, and a Self-Description Questionnaire (SDQ) in which the students rated their own perceptions of competence and interest in reading, mathematics, and all school subjects. In the SDQ, children also rated their competence and popularity with peers and reported problem behaviors. See sections 3.3 and 3.4 for more information on the SRS and SDQ and the scores that are available for analysis.

- **Sample design and weights.** The ECLS-K employs a complex sample design. See chapter 4 for a description of the sample design. In order to obtain accurate estimates, the user will need to select the appropriate weights. Section 4.7 describes the cross-sectional weights and provides advice for which weight to use for a given type of analysis. See exhibit 4-1 for a summary of the cross-sectional weights available for analysis. A description of the longitudinal weights is provided in chapter 9. Section 9.3.1 describes the K–5 longitudinal (panel) weights and provides advice for which panel weight to use for a given type of analysis. See exhibit 9-1 for a summary of the K–5 longitudinal (panel) weights.

- **Creating a longitudinal file**. It is possible to combine the fifth-grade data with data from earlier rounds. Instructions on how to create such a file are provided in chapter 9. Most users, however, will probably want to wait for the release of the public-use longitudinal data set. This data set will be available in 2006.

1. INTRODUCTION

This manual provides guidance and documentation for users of the fifth-grade data[1] of the Early Childhood Longitudinal Study, Kindergarten Class of 1998–99 (ECLS-K). It begins with an overview of the ECLS-K 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 of the data file.

The ECLS-K focuses on children’s early school experiences beginning with kindergarten. It is a multisource, multimethod study that includes interviews with parents, the collection of data from principals and teachers, and student records abstracts, as well as direct child assessments. The ECLS-K has been developed under the sponsorship of the U.S. Department of Education, National Center for Education Statistics (NCES). Westat is conducting this study with assistance provided by Educational Testing Service (ETS) in Princeton, New Jersey.

The ECLS-K is following a nationally representative cohort of children from kindergarten into high school. The base year data were collected in the fall and spring of the 1998–99 school year when the sampled children were in kindergarten. A total of 21,260 kindergartners throughout the nation participated.

Two more waves of data were collected in the fall and spring of the 1999–2000 school year when most, but not all, of the base year children were in first grade[2]. The fall-first grade data collection was limited to a 30 percent subsample of schools[3] (see exhibit 1-1). It was a design enhancement to enable researchers to measure the extent of summer learning loss and the factors that contribute to such loss and to better disentangle school and home effects on children’s learning. The spring-first grade data collection, on the full sample, was part of the original study design and can be used to measure annual school progress and to describe the first-grade learning environment of children in the study. All children assessed during the base year were eligible to be assessed in the spring-first grade data collection regardless of whether they repeated kindergarten, were promoted to first grade, or were promoted to second grade. In addition, children who were not in kindergarten in the United States during the 1998–99

[1] The term “fifth grade” is used throughout this document to refer to the data collections that took place in the 2003–04 school year, at which time most of the sampled children—but not all of them—were in fifth grade.
[2] Though the majority of base year children were in first grade during the 1999–2000 school year, about 5 percent of the sampled children were retained in kindergarten and a handful of others were in second grade during the 1999–2000 school year.
[3] Approximately 27 percent of the base year students who were eligible to participate in year 2 attended the 30 percent subsample of schools.

school year, and therefore did not have a chance to be selected to participate in the base year of the ECLS-K, were added to the spring-first grade sample.[4] Such children include immigrants to the United States who arrived after fall 1998 sampling, children living abroad during the 1998–99 school year, children who were in first grade in 1998–99 and repeated it in 1999–2000, and children who did not attend kindergarten. Their addition allows researchers to make estimates for all first-graders in the United States rather than just for those who attended kindergarten in the United States in the previous year.

A fifth wave of data was collected in the spring of the 2001–02 school year when most, but not all, of the sampled children were in third grade.[5] In addition to the school, teacher, parent, and child assessment data collection components, children were asked to complete a short self-description questionnaire, which asked them how they thought and felt about themselves both socially and academically. The spring-third grade data collection can be used to measure school progress and to describe the third-grade learning environment of children in the study.

Exhibit 1-1. ECLS-K waves of data collection: School years 1998–99, 1999–2000, 2001–02, and 2003–04

Data collection	Date of collection	Sample
Fall-kindergarten	Fall 1998	Full sample
Spring-kindergarten	Spring 1999	Full sample
Fall-first grade	Fall 1999	30 percent subsample[1]
Spring-first grade	Spring 2000	Full sample plus freshening[2]
Spring-third grade	Spring 2002	Full sample
Spring-fifth grade	Spring 2004	Full sample

[1] Fall data collection consisted of a 30 percent sample of schools containing approximately 27 percent of the base year students eligible to participate in year 2.
[2] See description of freshened sample in text preceding exhibit 1-1.
NOTE: See section 1.3 for a description of the study components.
SOURCE: U.S. Department of Education, National Center for Education Statistics, Early Childhood Longitudinal Study, Kindergarten Class of 1998–99 (ECLS-K), spring 2004.

A sixth wave of data was collected in the spring of the 2003–04 school year when most, but not all, of the sampled children were in fifth grade.[6] In addition to the school, teacher, parent, and child

[4] Their addition is referred to as "freshening" the sample. See chapter 4, section 4.3.2 for more detail on the freshening process.
[5] Approximately 89 percent of the children interviewed were in third grade during the 2001–02 school year, 9 percent were in second grade, and less than 1 percent were in fourth grade or higher.
[6] Approximately 90 percent of the children interviewed were in fifth grade during the 2003–04 school year, 9 percent were in fourth grade, and less than 1 percent were in third or some other grade such as second grade or sixth grade.

assessment data collection components, children were asked to complete a short self-description questionnaire, which asked them how they thought and felt about themselves both socially and academically. They were also asked about their food consumption at school and in the week prior to the interview. The spring-fifth grade data collection can be used to measure school progress and to describe the fifth-grade learning environment of children in the study.

The sample of children in the fifth-grade round of data collection of the ECLS-K represents the cohort of children who were in kindergarten in 1998–99 or in first grade in 1999–2000. Since the sample was not freshened after the first-grade year with third or fifth-graders who did not have a chance to be sampled in kindergarten or first grade (as was done in first grade), estimates from the ECLS-K third- and fifth-grade data are representative of the population cohort rather than all third-graders in 2001–02 or all fifth-graders in 2003–04. The estimated number of third-graders from the third-grade ECLS-K data collection is approximately 86 percent of all third-graders. From the fifth-grade ECLS-K data collection, the estimated number of fifth-graders is approximately 83 percent of all fifth-graders. While the vast majority of children in third grade in the 2001–02 school year and in fifth grade in the 2003–04 school year are members of the cohort, third-graders who repeated second or third grade, fifth graders who repeated third or fourth grade, and recent immigrants are not covered. Data were collected from teachers and schools to provide important contextual information about the school environment for the sampled children. The teachers and schools are not representative of fifth-grade teachers and schools in the country in 2003–04. For this reason, the only weights produced from the study are for making statements about children, including statements about the teachers and schools of those children.

The ECLS-K has several major objectives and numerous potential applications. The ECLS-K combines (1) a study of achievement in the elementary years; (2) an assessment of the developmental status of children in the United States at the start of their formal schooling and at key points during the elementary school years; (3) cross-sectional studies of the nature and quality of kindergarten programs in the United States; and (4) a study of the relationship of family, preschool, and school experiences to children's developmental status at school entry and their progress during the kindergarten and early elementary school years.

The ECLS-K is part of a longitudinal studies program comprising two cohorts—a kindergarten cohort and a birth cohort. The birth cohort (ECLS-B) is following a national sample of children born in the year 2001 from birth to kindergarten. The ECLS-B examines how early learning environments are associated with early cognitive, physical, and socioemotional development and thus prepare children for kindergarten success. Together these cohorts will provide the depth and breadth of data required to more fully describe and understand children's early learning, development, and education experiences.

The ECLS-K has both descriptive and analytic purposes. It provides descriptive data on children's status at school entry, their transition into school, and their progress into high school. The ECLS-K also provides a rich data set that enables researchers to analyze how a wide range of family, school, community, and individual variables affect children's early success in school; to explore school readiness and the relationship between the kindergarten experience and later elementary school performance; and to record children's academic growth as they move through secondary school.

1.1 Background

Efforts to expand and improve early education will benefit from insights gained through analyses of data from the large-scale, nationally representative ECLS-K data and the study's longitudinal design. The ECLS-K database contains information about the types of school programs in which children participate, the services they receive, and repeated measures of the children's cognitive skills and knowledge. The ECLS-K database also contains measures of children's physical health and growth, social development, and emotional well-being, along with information on family background and the educational quality of their home environments.

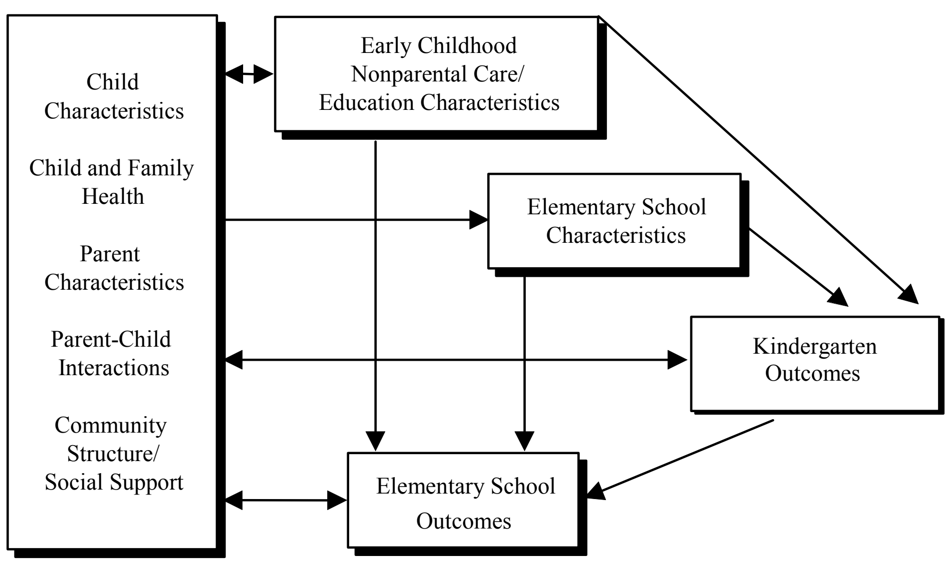
As a study of early achievement, the ECLS-K allows researchers to examine how children's progress is associated with such factors as placement in high or low ability groups, receipt of special services or remedial instruction, grade retention, and frequent changes in schools attended because of family moves. Data on these early school experiences are collected as they occur, with the exception of their experiences before kindergarten, which are collected retrospectively. Collecting this information as it occurs produces a more accurate measurement of antecedent factors and enables inferences to be made about their relationship to later academic progress. The longitudinal nature of the study enables researchers to study children's cognitive, social, and emotional growth and to relate trajectories of change to variations in children's experiences in kindergarten and the early grades to later grades.

The spring-fifth-grade data collection can be used to describe the diversity of the children in the study and the classrooms and schools they attend. It can also be used to study children's academic gains in the years following kindergarten. The ECLS-K sample includes substantial numbers of children from various minority groups. Thus, the ECLS-K data present many possibilities for studying cultural and ethnic differences in the educational preferences, home learning practices, and school involvement of families; the developmental patterns and learning styles of children; and the educational resources and opportunities that different groups are afforded in the United States.

1.2 Conceptual Model

The design of the ECLS-K has been guided by a framework of children's development and schooling that emphasizes the interrelationships between the child and family; the child and school; the family and school; and the family, school, and community. The ECLS-K recognizes the importance of factors that represent the child's health status and socioemotional and intellectual development and incorporates factors from the child's family, community, and school-classroom environments. The conceptual model is presented in exhibit 1-2. The study has paid particular attention to the role that parents and families play in helping children adjust to formal school and in supporting their education through the elementary grades. It has also gathered information on how schools prepare for and respond to the diverse backgrounds and experiences of the children and families they serve.

Exhibit 1-2. ECLS-K conceptual model



SOURCE: U.S. Department of Education, National Center for Education Statistics, Early Childhood Longitudinal Study, Kindergarten Class of 1998–99 (ECLS-K), fall 1998.

1.3 Study Components

The emphasis placed on measuring children's environments and development broadly has critical implications for the design of the ECLS-K. The design of the study includes the collection of data from the child, the child's parents/guardians, teachers, and schools.

- **Children** participate in various activities to measure the extent to which they exhibit those abilities and skills deemed important to success in school. They are asked to participate in activities designed to measure important cognitive (i.e., literacy, quantitative, and science) and noncognitive (i.e., fine motor and gross motor coordination and socioemotional) skills and knowledge. Most measures of a child's cognitive skills are obtained through an untimed one-on-one assessment of the child. Beginning with the third-grade data collection, children report on their own perceptions of their abilities and achievement as well as their interest in and enjoyment of reading, math, and other school subjects. Children are assessed in each round of data collection.

- **Parents/Guardians** are an important source of information about the families of the children selected for the study and about themselves. Parents provide information about children's development at school entry and their experiences both with family members and others. Information is collected from parents each time children are assessed using computer-assisted interviews (CAIs). Information is collected from parents/guardians in each round of data collection.

- **Teachers**, like parents, represent a valuable source of information on themselves, the children in their classrooms, and the children's learning environment (i.e., the classroom). Teachers are not only asked to provide information about their own backgrounds, teaching practices, and experience; they are also called on to provide information on the classroom setting for the sampled children they teach and to evaluate each sampled child on a number of critical cognitive and noncognitive dimensions. Special education teachers and service providers of sampled children with disabilities are also asked to provide information on the nature and types of services provided to the child. With the exception of the fall-first grade data collection, teachers complete self-administered questionnaires each time children are assessed.

- **School Administrators**, or their designees, are asked to provide information on the physical, organizational, and fiscal characteristics of their schools, and on the schools' learning environment and programs. Special attention is paid to the instructional philosophy of the school and its expectations for students. School administrators or their designees are also asked to provide basic information about the school grade level, school type (public or private), length of school year, and attendance recordkeeping practices. Prior to the third-grade data collection, the questions had been part of the school administrator questionnaire. These items were collected in a separate school fact sheet in third grade, but were reintegrated into the school administrator questionnaire in the fifth-grade data collection. Information is collected from school administrators via self-administered questionnaires during each spring data collection.

- **School Office Staff** are asked to complete a student records abstract form and provide basic information about the school. The student records abstract form includes questions about an individual child's enrollment and attendance at the school, transfer to another school (if applicable), and verifies whether the child has an Individualized Education Program (IEP) on record. A student records abstract form is completed for each child in the study during each spring data collection.

1.4 ECLS-K Data Files

The ECLS-K 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 here, followed by a description of the data files that are currently available.

1.4.1 Differences Between ECLS-K Restricted-Use and Public-Use Files

In preparing the public-use files, NCES takes steps to minimize the likelihood that an individual school, teacher, parent, or child participating in the study can be identified. Every effort is made to protect the identity of individual respondents. This is in compliance with the Privacy Act of 1974, as amended, the E-Government Act of 2002, the Education Sciences Reform Act of 2002, and the USA Patriot Act of 2001, which mandate the protection of confidentiality of NCES data that contain individually identifiable information. The process begins with a formal disclosure risk analysis. Variables identified as posing the greatest disclosure risk are altered (e.g., by combining categories), and in some instances, entirely suppressed.

The following data modifications account for the differences between the public-use and restricted-use data files:

- Outlier values are top- or bottom-coded;[7]
- Individual cases for which a particular variable poses an especially high risk of disclosure have the value of that variable altered (usually by no more than 5 to 10 percent for continuous variables) to reduce the risk;
- Some continuous variables are modified into categorical variables, and categories of certain categorical variables are collapsed;
- A small number of variables with too few cases and a sparse distribution are suppressed altogether, rather than modified; and
- A small number of variables are further masked to enhance confidentiality.

After modifying individual records that have the greatest risk of disclosure, the disclosure risk analysis is repeated to verify that the risk of disclosure has been reduced to acceptable levels. The modifications that are implemented to avoid identification of schools, teachers, parents, and children do not affect the overall data quality and most researchers should be able to find all that they need in the public-use files. While very few of the variables are suppressed, there are a few users who might require the restricted files. Those researchers examining certain rare subpopulations, such as children with disabilities, or children with specific non-English home languages or countries of birth, for example, will

[7] To understand top- and bottom-coding, consider a fictitious variable with the following frequency distribution:

Variable X frequency distribution

Value	Count	Percent
Total	4,641	100.00
0	45	0.97
1	193	4.16
2	2,846	61.32
3	1,318	28.40
4	220	4.74
5	18	0.39
6	1	0.02

The outlier values are 0, 5, and 6. Values 0 and 1 are bottom-coded and values 4, 5, and 6 are top-coded. The resulting masked variable has the following frequency:

Masked variable X frequency distribution

Value	Count	Percent
Total	4,641	100.00
≤ 1	238	5.13
2	2,846	61.32
3	1,318	28.40
≥ 4	239	5.15

find that the restricted-use files contain a few more variables with a wider range of data values. However, in many instances even though the detailed information on the restricted-use files may be of interest, the sample sizes will be too small to support these analyses. NCES recommends that researchers who are uncertain of which data release to use first examine the public-use files to ascertain whether their specific analytic objectives can be met using those data files.

1.4.2 Overview of Available Data Files

A variety of ECLS-K data files are available for use by analysts. These are described below beginning with the fifth-grade data files.

- **ECLS-K Fifth-Grade Restricted- and Public-Use Data Files.** The fifth-grade data are available only as a child-level file. The file includes all data collected from or about the children and their schools including data from the child assessments and from their parents, teachers, and schools. No fifth-grade teacher or school files are released because the sample of teachers and schools is not nationally representative of fifth-grade teachers and schools with fifth grades. Analysts who wish to examine children's experiences in fifth grade and the influence of their classroom or school characteristics on their fifth-grade experiences should use the fifth-grade file.

  The fifth-grade data file not only can be used to analyze data collected in the fifth grade but it also provides weights and variables that can be used in longitudinal data analysis of kindergarten, first, third, and fifth grades. In addition to the cross-sectional weights, cross-year (kindergarten–fifth grade) weights have been added to the fifth-grade data file for those analysts who wish to examine children's learning across school years. Instructions on how to create a longitudinal file using the base year, first-grade, third-grade, and fifth-grade data are provided in chapter 9. A longitudinal public-use file, however, is available that combines the base year, first-grade, third-grade, and fifth-grade data (see next bullet). Most analysts will find it more convenient to use the already created longitudinal file described below.

- **Longitudinal Kindergarten–Fifth Grade (K–fifth grade) Public-Use Data File.** This public-use data file combines data from the base, first-, third-, and fifth-grade years. It contains cross-year weights so that analysts can examine children's growth and development between kindergarten and fifth grade. In order to streamline the file, the household roster that lists all household members, their relationship to the sampled child, and selected other characteristics is not included on this longitudinal file. Instead, composite variables describing the children's family structure and selected characteristics of the family members have been added to the file. Analysts who wish to study children's learning across school years, but who do not require the detailed household roster information, should use the longitudinal file.

- **ECLS-K Third-Grade Restricted- and Public-Use Data Files.** The third-grade data are available only as a child-level file. The file includes all data collected from or about the children and their schools including data from the child assessments and from their parents, teachers, and schools. No third-grade teacher or school files are released because the sample of teachers and schools is not nationally representative of third-grade teachers and schools with third grades. Analysts who wish to examine children's experiences in third grade and the influence of their classroom or school characteristics on their third-grade experiences should use the third-grade file.

  The third-grade data file can be used not only to analyze data collected in the third grade but also to provide weights and variables that can be used in longitudinal data analysis of kindergarten, first grade, and third grade. In addition to the cross-sectional weights, cross-year (kindergarten–third grade) weights have been added to the third-grade data file for those analysts who wish to examine children's learning across school years. Instructions on how to create a longitudinal file using the base year, first-grade, and third-grade data are provided in chapter 9. A longitudinal public-use file, however, is available that combines the base year, first-grade, and third-grade data (see next bullet). Most analysts will find it more convenient to use the already created longitudinal file described below.

- **Longitudinal Kindergarten–Third Grade (K–Third Grade) Public-Use Data File.** This public-use data file combines data from the base, first-grade, and third-grade years. It contains cross-year weights so that analysts can examine children's growth and development between kindergarten and third grade. In order to streamline the file, the household roster that lists all household members, their relationship to the sampled child, and selected other characteristics, is not included on the file. Instead, composite variables describing the children's family structure and selected characteristics of the family members have been added to the file. Analysts who wish to study children's learning across school years, but who do not require the detailed household roster information, should use the longitudinal file. For information about this file, see the *User's Manual for the ECLS-K Longitudinal Kindergarten–Third Grade Public-Use Data File and Electronic Code Book* (NCES 2004–088) (Tourangeau, Nord, et al. 2004).

- **ECLS-K First-Grade Restricted- and Public-Use Data Files.** The first-grade data (fall and spring) are available only as a child-level file. The file includes all data collected from or about the children and their schools including data from the child assessments and from their parents, teacher, and schools. Although these data are freshened to be representative of first graders in the U.S. in 1999–2000, no first-grade teacher or school files are released because the sample of teachers and schools is not nationally representative of *first-grade teachers and schools with first grades.* Analysts who wish to examine children's experiences in first grade and the influence of their classroom or school characteristics on their first grade experiences should use the first-grade file.

  The first-grade data file can be used not only to analyze data collected in the first grade but also to provide weights and variables that can be used in longitudinal data analysis of both kindergarten and first grade. In addition to the cross-sectional weights, cross-year (kindergarten–first grade) weights have been added to the first-grade data file for those analysts who wish to examine children's learning across school years. A longitudinal public-use file, however, is available that combines the base year and first-grade data (see next bullet). Most analysts will find it more convenient to use the already created longitudinal file described below. For more information about the first-grade file, see the *User's Manual for the ECLS-K First Grade Public-Use Data Files and Electronic Codebook* (NCES 2002–135) (Tourangeau, Burke, et al. 2002).

- **Longitudinal Kindergarten–First Grade (K-First Grade) Public-Use Data File.** This public-use data file combines data from the base and first-grade years. It contains cross-year weights so that analysts can examine children's growth and development between kindergarten and first grade. In order to streamline the file, the household roster that lists all household members, their relationship to the sampled child, and selected other characteristics is not included on the file. Instead, composite variables describing the children's family structure and selected characteristics of the family members have been added to the file. Analysts who wish to study children's learning across school years or to study the extent of summer learning loss between kindergarten and the fall of the following school year, but who do not require the detailed household roster information, should use the longitudinal file. For information about this file, see the *User's Manual for the ECLS-K Longitudinal Kindergarten–First Grade Public-Use Data File and Electronic Codebook* (NCES 2002–149) (Tourangeau, Nord, et al. 2002).

- **ECLS-K Base Year Data Files.** There are three main and four supplementary files available for the base year. The three main files are the child-level file, the teacher-level file, and the school-level file. The supplementary files are the teacher salary and benefits file, the special education file, the student record abstract file, and the Head Start Verification Study file.

  The child file contains all the data collected from or about the children, including data from the child assessments, and from their teachers, parents, and schools. Analysts who wish to obtain descriptive information about U.S. kindergarten students or their families, or who want to examine relationships involving children and families, children and teachers, or children and schools, should make use of the child file. Analysts wishing to obtain descriptive information about the population of kindergarten teachers in the United States, or to study relationships involving teachers as the principal focus of attention, should use the teacher file. Analysts who want to obtain descriptive information about public and private schools that contain kindergarten classes, or who want to examine relationships among school characteristics, should make use of the school file. These child-, teacher-, and school-level files are available in public-use and restricted-use versions. For more information on these files, refer to the *ECLS-K Base Year Public-Use Data Files and Electronic Codebook: User's Manual* (NCES 2001–029r) (Tourangeau, Burke, Lê, et al. 2004).

- The **Salary and Benefits File** is collected at the school level and contains information on the base salary, merit pay, and benefit pay of teachers and principals. The salary and benefits data, when combined with other ECLS-K data, can be used to examine, for example, the relationship between student outcomes and school resource allocation and use. This file is only available as a restricted-use file. For more information about this file, see the *ECLS-K Base Year Restricted-Use Salary and Benefits File* (NCES 2001–014) (Tourangeau, Burke, et al. 2001b).

- The **Special Education File** is a child-based file that contains information on 784 children identified as receiving special education or related services in kindergarten. Special education teachers were asked to complete two questionnaires designed to collect information about their professional background and experience and about the nature of the special education program and special education services provided to each of the sampled children receiving services. It is only available as a restricted-use file. For more information about this file, see the *ECLS-K Base Year Restricted-Use Special Education Child File* (NCES 2001–015) (Tourangeau, Burke, et al. 2001c).

- The **Student Records Abstract File** contains information from school records about children's school enrollment and attendance; IEP and disability status; and home and school language. The student records abstract form was completed by school staff after the end of the school year. This file is useful in providing additional predictors and correlates of children's transitions to kindergarten and later progress in school. This file is only available as a restricted-use file. For more information about this file, see the *ECLS-K Base Year Restricted-Use Student Record Abstract File* (NCES 2001–016) (Tourangeau, Burke, et al. 2001d).

- The **Head Start Verification File** contains information from Head Start program providers. The purpose of the **Head Start Verification Study** was twofold: (1) to identify which of the children reported by either their parents or their schools as having attended Head Start the year prior to kindergarten did indeed attend a Head Start program and (2) to evaluate the process of identifying Head Start participation through parent and school reports and provide further information on the actual process of verifying these reports. This file is a restricted-use file. For more information about this file, see the *ECLS-K Base Year Restricted-Use Head Start File* (NCES 2001–025) (Tourangeau, Burke, et al. 2001a). The outcomes of the verification process are also included as data items on the ECLS-K first-grade and kindergarten–first grade longitudinal files.

1.5 Contents of Manual

This manual provides documentation for users of all three fifth-grade data files (fifth-grade restricted-use, public-use, and longitudinal K–5 public-use data files) of the ECLS-K. In previous rounds, separate manuals were issued for each data file. Please refer to the chapter *Getting Started* for a summary of which sections of the manual do not apply to all three files and for an overview of the major differences between the fifth-grade round of data collection and previous rounds.

The manual contains information about the data collection instruments (chapter 2) and the psychometric properties of these instruments (chapter 3). It describes the ECLS-K 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 the structure of the fifth-grade data file and provides definitions of composite variables (chapter 7); describes how to install and use the Electronic Codebook (chapter 8); and describes how to use and merge the base year, first-grade, third-grade, and fifth-grade files (chapter 9). Finally, chapter 10 presents information on the longitudinal kindergarten–fifth grade public-use data file. The Electronic Codebook contains unweighted frequencies for all variables. Because this manual focuses on the fifth-grade data collection, minimal information is provided about the base year, first-grade, or third-grade data. Users who wish to learn more about these data collections should refer to the *ECLS-K Base Year Public-Use Data Files and Electronic Codebook: User's Manual* (NCES 2001–029r) (Tourangeau, Burke, Lê, et al. 2004); the *User's Manual for the ECLS-K First Grade Public-Use Data Files and Electronic Codebook* (NCES 2002–135) (Tourangeau, Burke, et al. 2002) or the *User's Manual for the ECLS-K Third Grade Public-Use Data File and Electronic Code Book* (NCES 2004–001) (Tourangeau, Brick, Lê, et al. 2004). Additional information about the ECLS program can be found on the World Wide Web at http://nces.ed.gov/ecls.

This page is intentionally left blank.

2. DESCRIPTION OF DATA COLLECTION INSTRUMENTS

This chapter describes the survey instruments used during the fifth-grade data collection of the Early Childhood Longitudinal Study, Kindergarten Class of 1998–99 (ECLS-K). Exhibit 2-1 lists all the instruments used during the fifth-grade data collection. The instrumentation for the base year, first-grade, and third-grade data collections are also shown. Similarities and differences between the fifth-grade instruments and those used in the previous rounds are highlighted throughout this chapter.

The ECLS-K fifth-grade data collection occurred in the spring of the 2003–04 school year. Data were collected using computer-assisted interviewing (CAI) for parent interviews and child assessments. As part of the direct child assessments, children completed a short self-description questionnaire on their own and were interviewed using a food consumption questionnaire. Self-administered questionnaires were used to collect information from teachers (teacher questionnaires, special education teacher questionnaires) and school administrators or their designees (school administrator questionnaire and student records abstract). Field staff completed the school facilities check list.

The fifth-grade data collection instruments, with some exceptions, are available on the CD-ROM as appendix A. The exceptions are the direct child assessment, the Social Rating Scale (SRS)[1] in the teacher questionnaire, and the self-description questionnaire (SDQ).[2] These latter measures contain copyright-protected materials and agreements with the test publishers that restrict their distribution.

[1] Adapted with permission from *Social Skills Rating System, Elementary Scale A ("How Often?")*, F.M. Gresham and S.N. Elliott. (1990). Circle Pines, MN: American Guidance Services, Inc.
[2] Adapted with permission from *Self-Description Questionnaire I,* H.W. Marsh. (1992). Campbelltown, N.S.W.: Australia: University of Western Sydney, Macarthur.

Exhibit 2-1. Instruments used in the ECLS-K, by round of data collection: School years 1998–99, 1999–2000, 2001–02, and 2003–04

Instruments	1998–99 school year		1999–2000 school year		2001–02 school year	2003–04 school year
	Fall-kindergarten	Spring-kindergarten	Fall-first grade[1]	Spring-first grade[2]	Spring-third grade	Spring-fifth grade
Parent interview	X	X	X	X	X	X
Child assessments	X	X	X	X	X	X
Teacher questionnaire part A	X	X	X	X[2]	X	
Teacher questionnaire part B	X	X	X	X[2]	X	
Teacher questionnaire part C	X	X	X	X[2]	X	
Teacher questionnaire (teacher level)						X[3]
Reading teacher questionnaire						X
Mathematics teacher questionnaire						X
Science teacher questionnaire						X
Special education teacher questionnaire part A		X		X	X	X
Special education teacher questionnaire part B		X		X	X	X
Adaptive Behavior Scale[4]		X		X		
Self-description questionnaire					X	X
Food consumption questionnaire						X
School administrator questionnaire		X		X[5]	X	X[6]
Student record abstract		X		X	X	X
School fact sheet					X[7]	
School facilities checklist		X		X	X	X
Salary and benefits questionnaire[8]		X				
Head Start verification[9]		X				

X Round that included the instrument.

[1] The fall-first grade data collection consisted of a 30 percent subsample of the study schools. See the *User's Manual for the ECLS-K First Grade Public-Use Data Files and Electronic Code Book* (NCES 2002-135) (Tourangeau, Burke et al. 2002) for information about the purposes and methods of the fall-first grade data collection.

[2] In spring-first grade, there were two sets of teacher questionnaires—one for the teachers of children who had made the transition to the first grade or any higher elementary school grade, and the second for teachers of children who were repeating or attending the second year of kindergarten.

[3] In spring-fifth grade, teacher questionnaires part A, B, and C were replaced by a teacher-level questionnaire and questionnaires for reading, math, and science teachers.

[4] The Adaptive Behavior Scale was completed only for children with disabilities who could not otherwise be directly assessed.

[5] In spring-first grade, there were two different school administrator questionnaires—one for school administrators in schools new to the study and one for school administrators in schools that participated in the base year data collection.

[6] In spring-fifth grade, questions from the School Fact Sheet used in spring-third grade were included in the school administrator questionnaire.

[7] The items in the school fact sheet were included in the school administrator questionnaire in kindergarten and in first grade. These items were reintegrated into the school administrator questionnaire in the fifth-grade data collection.

[8] The salary and benefits questionnaire collected information on the base salary, merit pay, and health benefit pay of teachers and principals. It was completed by the school or district business administrator or by a private school administrator or headmaster.

[9] The Head Start Verification Study confirmed parent and school reports of children's Head Start participation by matching information on the name and location of the Head Start facilities the children were reported to have attended against a database of Head Start centers. For each match, the center was contacted to confirm that the child had attended the center in the year before kindergarten.

SOURCE: U.S. Department of Education, National Center for Education Statistics, Early Childhood Longitudinal Study, Kindergarten Class of 1998–99 (ECLS-K), fall 1998, spring 1999, fall 1999, spring 2000, spring 2002, and spring 2004.

2.1 Direct Child Assessments

One-on-one, untimed direct child assessments were administered using both hard-copy instruments and computer-assisted interviewing (CAI) in the spring of the 2003–04 school year.[3] The children were assessed regardless of whether they were on grade level (i.e., in fifth grade). On average, the assessments took about 96 minutes to administer. Exhibit 2-2 displays the major domains measured during the direct child assessments from all six rounds of data collection. As in the previous rounds, the fifth-grade assessments included cognitive and physical components. In addition, the fifth-grade assessment contained one questionnaire completed by the children: the self-description questionnaire (SDQ), with questions about children's socioemotional development. Children were interviewed using the food consumption questionnaire (FCQ), with questions about the kinds of food the children could buy at school and the food that they had eaten in the past week. The spring-fifth grade cognitive assessment scores include measures that can be compared with the base year assessments conducted in the fall of 1998 and the spring of 1999, with the first-grade assessments conducted in the fall of 1999 and the spring of 2000, and with the third-grade assessments conducted in the spring of 2002 to study children's gains in reading, mathematics, and science. (Measuring gains in science can only be compared for the third- and fifth-grade rounds, since science items were not administered in kindergarten or first grade.) Chapter 3 contains a detailed description of the scores and information on their use and interpretation.

The fifth-grade direct child assessment began by verifying the child's name and administering a short set of warm-up exercises similar in form to the items used in the SDQ. The assessor then administered the SDQ followed by the reading, mathematics, and science assessments; the FCQ; and then the physical measurements (i.e., height and weight).

[3] The majority of fifth-grade assessments were conducted at school (11,024), but some were assessed elsewhere (270), such as at home.

Exhibit 2-2. Direct child assessments, by domain and round of data collection: School years 1998–99, 1999–2000, 2001–02, and 2003–04

Direct child assessment	1998–99 school year		1999–2000 school year		2001–02 school year	2003–04 school year
	Fall-kindergarten	Spring-kindergarten	Fall-first grade	Spring-first grade	Spring-third grade	Spring-fifth grade
Language screener (Oral Language Development Scale[OLDS])	X	/	/	/		
Food consumption questionnaire (FCQ)						X
Reading (language and literacy)	X	X	X	X	X	X
Mathematical thinking	X	X	X	X	X	X
Socioemotional development					X	X
General knowledge (science and social studies)	X	X	X	X		
Science					X[2]	X
Psychomotor	X					
Height and weight	X	X	X	X	X	X

X Round that included the instrument.
/ OLDS (Oral Language Development Scale) was given to language-minority students new to the study in the spring, or who did not pass the cut score in the English version during the previous OLDS administration. The screener determined if the children understood English well enough to receive the direct child assessments in English. For further information on the language screener, please refer to the *ECLS-K Base Year Public-Use Data Files and Electronic Code Book: User's Manual* (NCES 2001-029r) (Tourangeau, Burke, Lê, et al. 2004). The screener was not used in third or fifth grade because the majority of minority-language children (86 percent) passed it by spring-first grade.
[2] In spring-third grade, the general knowledge assessment was replaced with a science assessment. Children received a science assessment in third and fifth grade that measured their understanding of science concepts and scientific investigation skills.
SOURCE: U.S. Department of Education, National Center for Education Statistics, Early Childhood Longitudinal Study, Kindergarten Class of 1998–99 (ECLS-K), fall 1998, spring 1999, fall 1999, spring 2000, spring 2002, and spring 2004.

2.1.1 Socioemotional Development

To measure children's socioemotional development, the ECLS-K assessors administered the Self-Description Questionnaire (SDQ), which is used to determine how children think and feel about themselves both socially and academically. The SDQ consists of 42 statements. Children rated their perceptions of their competence and interest in reading, mathematics, and "all school subjects." They also rated their perceptions of competence and popularity with peers and reported on problem behaviors with which they might struggle.

Each behavior was rated in relation to their perception of themselves on a one to four response scale: "not at all true," "a little bit true," "mostly true," or "very true." The 42 items factored into six scales.

- **SDQ Reading** scale includes eight items on reading grades, the difficulty of reading work, and their interest in and enjoyment of reading.

- **SDQ Mathematics** scale includes eight items on mathematics grades, the difficulty of mathematics work, and their interest in and enjoyment of mathematics.

- **SDQ School** scale includes six items on how well they do in “all school subjects” and their enjoyment of “all school subjects.”

- **SDQ Peer** scale includes six items on how easily they make friends and get along with children as well as their perception of their popularity.

- **SDQ Anger/Distractibility** scale includes six items on externalizing problem behaviors such as fighting and arguing “with other kids,” talking and disturbing others, and problems with distractibility.

- **SDQ Sad/Lonely/Anxious** scale includes eight items on internalizing problem behaviors such as feeling “sad a lot of the time,” feeling lonely, feeling ashamed of mistakes, feeling frustrated, and worrying about school and friendships.

The items on the first four scales were adapted with permission from the *Self-Description Questionnaire I* (Marsh 1992). The items in the two problem behavior scales were developed specifically for the ECLS-K.

In order to pace the assessment, assessors read the SDQ questions to each child even if a child said that he or she could read them. In this way, children’s responses were not affected by their reading ability, or differences in the amount of time children might have spent reflecting on their responses. Children were given a few seconds after each statement was read to mark their response in the SDQ questionnaire. Assessors were trained to maintain a brisk pace so that the children were not tempted to move ahead and so that the child’s overall evaluation was obtained. The assessors were also trained not to look at the children’s answers so that the children would not be tempted to answer in a more positive way then they would have otherwise. The entire questionnaire took about 5 minutes to administer. Assessors entered the answers into the computer after the child had completed the remaining assessments and had left the room.

2.1.2 Cognitive Components

The direct cognitive assessments were individually administered at all six time points. A two-stage cognitive assessment approach was used to maximize the accuracy of measurement and reduce administration time by using the children's responses from a brief first stage routing test to select a second stage form of the appropriate level of difficulty.[4] The kindergarten-first grade (K-1) cognitive assessment focused on three general content areas: (1) reading; (2) mathematics; and (3) knowledge of the social and physical world, referred to as "general knowledge." The K-1 assessment did not ask the children to write anything or to explain their reasoning; rather, children pointed to their answers or responded orally to complete the tasks. The assessment battery was administered using small easels with the items printed on one side and administration instructions for the assessor on the other side. Assessors entered children's responses on a laptop computer.

The direct cognitive domains measured in kindergarten and first grade included reading, mathematics, and general knowledge. In third and fifth grades, the direct cognitive domains measured reading, mathematics, and science. In third and fifth grades, general knowledge was replaced with science because the curriculum at these grades is more differentiated and the amount of time available to administer the assessments was limited. The fifth-grade assessments also utilized a two-stage design. Easels were used to administer items in reading, mathematics, and science. The students also completed workbooks with open-ended mathematics questions. The reading passages were in a booklet format to allow the student to refer back to the story when answering the questions. All questions were read by the assessor. Although the child read the response options to him/herself in the reading assessment, the assessor read all the response options to the child in the mathematics and science assessments.

The ECLS-K fifth-grade direct cognitive assessment battery was designed to assess children's academic achievement in spring of fifth grade, and to provide a means of measuring academic growth since kindergarten entry. Child development, elementary education, and content area experts were consulted on the design and development of the assessment instruments. They recommended that the knowledge and skills assessed by the ECLS-K fifth-grade assessments should represent the typical and important cognitive goals of elementary schools' curricula. The subject matter domains of language use and literacy skills (reading), mathematics, and science were selected. This focus on the main academic

[4] For details on the two-stage assessment design, see the *ECLS-K Base Year Public-Use Data Files and Electronic Code Book: User's Manual* (NCES 2001-029r) (Tourangeau, Burke, Lê, et al. 2004) or the *ECLS-K, Psychometric Report for Kindergarten Through First Grade* (NCES 2002–05) (Rock and Pollack 2002b).

subjects of the elementary grades was made because of the central nature of these skills as antecedents of individuals' later educational outcomes.

In order to measure growth across time, a longitudinal scale is needed. Therefore, the cognitive assessments were designed to have overlapping items, i.e., items that were included in at least two rounds of data collection.

Pools of test items in each of the content domains were developed by a team of elementary education specialists. Items were chosen to extend the longitudinal scales initiated in kindergarten, first grade, and third grade, but there were grade-appropriate changes in content and format. For example, in the kindergarten and first-grade reading assessment, children read short sentences. By fifth grade, the new passages were more complex and more text was presented on a single page than had been the case in the third-grade reading assessment. Test items were reviewed by elementary school curriculum and content area specialists for appropriateness of content and difficulty, and for relevance to the test framework. In addition, items were reviewed for issues related to sensitivity to minority concerns. Items that passed these content, construct, and sensitivity screenings were field tested in the spring of both 2000 and 2002. The validity of the content in the ECLS-K item pools was established by comparing the results of the ECLS-K with scores on the Woodcock-McGrew-Werder Mini-Battery of Achievement (MBA)[5] that was also administered during the field test. Additional information about the development of the fifth-grade cognitive assessment battery can be found in the *ECLS-K Psychometric Report for the Fifth Grade* (NCES 2006–036) (Pollack et al. 2005).

Reading. The K-1 reading (language and literacy) assessment included questions designed to measure basic skills (print familiarity, letter recognition, beginning and ending sounds, creating rhyming words, "sight" word recognition), vocabulary (receptive vocabulary), and comprehension (listening comprehension, words in context). Comprehension items were targeted to measure skills in initial understanding, developing interpretation, personal reflection, and demonstrating critical stance (i.e., analyzing the way the author wrote the text).

The K-1 reading assessment contained five proficiency levels. These five levels reflect a progression of skills and knowledge. Children were thought to master a level if they passed the items within a level.[6] If a child mastered one of the higher proficiency levels, he or she was very likely to have

---

[5] Woodcock, McGrew, and Werder (1994). *Woodcock-McGrew-Werder Mini Battery of Achievement*, Itasca, IL: Riverside Publishing.

[6] See section 3.1.4.3 for a discussion of highest proficiency mastered.

passed the items that made up the earlier levels as well. The five levels were as follows: (1) identifying upper- and lower-case letters of the alphabet by name; (2) associating letters with sounds at the beginning of words; (3) associating letters with sounds at the end of words; (4) recognizing common "sight" words; and (5) reading words in context.

The third-grade reading assessment included items that were designed to measure phonemic awareness, single word decoding, vocabulary (reading), and passage comprehension. The comprehension items measured skills in initial understanding, developing interpretation, personal reflection, and demonstrating a critical stance. The passage reading section examined sentence, paragraph, and story comprehension and comprised a variety of literary genres including poetry, letters, informational text, and narrative text. The test items marking the highest two K-1 proficiency levels, recognizing common "sight" words and reading words in context, were retained in the third-grade assessment to assess the skills of the lowest-achieving third-graders. Three higher proficiency levels were added: literal inference, extrapolation, and evaluation at the third-grade level.

Thus, the third-grade reading assessment contained five proficiency levels: the two retained from K-1 plus three new levels. These five levels reflected a progression of skills and knowledge: if a child had mastered one of the higher levels, he or she was very likely to have passed the items from the earlier levels as well. The third-grade proficiency levels were as follows: (1) recognizing common "sight" words; (2) reading words in context; (3) making inferences using cues that were directly stated with key words in text (literal inference); (4) identifying clues used to make inferences (extrapolation), and using personal background knowledge combined with cues in a sentence to understand use of homonyms; and (5) demonstrating understanding of author's craft and making connections between a problem in the narrative and similar life problems (evaluation).

The fifth-grade reading assessment included items from the third-grade reading assessment. Items from the third-grade proficiency levels were level 3 making inferences using cues that were directly stated with key words in text (literal inference); level 4 identifying clues used to make inferences (extrapolation), and using personal background knowledge combined with cues in a sentence to understand use of homonyms; and level 5 demonstrating understanding of author's craft and making connections between a problem in the narrative and similar life problems (evaluation). In the fifth-grade, new items were added to the reading assessment. These items were more difficult and contributed to the formation of a proficiency level where children demonstrated their ability to comprehend biographical and expository text (evaluating nonfiction). Children were required to identify the tone of a remark, the author's purpose for a selection, and evidence for and against theories discussed in the text.

The Kindergarten and First-Grade Reading battery links with the Third- and Fifth-Grade Reading battery, creating the following reading proficiency levels over time: (1) identifying upper- and lower-case letters of the alphabet by name (letter knowledge); (2) associating letters with sounds at the beginning of words (beginning sounds); (3) associating letters with sounds at the end of words (ending sounds); (4) recognizing common "sight" words (sight words); (5) reading words in context (words in context); (6) making inferences using cues that were directly stated with key words in text (literal inference); (7) identifying clues used to make inferences (extrapolation); (8) demonstrating understanding of author's craft and making connections between a problem in the narrative and similar life problems (evaluation), and (9) comprehension of biographical and expository text (evaluating non-fiction).

Mathematical Thinking. The K-1 mathematics assessment was designed to measure skills in conceptual knowledge, procedural knowledge, and problem solving. Approximately one-half of the mathematics assessment consisted of questions on number sense and number properties and operations. The remainder of the assessment included questions in measurement; geometry and spatial sense; data analysis, statistics, and probability; and patterns, algebra, and functions. The mathematics assessment contained several items for which manipulatives were available for children to use in solving the problems. Paper and pencil were also offered to the children to use with specific items on the assessment.

The items in the K-1 mathematics assessment could also be grouped into five proficiency levels, though the mathematics clusters were less homogeneous in content than the reading clusters. The clusters of mathematical items included the following: (1) identifying some one-digit numerals, recognizing geometric shapes, and one-to-one counting up to ten objects; (2) reading all one-digit numerals, counting beyond ten, recognizing a sequence of patterns, and using nonstandard units of length to compare the size of objects; (3) reading two-digit numerals, recognizing the next number in a sequence, identifying the ordinal position of an object, and solving a simple word problem; (4) solving simple addition and subtraction problems; and (5) solving simple multiplication and division problems and recognizing more complex number patterns.

The third- and fifth-grade mathematics assessments addressed the following content strands: number sense, properties, and operations; measurement; geometry and spatial sense; data analysis, statistics, and probability; and pattern, algebra, and functions. The cognitive processes (conceptual, procedural, and problem solving) were assessed in each of the strands. Some of the items drew upon knowledge from more than one strand. For example, an item might require that a child apply knowledge about geometry, measurement, and number operations to answer the question correctly. Proficiency levels defined in the third-grade assessment included levels 4 and 5 retained from the earlier test forms, plus two new levels: place value, and rate and measurement. The fifth-grade mathematics assessment retained level 5 from K-1, levels 6 and 7 from third grade, and added two new levels, 8 and 9.

Thus, the items in the fifth-grade mathematics assessment could be grouped into five proficiency levels. The lower level mathematics clusters (i.e., levels 1 and 2, number and shape, relative size) tended to be less homogeneous in content than the reading clusters, while the reverse was true for higher level skills (levels 7, 8, and 9, rate and measurement, fractions, area and volume). The clusters of fifth-grade mathematics items included the following: (1) solving simple multiplication and division problems and recognizing more complex number patterns; (2) demonstrating understanding of place value in integers to hundreds place; (3) using knowledge of measurement and rate to solve word problems; (4) solving problems using fractions; and (5) solving word problems involving area and volume.

The Kindergarten and First-Grade Mathematics battery links with the Third- and Fifth-Grade Mathematics battery, creating the following mathematics proficiency levels over time: (1) identifying some one-digit numerals, recognizing geometric shapes, and one-to-one counting up to ten objects (number and shape); (2) reading all one-digit numerals, counting beyond ten, recognizing a sequence of patterns, and using nonstandard units of length to compare the size of objects (relative size); (3) reading two-digit numerals, recognizing the next number in a sequence, identifying the ordinal position of an object, and solving a simple word problem (ordinality and sequence); (4) solving simple addition and subtraction problems (addition and subtraction); (5) solving simple multiplication and division problems and recognizing more complex number patterns (multiplication and division); (6) demonstrating understanding of place value in integers to hundreds place (place value); (7) using knowledge of measurement and rate to solve word problems (rate and measurement); (8) solving problems using fractions (fractions); and (9) solving word problems involving area and volume (area and volume).

Science. The K-1 assessment battery differed from the third- and fifth-grade batteries. The K-1 battery included a measure of general knowledge whereas the third and fifth grade included a measure of science. The K-1 general knowledge assessment battery consisted of items that measured knowledge in the natural sciences and social studies in a single scale. The science subdomain measured two broad classes of science competencies: (1) conceptual understanding of scientific facts, and (2) skills and abilities to form questions about the natural world, to answer such questions on the basis of the tools and the evidence collected, to communicate answers and to explain how the answers were obtained. The social studies subdomain included questions that measured children's knowledge in a wide range of disciplines such as history, government, culture, geography, economics, and law. The science subdomain included questions from the fields of earth, space, physical, and life sciences. The assessment items drew on children's experiences with their environment, and many questions related to more than one of the categories. The items captured information on children's conception and understanding of the social, physical, and natural world and of their ability to draw inferences and comprehend implications. The skills children need to establish relationships between and among objects, events, or people and to make inferences and to comprehend the implications of verbal and pictorial concepts were measured.

The subject matter content of the K-1 general knowledge assessment domain was too diverse and the items insufficiently ranked or graded to permit the formation of a set of proficiency levels. It was also not possible to develop separate scores for science and social studies. Instead, a single score was calculated to represent each child's breadth and depth of understanding and knowledge of the world around them.

As noted previously, the third- and fifth-grade batteries addressed the science domain. Equal emphasis was placed on life science, earth and space science, and physical science. Similar to the K-1 assessment of general knowledge, children needed to demonstrate understanding of the physical and natural world, draw inferences, and comprehend relationships. In addition, third- and fifth-graders needed to interpret scientific data, formulate hypotheses, and identify the best plan to investigate a given question. As with the K-1 general knowledge assessment, no set of proficiency levels was developed.

2.1.3 Physical Components

In the fall of the base year there were two parts to the physical component of the child assessment: psychomotor and anthropometric. The psychomotor component (fine and gross motor) was not administered beyond fall kindergarten. The anthropometric component was administered in all six rounds. The anthropometric component consisted of recording the children's height (in inches to the nearest quarter inch) and weight (in pounds) to measure their physical growth and development. A Shorr Board (for measuring height) and a digital scale were used to obtain the height and weight measurements, which were recorded on a height and weight recording form and entered into a laptop computer by field staff. Height and weight were measured twice. For additional detail on the procedures used to collect height and weight, see the *ECLS-K Fifth-Grade Methodology Report* (NCES 2006–037) (Tourangeau, Lê, and Nord 2005).

2.1.4 Food Consumption

To measure children's food consumption, in spring-fifth grade the ECLS-K assessors administered the food consumption questionnaire (FCQ), a questionnaire used to determine the kinds of food the children can buy at school and the food they have eaten in the past week. The FCQ for children consisted of 19 questions. In spring-fifth grade, there were also food consumption questions for school administrators. Those are described in section 2.5 below.

In the FCQ for children, the first set of questions was about foods that are high in fat, sodium, and/or added sugars (e.g., candy, salty snacks, soda pop). Children were asked if they could buy these foods at school, and, if so, how often they bought the food in the past week and where they bought the food (vending machine, cafeteria, or somewhere else in school). In the second set of questions, children were asked about whether they ate particular key foods and beverages in the past 7 days, such as milk, sweetened beverages (e.g., soft drinks), fruits and vegetables, and fast food. They were asked to include food they ate at home, at school, at restaurants, or anywhere else.

Items for the FCQ were taken mainly from existing surveys, although some were developed for the ECLS-K. Two main sources for questions were two surveys by the Center for Disease Control/Division of Adolescent and School Health Surveys: the Youth Risk Behavior Surveillance Survey (YRBSS) and the School Health Programs and Policies Survey (SHPPS).[7] The question on fast-food meals was taken from the California Children's Healthy Eating and Exercise Practices Survey (CalCheeps). Questions on soft drinks and children's at-school consumption of snack foods were developed by the U.S. Department of Agriculture (USDA), using YRBSS and CalCheeps questions as models.

Assessors read each question of the FCQ to the child, along with the response categories, and the child circled his or her answer. The child was asked to tell the assessor what he or she circled so the assessor could enter the answer into the computer. At the beginning of the FCQ, there is an example

[7] Information on these CDC surveys is available at http://www.cdc.gov/HealthyYouth/

question to show the child the kinds of questions that would be asked. The example was also used to show the child how to circle a response and to practice telling the assessor what answer had been chosen. After the first few questions of the FCQ, if the child appeared to understand the response categories and was in one of the higher reading categories in the reading assessment, the child was allowed to read the response categories if he or she wanted to do so. For children who were homeschooled by their parents or another adult and did not attend school, questions about food that could be purchased at school were not asked. For these cases, assessors were told to skip questions 1 through 9 and enter “Don’t know” into CAPI for each of these questions and then begin with the statement after question 9. Not all assessors followed these instructions and, therefore, some of the 28 homeschoolers gave responses to the first nine questions.

2.2 Parent Interview

The fifth-grade parent interview was conducted using a computer-assisted interview (CAI). The parent interview was conducted primarily in English, but provisions were made to interview parents who spoke other languages with bilingual English-Spanish interviewers or interpreters for other languages. Most of the interviews were conducted by telephone, but a small percentage (2.5 percent) were conducted in person.

The parent interview for the spring-fifth grade data collection lasted on average 44 minutes and contained approximately 330 questions covering fifth-grade school experiences, child care, parent characteristics, and child health. Exhibit 2-3 provides an overview of the topics covered in the fifth grade and in the previous rounds of data collection. Key topics such as family structure, parental involvement in school, and the child’s home environment and cognitive stimulation are covered in most rounds. Other topics, such as parent income, employment, and education, are measured at least once in each school year. The general content areas are similar across the questionnaires, though some topics were added and a few were dropped. For example, in spring-fifth grade, among the questions added were ones on prescription medicines taken if the child had attention deficit disorder (ADD), attention deficit disorder with hyperactivity (ADHD), or hyperactivity. Questions about family therapy were also added. Topics that were dropped included home learning activities, social support, and parental emotional well-being.

Exhibit 2-3. ECLS-K parent interview by major content topics and round of data collection: School years 1998–99, 1999–2000, 2001–02, and 2003–04

Parent questionnaire	1998–99 school year		1999–2000 school year		2001–02 school year	2000–04 school year
	Fall-kindergarten	Spring-kindergarten	Fall-first grade	Spring-first grade	Spring-third grade	Spring-fifth grade
Family structure	X	X	X	X	X	X
Demographics	X	X	X	X	X	X
Household roster	X	X	X	X	X	X
Marital status	X	X	X	X	X	X
Immigration status		X		X	X[1]	X
Primary language(s) spoken in home	X	/	/	/	/	
Parent's involvement with child's school		X	X	X	X	X
Child care	X		X	X	X	X
Current arrangements with relatives	X		X	X	X	X
Current arrangements with nonrelatives	X		X	X	X	X
Current arrangements with centers	X		X	X	X	X
Head Start attendance year before kindergarten	X	/	/	/		
Child care arrangements year before kindergarten	X	/	/	/		
Child's health and well-being	X	X		/	X	X
Birth weight	X	/	/	/		
Physical functioning	X	/	/	X	X	X
Services for children with special needs	X	/	/	X	X	X
Prescription medicine for attention and/or hyperactivity disorders						X
Family therapy						X
Social skills rating	X	X		X		
Home environment and cognitive stimulation	X	X	X	X	X	X
Frequency of literacy activities	X	X	X	X	X	X
Computer use		X	X	X	X	X
Television viewing		X	X	X	X	X
Homework				X	X	X
Family routines					X	X
Summer activities and time use			X			

See notes at end of exhibit.

Exhibit 2-3. ECLS-K parent interview by major content topics and round of data collection: School years 1998–99, 1999–2000, 2001–02, and 2003–04—Continued

Parent questionnaire	1998–99 school year		1999–2000 school year		2001–02 school year	2000–04 school year
	Fall-kindergarten	Spring-kindergarten	Fall-first grade	Spring-first grade	Spring-third grade	Spring-fifth grade
Parental educational expectations for child	X		X	X	X	X
Neighborhood		X	X	X	X	X
Safety		X		X	X	X
Resources (e.g., community center, library)			X			
Parent education	X	/	/	X[2]	X[2]	X[2]
Parent employment	X			X[2]	X[2]	X[2]
Parent income		X		X	X	X
Welfare and other public assistance use	X	X		X	X	X
Parent/child interaction		X		X	X	X
Parent discipline		X			X	X
Parent health and emotional well-being		X			X	X[3]
Relationships and social support	X	X			X	X[4]
Marital satisfaction		X			X	X
Background data	X	X		X		
Mother's age at first birth	X					
Mother's age at child's birth				/		
WIC benefits during pregnancy	X	/	/	/	/	
Whether mother worked for pay between when child was born and time child entered kindergarten	X	/	/	/		
Nonresident parent						
Contact with child	X	X		X	X	X
School involvement		X			X	X
Paternity		X		X	X	
Child support		X		X	X	X

X Rounds that included the construct.
/ Content area asked only of new parent respondents in each round.
1 Asked if new person added to roster or an existing person has missing information on this item.
2 Updated if changed from previous round.
3 In spring, fifth-grade there is a measure of parent health but not well-being.
4 In spring-fifth grade, there is a measure of marital satisfaction but not social support.
SOURCE: U.S. Department of Education, National Center for Education Statistics, Early Childhood Longitudinal Study, Kindergarten Class of 1998–99 (ECLS-K), fall 1998, spring 1999, fall 1999, spring 2000, spring 2002, and spring 2004.

The order of preference for the respondent to the parent interview was the same as in previous rounds: (1) the respondent from the previous round (if there was one), (2) the child's mother, (3) another parent or guardian, or (4) some other adult household member. In a majority of the cases (91 percent), the fifth-grade respondent was the same as the respondent from the previous round. The child's mother was the respondent in 81 percent of the cases and the child's father in 8 percent. Other adults completed the parent interview in 11 percent of the cases (typically grandparents of the sample child).

2.3 Teacher Questionnaires

In the first five rounds of data collection, each sampled child's regular classroom teacher (i.e., the teacher who taught the child for the majority of the day) completed the teacher questionnaires. In spring-fifth grade, each sampled child's reading teacher and either his or her mathematics or science teacher completed questionnaires. In some schools, the sampled children were taught reading, mathematics, and science by the same teacher in one classroom. In other schools, different teachers taught these subjects to the sampled children.

During the spring-fifth grade data collection, each child's teacher received a self-administered teacher-level questionnaire about a variety of topics, including instructional practices, classroom resources, views on teaching and the school, and teacher background. Three additional questionnaires specifically about the focal child were also distributed for teachers in reading, mathematics, and science. Each teacher received a teacher questionnaire in addition to at least one child-level questionnaire in reading, mathematics, or science. All students were assigned to have a reading teacher complete questionnaires. Half of the students were randomly assigned to have a mathematics teacher complete questionnaires, and the other half of the students were assigned to have a science teacher complete questionnaires. In cases where the same teacher taught the sample child reading, mathematics, and science, the teacher was asked to complete a reading questionnaire and either a mathematics or science questionnaire, depending upon the domain to which the child was assigned.

The reading teacher questionnaire had three different sections. The first section included questions from the Social Rating Scale (SRS) that collected data on five areas of children's social skills. The second section had questions from the Academic Rating Scale (ARS) and gathered data on each sampled child's skills in areas of language and literacy. The third section asked child-specific instructional information (e.g., child's grade, additional tutoring or services the child received), asked the teacher to rate how this child behaved and performed in language and literacy relative to the other children in the class, and asked about the teacher's classroom and the characteristics of the students, instructional activities and curricular focus, and instructional practices in language arts. The mathematics teacher questionnaire included questions from the Academic Rating Scale (ARS) gathering data on each sampled child's skills in mathematics, asked child-specific specific instructional information (e.g., child's grade, additional tutoring or services the child received), asked the teacher to rate how this child behaved and performed in mathematics class relative to the other children in the class, and asked about the teacher's classroom and the characteristics of the students, instructional activities and curricular focus, and instructional practices in mathematics. The science teacher questionnaire was similar to the mathematics teacher questionnaire with the questions focusing on science rather than mathematics. Teachers responded to two of these questionnaires for each sampled child. Therefore, data were gathered on each sampled child's skills in the areas of language and literacy and mathematical thinking, or in the areas of language and literacy and science. The ARS and SRS are described in more detail in sections 2.3.1 and 2.3.2, respectively.

In addition to the teacher questionnaire described above, the ECLS-K also included special education teacher questionnaires described in section 2.4.

Exhibit 2-4 shows the distribution of topics covered in the spring-fifth grade teacher questionnaires and previous rounds of data collection.

Exhibit 2-4. Teacher questionnaires by major content topics and round of data collection: School years 1998–99, 1999–2000, 2001–02, and 2003–04

Topic	1998–99 school year		1999–2000 school year		2001–02 school year	2003–04 school year
	Fall-kindergarten	Spring-kindergarten	Spring-first grade (First grade teacher)	Spring-first grade (Kindergarten teacher)	Spring-third grade (Third grade teacher)	Spring-fifth grade (Fifth grade teacher)
Description of class—age, race/ethnicity, and sex distribution	X[1]		X[1]	X[1]	X[1]	X[3]
Class organization						
Activities/interest areas	X[2]	/	X[1]	X[2]	X[1]	X[4]
Types of materials/ resources		X[1]	X[1]	X[1]	X[1]	X[4]
Instructional time in different subjects		X[1]	X[1]	X[1]	X[1]	X[4]
Child vs. teacher-initiated activities	X[2]	X[1]	X[1]	X[2]	X[1]	X[3]
Homework time in different subjects					X[1]	X[4]
Time in reading and math achievement groups		X	X	X	X[1]	X[3]
Classroom characteristics						
Children with special needs		X[1]	X[1]	X[1]	X[1]	X[3]
Classroom aides		X[1]	X[1]	X[1]	X[1]	X[4]
Class assignment and grouping		X[1]	X[1]	X[1]	X[1]	X[3]
Behavior of children in classroom	X[1]	X[1]	X[1]	X[1]	X[1]	X[3]
Instructional information					X[1]	X[4]
Language arts		X	X	X	X[1]	X[4]
Mathematics		X	X	X	X[1]	X[4]
Science		X	X	X	X[1]	X[4]
Social studies		X	X	X	X[1]	X[4]
Parental involvement		X[1]	X[1]	X[1]	X[1]	X[5]
Share progress information with parents		X[1]	X[1]	X[1]		X[5]
Professional development		X[1]	X[2]	X[1]	X[2]	X[4]

See notes at end of exhibit.

Exhibit 2-4. Teacher questionnaires by major content topics and round of data collection: School years 1998–99, 1999–2000, 2001–02, and 2003–04—Continued

Topic	1998–99 school year		1999–2000 school year		2001–02 school year	2003–04 school year
	Fall-kindergarten	Spring-kindergarten	Spring-first grade (First grade teacher)	Spring-first grade (Kindergarten teacher)	Spring-third grade (Third grade teacher)	Spring-fifth grade (Fifth grade teacher)
Teachers' evaluation and grading practices	X[2]	/	X[1]	X[2]	X[1]	X[4]
Teachers' views on school readiness	X[2]	/	X[2]	X[2]		
Perceptions about school climate	X[2]	/	X[2]	X[2]	X[2]	X[4]
Perception of personal influence on policies and classroom planning	X[2]	/	X[2]	X[2]	X[2]	X[4]
Teacher demographic information	X[2]	/	X[2]	X[2]	X[2]	X[4]
Teacher experience and education	X[2]	/	X[2]	X[2]	X[2]	X[4]
Job satisfaction	X[2]	/	X[2]	X[2]	X[2]	X[4]
Transition to school activities	X[2]	/	X[2]	X[2]		
Indirect child cognitive evaluation by teacher (ARS)	X[6]	X[6]	X[6]	X[6]	X[6]	X[3]
Language and literacy, and mathematics in all grades; general knowledge (science and social studies) in grades K and 1; science in grades 3 and 5	X[6]	X[6]	X[6]	X[6]	X[6]	X[3]
Child's social skills (SRS)	X[6]	X[6]	X[6]	X[6]	X[6]	X[5]
Additional information on sampled child		X[6]	X[6]	X[6]	X[6]	X[7]
Participation in special services and programs		X[6]	X[6]	X[6]	X[6]	X[7]

See notes at end of exhibit.

Exhibit 2-4. Teacher questionnaires, by major contact topics and round of data collection: School years 1998–99, 1999–2000, 2001–02, and 2003–04—Continued

Topic	1998–99 school year		1999–2000 school year		2001–02 school year	2003–04 school year
	Fall-kindergarten	Spring-kindergarten	Spring-first grade (First grade teacher)	Spring-first grade (Kindergarten teacher)	Spring-third grade (Third grade teacher)	Spring-fifth grade (Fifth grade teacher)
Overall academic skills		X[6]	X[6]	X[6]	X[6]	X[3]
Physical activity levels		X[6]	X[6]	X[6]	X[6]	X[5]
Reading group participation		X[6]	X[6]	X[6]	X[6]	X[5]
Parental involvement		X[6]	X[6]	X[6]	X[6]	X[5]

X Rounds that included the construct.
/ Content area asked only of new teacher participants in each round.
[1] Topic is in teacher questionnaire part A that was used in grades K through 3.
[2] Topic is in teacher questionnaire part B that was used in grades K through 3.
[3] Topic is in reading, math, and science questionnaires specific to the child. These were used only in grade 5.
[4] Topic is in teacher-level questionnaire used in grade 5.
[5] Topic is in the reading teacher questionnaire used in grade 5.
[6] Topic is in teacher questionnaire part C that was used in grades K through 3.
[7] Topic is in the reading and mathematics teacher questionnaires used in grade 5.
SOURCE: U.S. Department of Education, National Center for Education Statistics, Early Childhood Longitudinal Study, Kindergarten Class of 1998–99 (ECLS-K), fall 1998, spring 1999, fall 1999, spring 2000, spring 2002, and spring 2004.

2.3.1 Content of the Academic Rating Scale (ARS)

The kindergarten and first-grade ARS contained three scales: language and literacy, mathematics, and general knowledge. There were four scales on the third-grade ARS: language and literacy, mathematical thinking, science, and social studies. Three scales were used on the fifth-grade ARS: language and literacy, mathematical thinking, and science. In spring of fifth grade every child's reading teacher completed child level information and a language and literacy ARS. Children were also rated by either their science teacher or their mathematics teacher. Thus, each child was rated on two of the content areas—either language and literacy and mathematics or language and literacy and science.

The areas measured in the ARS overlap and augment what is measured in the direct cognitive assessment. The items were designed to ascertain the current skill levels, knowledge, and behaviors of the child in fifth grade based on the teacher's past observations of the child with the selected content. In the fifth grade, the teacher most knowledgeable of each sampled child's skills and knowledge in each of the content areas was asked to complete the ratings. The questionnaires were mailed to the appropriate content area teacher to complete.

Although the topics covered in the ARS are similar across years, the skills that children exhibit for a particular topic, such as reads fluently, increase by grade. Teachers were provided with examples that helped them establish the level of difficulty of a particular item. For example, reading fluency is covered in first, third, and fifth grade, but the third- and fifth-grade items set a higher difficulty level, as seen below:

- **Spring-first grade: Reads first grade books fluently—**for example, easily reads words in meaningful phrases rather than reading word by word.
- **Spring-third grade: Reads fluently—**for example, easily reads words as part of meaningful phrases rather than word by word, including words with three or more syllables, such as rambunctious, residential, genuinely, and pneumonia.
- **Spring-fifth grade: Reads fluently**—for example, utilizes vocal expression and appropriate pacing when reading aloud, or does dramatic readings increasing pace to denote excitement.

Similarly, in mathematics, children are asked to demonstrate an understanding of place value across grades, but the third- and fifth-grade items set a higher level of difficulty:

- **Spring-first grade: Demonstrates an understanding of place value—**for example, by explaining that fourteen is ten plus four, or using two stacks of ten and five single cubes to represent 25.
- **Spring-third grade: Shows understanding of place value with whole numbers—**for example, correctly orders the numbers 19,321, 14,999, 9,900, and 20,101 from least to greatest, or correctly regroups when adding and subtracting.
- **Spring-fifth grade: Shows understanding of place value**—for example, compares decimals to the thousandths place (1.04 > 1.009).

Below is a description of the content of the fifth-grade ARS.

- The **Language and Literacy** section of the ARS consists of nine items. Teachers are asked to rate each child's proficiency in expressing ideas, use of strategies to gain information, reading on grade level, and writing. This section also includes a question about the child's use of the computer for a variety of purposes. This question is not included in the Language and Literacy scale.
- The **Mathematical Thinking** section of the ARS consists of 10 items. Teachers are asked to rate each child's proficiency in the following areas: number concepts (place value, fractions, and estimation), measurement, operations, geometry, application of mathematical strategies, and beginning algebraic thinking.

- The **Science** section of the ARS consists of seven items. Teachers are asked to rate each child's ability to make predictions, form explanations and conclusions based on observation and investigation, communicate scientific information, apply scientific principles, and demonstrate understanding of life science and physical science.

See chapter 3, section 3.2.2 for scale scores, value ranges, means, and standard deviations for the ARS.

2.3.2 Teacher Social Rating Scale

Teachers rated individual students' social development on part C of the teacher questionnaire. In the fifth-grade data collection, the reading teacher completed these social ratings for each student. These items were intended to measure approaches to learning, self-control, and interpersonal skills. The items were rated on a scale of 1 (never) to 4 (very often). The same five scales defined for the K-1 assessments are formed from these items. Three of the scales capture positive aspects of children's development, and two represent problem behaviors. Two items were added to the third and fifth-grade scales due to a high number of maximum (positive) scores on the third-grade field test of these items. One item was added to the externalizing problem behavior scale ("child talks during quiet study time"). The other additional item, "child follows classroom rules," was added to the Social Rating Scale (SRS) in an attempt to increase variance in the self-control scale. Analysis of the item responses indicated that it contributed strongly to the approaches to learning scale, increasing the variance and reliability of that scale. Thus, this item was included in the Approaches to Learning scale.

In third grade, examination of the responses suggested a different perception of student's self-control and interpersonal social abilities. The self-control scale included items on control of attention as well as control of emotions and behavior in interactions. Third-grade students who were rated higher on self-control were also rated higher on interpersonal skills that involved peers. A peer relations score that combines responses on both the interpersonal items and self-control items that relate to peers was computed and reported in the third-grade files, as well as these scales reported separately to facilitate comparison with earlier rounds of data collection. All of the scales available for third grade were computed for fifth grade and are available in the file. See chapter 3, section 3.3 for variable names, ranges, means, and standard deviations for these scales.

- The **Approaches to Learning** scale (Teacher SRS) measures behaviors that affect the ease with which children can benefit from the learning environment. It includes six items that rate the child's attentiveness, task persistence, eagerness to learn, learning independence, flexibility, and organization. In the third- and fifth-grade administration, an item "child follows classroom rules" was added to the SRS to increase variance in the self-control scale.

- The **Self-Control** scale (Teacher SRS) has four items that indicate the child's ability to control behavior by respecting the property rights of others, controlling temper, accepting peer ideas for group activities, and responding appropriately to pressure from peers.

- The **Interpersonal Skills** scale (Teacher SRS) has five items that rate the child's skill in forming and maintaining friendships; getting along with people who are different; comforting or helping other children; expressing feelings, ideas, and opinions in positive ways; and showing sensitivity to the feelings of others.

- The **Peer Relations** scale (grade three Teacher SRS) has nine items. The scale is a combination of the items from the **interpersonal skills** and **self-control scales**. This scale represents the self-control and interpersonal skills that are important in establishing and maintaining peer relationships.

The two problem behavior scales reflect behaviors that may interfere with the learning process and the child's ability to interact positively in the classroom.

- **Externalizing Problem Behaviors** scale (Teacher SRS) includes acting out behaviors. The kindergarten and first-grade forms have five items on this scale that rate the frequency with which a child argues, fights, gets angry, acts impulsively, and disturbs ongoing activities. To increase the variance on this scale, an item was added in third and fifth grade asking about the frequency with which a child talks during quiet study time.

- The **Internalizing Problem Behavior** scale (Teacher SRS) asks about the apparent presence of anxiety, loneliness, low self-esteem, and sadness. This scale comprises four items.

These measures are adapted with permission from the instrument *Social Skills Rating Scale: Elementary Scale A ("How Often?")* (SSRS) by Gresham and Elliott (1990). The order of presentation of items was different on the SRS. Unlike the SSRS, the SRS did not separate the problem behavior items from the social skill items. On the SRS, the problem behavior items were interspersed throughout the SRS questionnaire to break any response sets. The SSRS uses a three point response scale while the SRS used a four point scale (never, sometimes, often, very often) and allowed respondents to indicate "no opportunity to observe." Only three of the SSRS social skills items are the same on the SRS. The remainder of the social skills items were adapted (N=6) or new (N=7). Some items were adapted completely to tap a wider representation of the skill (e.g., "keeps belonging organized," "forms and maintains friendships," "easily adapts to changes in routine," "pays attention well," "follows classroom rules"). One item was abbreviated to cover a wider range of situations ("controls temper"). Seven of the social skills items were new items developed for ECLS-K (i.e., “is sensitive to the feelings of others,” “respects the property rights of others,” “shows eagerness to learn new things,” “persists in completing tasks,” “works independently,” “expresses own feelings, opinions, and ideas without putting down those of others,” “comforts or helps other children”). The SRS problem behavior scales were much shorter than the SSRS (ten items on the SRS compared with eighteen on the SSRS). Seven of the items on the SRS problem scales are identical to SSRS problem behavior items. The remaining three items are new (i.e., “worries about things,” “talks during quiet study time”) or adapted from the SSRS (“shows low self-esteem”).

2.4 Special Education Teacher Questionnaires

In the spring-fifth-grade data collection, ECLS-K supervisors reviewed accommodation and inclusion information for children who received special education services. During the preassessment phone call with the school coordinator, the field supervisors asked for the names of sampled children receiving special education services, and the names of the teachers providing this service. The supervisor then listed special education staff working with each child (e.g., speech pathologists, reading instructors, and audiologists). These special education teachers and related services providers received questionnaires. If a child received special education services from more than one special education teacher/provider, a field supervisor determined the child’s primary special education teacher/service provider. The primary special education teacher/service provider was defined as:

- The teacher who managed the child’s Individualized Education Plan (IEP);
- The teacher who spent the most amount of time providing special education services to the child; or
- The teacher who was most knowledgeable about the child’s special needs and use of assistive technologies.

Except for one change, the spring-fifth grade special education teacher questionnaires were identical to the ones used in spring-third grade. A question on the receipt of special education or related services due to an attention deficit/hyperactivity disorder (ADHD) was added to the spring-fifth grade questionnaire. Exhibit 2-5 provides a summary of the content areas addressed in the special education teacher questionnaires in spring-fifth grade and in the previous rounds. The questionnaires addressed topics such as the child's disability, IEP goals, the amount and type of services used by sampled students, and communication with parents and general education teachers.

Exhibit 2-5. Special education teacher questionnaires by major content topics and round of data collection: School years 1998–99, 1999–2000, 2001–02, and 2003–04

Topic	1998-1999 school year Spring-kindergarten	1999-2000 school year Spring-first grade	2001–02 school year Spring-third grade	2003–04 school year Spring-fifth grade
Part A (Teacher Level)				
Teacher's sex	X	X	X	X
Teacher's age	X	X	X	X
Teacher's race/ethnicity	X	X	X	X
Teaching experience	X	X	X	X
Educational background	X	X	X	X
Special education teacher background	X	X	X	X
Location of service provision	X	X	X	X
Student load per week	X	X	X	X
Teacher's main assignment			X	X
Part B (Child Level)				
Disability category	X	X	X	X
IEP goals for the school year	X	X	X	X
Extent of services	X	X	X	X
Types of services provided for the year	X	X	X	X
Primary placement	X	X	X	X
Teaching practices, methods, and materials	X	X	X	X
Assistive technologies used by child	X	X	X	X
General education goals, expectations, and assessments	X	X	X	X
Collaboration/communication with child's general education teacher	X	X	X	X
Frequency of communicating with child's parents	X	X	X	X
Receipt of formal evaluations in the past year	X	X	X	X
When child first had an IEP			X	X
Likelihood child will have an IEP next school year			X	X
Percentage of IEP goals that have been met this school year			X	X
Receipt of special education or related services because of attention deficit/hyperactivity disorder				X

NOTE: Data collected only in the spring of each school year.
SOURCE: U.S. Department of Education, National Center for Education Statistics, Early Childhood Longitudinal Study, Kindergarten Class of 1998–99 (ECLS-K), fall 1998, spring 1999, fall 1999, spring 2000, spring 2002, and spring 2004.

Part A of the special education teacher questionnaire was designed to collect information about the special education teacher's professional background and experience. Part B asked about the special education services provided to the child and the nature of the child's special education curriculum. The special education teacher of a sampled child was asked to complete a copy of part B for each sampled child she or he was responsible for overseeing.

2.5 School Administrator Questionnaire

The principal, administrator, or headmaster at the school attended by the sampled child was asked to complete the school administrator questionnaire in the spring of 2004. This self-administered questionnaire was intended to gather information about the school, student body, teachers, school policies, and characteristics of the administrator. In spring-fifth grade, it also included items that in spring-third grade had been in a questionnaire called the school fact sheet (e.g., the grades taught in the school, school sector and focus,[8] the length of the school year). The school administrator questionnaire was divided into seven sections. The first five sections requested mainly factual information about each school and the programs offered at the school. Either a principal or a designee who was able to provide the requested information could complete these sections. The school's principal was asked to complete the remaining two sections concerning his or her background and evaluations of the school climate. If a designee was chosen to do the last two sections, he or she was instructed to answer the background and education questions about the school's principal or headmaster. Exhibit 2-6 summarizes the content areas addressed in this questionnaire in spring-fifth grade and previous rounds.

In spring-fifth grade, a new content area on food consumption was added. The main purpose of these questions was to determine the availability at school of various foods, including those that are healthy and those that are high in fat, sodium, and/or added sugars. Questions were asked about whether students could purchase food or beverages from vending machines at the school or a school store, canteen, or snack bar. School administrators were also asked if the school offered a la carte lunch or breakfast items to students that were not sold as part of the National School Lunch or the School Breakfast Program. In addition, questions were asked about whether children could buy particular foods and beverages at school, such as milk, sweetened beverages (e.g., soft drinks), fruits and vegetables, candy, and salty snacks; where these foods could be obtained in the school (e.g., a school store, a vending

[8] School focus refers to whether the school is a regular school or a school with a particular focus such as a magnet, charter, tribal, special education, or other type of school.

machine); and how full the cafeteria is at peak meal times. Questions on the availability of foods not part of U.S. Department of Agriculture (USDA) meal programs and cafeteria crowding were taken from the School Health Policies and Programs Study (SHPPS). The sources for the other food consumption questions in the school administrator questionnaire are the same as those described in section 2.1.4 for the children's food consumption questionnaire.

2.6 School Facilities Checklist

ECLS-K supervisors completed the school facilities checklist during their visits to the school in the spring of fifth grade. The facilities checklist collects information about the (1) presence of security measures, and (2) school neighborhood characteristics. The school facilities checklist was shorter in spring-fifth grade than in spring-third grade.[9]

2.7 Student Records Abstract Form

School staff completed the student records abstract form for each sampled child in the spring of kindergarten, first grade, third, and fifth grade. This instrument was used to obtain information about the child's attendance record and, if applicable, details on a child's IEP. The spring-fifth grade version was the same as the spring-third grade version. Both spring-fifth and spring-third versions of the student records abstract form differed from the spring-kindergarten version in two ways. First, no data were collected on the pre-kindergarten Head Start status of children in the third grade. Second, two questions on the form were modified to enable the school to provide more comprehensive answers to the question of the status of the child in the previous school year (1998–99) and whether a student had an IEP. (See chapter 5, section 5.5.5 for more details on the collection of these forms.)

---

[9] The spring-third grade version had additional questions about the number of portable classrooms on school grounds, the presence of environmental factors that may affect the learning environment, and the overall learning climate of the school.

Exhibit 2-6. School administrator questionnaire, by major content topics and round of data collection: School years 1998–99, 1999–2000, 2001–02, and 2003–04

Topic	1998–99 school year	1999-2000 school year		2001–02 school year	2003–04 school year
		Spring-first grade			
	Spring-kindergarten	Returning schools	New schools	Spring-third grade	Spring-fifth grade
School characteristics	X	--	X	X	X
Type of school	X		X	X	X
Admission requirements	X				
School size	X	X	X	X	X
Average daily attendance				X	X
Student characteristics	X	X	X	X	X
Race/ethnicity of students	X	X	X	X	X
Children eligible for special services	X	X	X	X	X
Types of kindergarten programs	X				
School facilities and resources	X	--	X	X	X
Computer equipment	X	X	X	X	X
Community characteristics and school safety	X	X	X	X	X
Teaching and other school staff characteristics	X	X	X	X	X
Range of salary paid to teachers	X		X		
Race/ethnicity of staff	X	X	X	X	
Full- and part-time staff in different specialties				X	X
School policies and programs	X	--	X	X	X
Assessments, testing, and retention	X	X	X	X	X
School-family-community connections	X	--	X	X	X
Programs and activities for families	X		X	X	
Parental involvement and participation	X	X	X	X	X
Programs for special populations	X	X	X	X	X
ESL and bilingual education	X	X	X	X	
Special education	X	--	X	X	X
Gifted and talented	X		X	X	X
Principal characteristics	X	X	X	X	X
Sex, race/ethnicity, age of principal	X	X	X	X	X
Experience and education	X	X	X	X	X
School governance and climate	X	X	X	X	X
Goals and objectives for teachers	X	X	X	X	X
School functioning and decisionmaking	X	X	X	X	X
School practices related to student food consumption					X

NOTE: "--" indicates that fewer details on the topic were collected than for new schools.
SOURCE: U.S. Department of Education, National Center for Education Statistics, Early Childhood Longitudinal Study, Kindergarten Class of 1998–99 (ECLS-K), fall 1998, spring 1999, fall 1999, spring 2000, spring 2002, and spring 2004.

3. ASSESSMENT AND RATING SCALE SCORES USED IN THE ECLS-K

Several types of scores were used in the Early Childhood Longitudinal Study, Kindergarten Class of 1998–99 (ECLS-K) to describe children's cognitive and social development during kindergarten through fifth grade. These scores were for the direct cognitive assessment, the Academic Rating Scale (ARS), the Social Rating Scale (SRS), and the self-description questionnaire (SDQ). Descriptions of the scores for each assessment or scale follow, along with variable names, variable descriptions, and descriptive statistics from the ECLS-K data files.[1] Guidelines for when and how to use each cognitive assessment score are also provided in this chapter.

3.1 Direct Cognitive Assessment

The fifth-grade direct cognitive assessment contained items in reading, mathematics, and science. In each subject area, children received an 18- to 25-item routing test. Performance on the routing items guided the selection and administration of one of three second-stage forms. The second-stage form contained items of appropriate difficulty for the level of ability indicated by the routing items.[2]

The fifth-grade direct cognitive assessment built on the framework established in the kindergarten through third-grade rounds of data collection. The design and administration of the assessment instruments, and the scores derived from them, evolved over time to keep pace with children's growth and the objectives of the survey. Changes in the assessments include the following:

- **English language screening:** In kindergarten and first grade, children who were identified as coming from a language minority background were administered a language-screening assessment, the Oral Language Development Scale (OLDS), prior to administration of the direct cognitive assessments. Scores on the OLDS were used to determine children who would be administered the direct cognitive assessments. English language screening was discontinued after spring-first grade because nearly all children in the sample had demonstrated sufficient English proficiency to participate in the full assessment by that time.

[1] This user's manual is applicable to the data gathered during the 2003–04 school year; information contained in this manual about data gathered during the 1998–1999 school year (base year of the study), 1999–2000 school year (first grade), and 2001-02 school year (third grade) is provided primarily for background and comparison purposes.
[2] See chapter 2, section 2.1 of the *ECLS-K Psychometric Report for the Fifth Grade* (NCES 2006–036) (Pollack et al. 2005) for additional information on the two-stage process for the direct cognitive assessments.

- **New assessment instruments:** The four rounds of data collection in kindergarten and first grade used the same set of assessment instruments in reading, mathematics, and general knowledge. Children were routed to different levels of difficulty within each assessment domain depending on their performance on a short routing test in each subject area. Because children's academic skills in the subsequent rounds could be expected to have advanced beyond the levels covered by the original forms, new sets of assessment instruments were developed for the third grade, and again for the fifth grade. Some of the assessment items were retained across rounds to support the development of longitudinal score scales in each subject area.

- **Science assessment:** The K-1 general knowledge assessment included basic natural science concepts as well as concepts in social studies. For third and fifth grades, a science assessment replaced the general knowledge assessment. The longitudinal scale for measuring gains in science spans only the third- and fifth-grade rounds.

- **Assessment format:** The format of the fifth-grade assessment was similar to that of prior rounds, with some changes to accommodate the more advanced level of the questions. As before, a survey administrator presented the questions to the child and entered responses into a computer for each individually administered assessment. As was the case in third grade, the fifth-grade mathematics assessment included a workbook for the questions that required computations or written responses. The reading assessment in third grade was administered in booklet format instead of on an easel to accommodate the length of the reading passages used in the assessment, while the fifth-grade reading assessment had both a booklet containing the reading passages and an easel for the presentation of questions.

- **Item cluster scores:** The K-1 assessment scores included a count of the number right on three questions related to familiarity with conventions of print. Additional cluster scores, based on small numbers of reading and science items, are reported for the third and fifth-grade assessments and are described in detail below.

- **Bridge sample:** Field test results after spring-first grade suggested that the growth in skills between the first- and third-grade assessments might make measurement of gain problematic. Data were collected for a small "bridge sample" of second-graders to support development of longitudinal scales in reading and mathematics. A bridge sample of fourth-graders was not necessary to bridge the gap between the third- and fifth-grade assessments, because field test results showed a sufficient amount of overlap between high achieving third-graders and low achieving fifth-graders.

The scores used to describe children's performance on the direct cognitive assessment included broad-based measures that reported performance in each domain as a whole, as well as targeted scores reflecting knowledge of selected content or mastery within a set of hierarchical skill levels. Some of the scores were simple counts of correct answers, while others were based on Item Response Theory (IRT), which uses patterns of correct and incorrect answers to obtain estimates that are comparable across different assessment forms. The different types of scores that were used to describe children's performance on the direct cognitive assessment are described in detail in this chapter. Number-right scores and IRT scale scores measured children's performance on a set of questions with a broad range of difficulty. Standardized scores (T-scores) reported children's performance relative to their peers. Criterion-referenced proficiency scores and item cluster scores evaluated children's performance with respect to subsets of items that mark specific skills.

Tables 3-1 through 3-10 show the types of scores, variable names, descriptions, and summary statistics for the direct cognitive assessment. The name and description for each variable in the tables begin with a "C," indicating that it is a child variable, and a data collection round number, either 1 (fall-kindergarten), 2 (spring-kindergarten), 3 (fall-first grade), 4 (spring-first grade), 5 (spring-third grade), or 6 (spring-fifth grade). Weighted means in tables containing only fifth-grade scores use the round 6 cross-sectional weight, C6CW0, to represent population estimates for fifth grade. Weighted estimates in tables containing scores for all earlier rounds are based on C1_6SC0, the round 1-2-3-4-5-6 panel weight, while tables containing only scores for science, assessed only in third and fifth grades, use C56CW0, the round 5-6 panel weight. Kindergarten through third-grade scores in this data base differ somewhat from the corresponding scores in the previously released data files because they were re-estimated along with the fifth-grade scores (see section 3.1.2). In addition, all kindergarten through third-grade score statistics presented here differed from previous estimates because the panel weight used restricted estimates to children who participated in all six rounds of data collection (for reading and mathematics scores), or rounds 5 and 6 (science scores).

3.1.1 Number-Right Scores

Number-right scores are counts of the raw number of items a child answered correctly. These scores are useful for descriptive purposes only for assessments that are the same for all children. However, when these scores are for assessments that differ in difficulty, they are not comparable to each other. For example, a student who took the middle difficulty mathematics second-stage form would probably have answered more questions correctly if he or she had taken the easier low form, and fewer if the more difficult high form had been administered. For this reason, raw number-right scores were reported in the database only for the first-stage (routing) tests, which were the same for all children being assessed in that round of data collection. The routing test in each subject area consisted of sets of items spanning a wide range of skills. For example, the kindergarten-first grade (K-1) reading routing test emphasized pre-reading skills, while the routing tests in third and fifth grade contained questions based on reading passages as well as progressively more difficult decoding words. An analyst might use the routing test number-right scores to report actual performance on these particular sets of tasks. Note that because the same routing test was used for the fall-kindergarten through spring-first grade data collections, rounds 1 through 4, score comparisons *may* be made among these rounds. However, the routing test scores in the third and fifth grades, which contained more difficult items, are *not* comparable with the kindergarten or first-grade number-right scores, nor with each other. The third-grade routing test number-right scores should be used only for comparisons *within third grade*, and the fifth-grade scores only *within fifth grade*, but not across grades.

See table 3-1 for the variable names, descriptions, ranges, weighted means, and standard deviations for the routing test number-right scores for the kindergarten and first-grade surveys. Table 3-2 has the same information for the third-grade routing tests, and table 3-3 for the fifth-grade routing tests.

Table 3-1. Direct cognitive assessment: routing test number-right, kindergarten and first grade (K-1) assessments: School years 1998–99 and 1999–2000

Variable	Description	Range of values	Weighted mean	Standard deviation
C1R3RNOR	C1 RC3 Reading Routing #Right - K-1 Assmt	0 - 20	5.8	3.9
C2R3RNOR	C2 RC3 Reading Routing #Right - K-1 Assmt	0 - 20	10.0	4.0
C3R3RNOR	C3 RC3 Reading Routing #Right - K-1 Assmt	0 - 20	11.7	4.1
C4R3RNOR	C4 RC3 Reading Routing #Right - K-1 Assmt	0 - 20	16.3	3.7
C1R3MNOR	C1 RC3 Mathematics Routing #Right - K-1 Assmt	0 - 16	4.6	3.0
C2R3MNOR	C2 RC3 Mathematics Routing #Right - K-1 Assmt	0 - 16	7.2	3.4
C3R3MNOR	C3 RC3 Mathematics Routing #Right - K-1 Assmt	0 - 16	8.9	3.3
C4R3MNOR	C4 RC3 Mathematics Routing #Right - K-1 Assmt	0 - 16	11.8	2.9

NOTE: Table estimates based on C1_6SC0 panel weight. Table estimates may differ from those reported in earlier user's manuals and the *ECLS-Psychometric Report for Kindergarten Through First Grade* (NCES 2002–05) (Rock and Pollack 2002b) because of sample attrition. See chapter 7, section 7.4 for variable naming conventions.
SOURCE: U.S. Department of Education, National Center for Education Statistics, Early Childhood Longitudinal Study, Kindergarten Class of 1998–99 (ECLS-K), fall 1998, spring 1999, fall 1999, and spring 2000.

Table 3-2. Direct cognitive assessment: routing test number-right, third-grade assessment: School year 2001–02

Variable	Description	Range of values	Weighted mean	Standard deviation
C5R3RNR3	C5 RC3 Reading Routing #Right - Gr3 Assmt	0 - 15	10.0	2.8
C5R3MNR3	C5 RC3 Mathematics Routing #Right - Gr3 Assmt	0 - 17	8.9	4.4
C5SROUNR	C5 Science Routing #Right - Gr3 Assmt	0 - 15	8.2	3.4

NOTE: Table estimates based on C1_6SC0 panel weight. Table estimates may differ from those reported in earlier user's manuals and the *ECLS-K Psychometric Report for the Third Grade* (NCES 2005–062) (Pollack et al. 2005) because of sample attrition. See chapter 7, section 7.4 for variable naming conventions.
SOURCE: U.S. Department of Education, National Center for Education Statistics, Early Childhood Longitudinal Study, Kindergarten Class of 1998–99 (ECLS-K), spring 2002.

Table 3-3. Direct cognitive assessment: routing test number-right, fifth-grade assessment: School year 2003–04

Variable	Description	Range of values	Weighted mean	Standard deviation
C6R3RNR5	C6 RC3 Reading Routing #Right - Gr5 Assmt	0 - 25	11.4	5.4
C6R3MNR5	C6 RC3 Mathematics Routing #Right - Gr5 Assmt	0 - 18	9.6	4.9
C6R1SNR5	C6 RC1 Science Routing #Right - Gr5 Assmt	0 - 21	13.2	4.2

NOTE: Table estimates based on C6CW0 cross-sectional weight. See chapter 7, section 7.4 for variable naming conventions.
SOURCE: U.S. Department of Education, National Center for Education Statistics, Early Childhood Longitudinal Study, Kindergarten Class of 1998–99 (ECLS-K), spring 2004.

3.1.2 Item Response Theory Scale Scores; Standardized Scores (T-Scores)

Broad-based scores using the full set of assessment items in reading, mathematics, and science were calculated using IRT procedures. The IRT scale scores estimated children's performance on the whole set of assessment questions, while standardized scores (T-scores) reported children's performance relative to their peers on the content domains. IRT makes it possible to calculate scores that can be compared regardless of which second-stage form a child takes. IRT uses the pattern of right, wrong, and omitted responses to the items actually administered in an assessment and the difficulty, discriminating ability, and "guess-ability" of each item to place each child on a continuous ability scale. The items in the routing tests, plus a core set of items shared among the different second-stage forms and different rounds of data collection, made it possible to establish a common scale. It is then possible to estimate the score the child would have achieved if all of the items in all of the assessment forms had been administered.

IRT has several other advantages over raw number-right scoring. By using the overall pattern of right and wrong responses and the characteristics of each item to estimate ability, IRT can compensate for the possibility of a low-ability student guessing several hard items correctly. If answers on several easy items are wrong, the probability of a correct answer on a difficult item would be quite low. Omitted items are also less likely to cause distortion of scores, as long as enough items have been answered right and wrong to establish a consistent pattern. Unlike raw scoring, which treats omitted items as if they had been answered incorrectly, IRT procedures use the pattern of responses to estimate the probability of correct responses for all assessment questions. Finally, IRT scoring makes possible longitudinal measurement of gain in achievement over time, even though the assessments that are administered are not identical at each point. The common items present in the routing test and in overlapping second-stage forms allow the scores to be placed on the same scale, even as the two-stage design adapts to children's growth over time. As noted earlier, kindergarten and first-grade responses were pooled with third- and fifth-grade data to stabilize the longitudinal estimates. In addition, the maximum values of the scale scores have been extended to include the more difficult items administered in the fifth-grade assessments. The scale scores for each round of user files are defined based on performance on all tasks administered *up to and including* the current round. The re-estimated kindergarten/first-grade and third-grade IRT scores in this database differ from the IRT scores in the kindergarten/first-grade and third-grade files previously released. For example, the reading scale score in the third-grade file is based on test items used in kindergarten through third grade, while the current reading score is an estimate based on an expanded set of items, all of those used in kindergarten through fifth grade. In order to compute meaningful estimates of gains over time, scores for different rounds must be based on comparable sets of tasks. As a result, scores for all previous rounds have been re-estimated so that comparisons can be made.

The IRT scale scores in the database represent estimates of the number of items students would have answered correctly at each point in time if they had taken all of the 186 questions in all of the first- and second-stage reading forms administered in all rounds, the 153 questions in all of the mathematics forms, and the 92 science items. These scores are not integers because they are probabilities of correct answers, summed over all items in the pools. Reading and mathematics gain scores may be obtained by subtracting the re-estimated IRT scale scores at fall-kindergarten from the IRT scale scores at spring-first grade, spring-first grade from spring-third grade, spring-third grade from spring-fifth grade, and so forth. For the science assessment, which was not administered in kindergarten/first grade, gain scores may be computed for third to fifth grade only. The general knowledge test administered in the earlier rounds is not on the same scale. (Note that scores for different subject areas are not comparable to each other because they are based on different numbers of questions and content that is not necessarily equivalent in difficulty (i.e., it would not be correct to assume that a child is doing better in reading than in mathematics because his or her IRT scale score in reading is higher than in mathematics).

See table 3-4 for variable names, descriptions, ranges, weighted means, and standard deviations for the IRT scale scores.

Table 3-4. Direct cognitive assessment: item response theory (IRT) scale scores, fifth-grade assessment: School year 2003–04

Variable	Description	Range of values	Weighted mean	Standard deviation
C1R3RSCL	C1 RC3 Reading IRT Scale Score	0 – 186	29.3	9.8
C2R3RSCL	C2 RC3 Reading IRT Scale Score	0 – 186	40.7	13.2
C3R3RSCL	C3 RC3 Reading IRT Scale Score	0 – 186	46.8	16.0
C4R3RSCL	C4 RC3 Reading IRT Scale Score	0 – 186	70.4	21.9
C5R3RSCL	C5 RC3 Reading IRT Scale Score	0 – 186	116.4	25.5
C6R3RSCL	C6 RC3 Reading IRT Scale Score	0 – 186	137.5	23.6
C1R3MSCL	C1 RC3 Mathematics IRT Scale Score	0 – 153	22.7	9.1
C2R3MSCL	C2 RC3 Mathematics IRT Scale Score	0 – 153	32.6	11.6
C3R3MSCL	C3 RC3 Mathematics IRT Scale Score	0 – 153	39.8	13.5
C4R3MSCL	C4 RC3 Mathematics IRT Scale Score	0 – 153	57.2	16.5
C5R3MSCL	C5 RC3 Mathematics IRT Scale Score	0 – 153	91.0	21.9
C6R3MSCL	C6 RC3 Mathematics IRT Scale Score	0 – 153	112.1	22.0
C5SR1SSCL	C5 RC1 Science IRT Scale Score	0 – 92	43.7	14.2
C6SR1SSCL	C6 RC1 Science IRT Scale Score	0 – 92	56.6	14.3

NOTE: Table estimates based on C1_6SC0 panel weight. Table estimates may differ from those reported in earlier user's manuals, the *ECLS-K Psychometric Report for Kindergarten Through First Grade* (NCES 2002–05) (Rock and Pollack 2002b), and the *ECLS-K Psychometric Report for the Third Grade* (NCES 2005–062) (Pollack et al. 2005) because of re-estimation of scores on a longitudinal scale that includes fifth grade, and because of sample attrition. See chapter 7, section 7.4 for variable naming conventions.
SOURCE: U.S. Department of Education, National Center for Education Statistics, Early Childhood Longitudinal Study, Kindergarten Class of 1998–99 (ECLS-K), spring 2004.

Standardized scores (T-scores) provide norm-referenced measurements of achievement, that is, estimates of achievement *relative to the population as a whole*. A high mean T-score for a particular subgroup indicates that the group's performance is high in comparison to other groups. It does not represent mastery of a particular set of skills, only that the subgroup's mastery level is greater than a comparison group. Similarly, a change in mean T-scores over time reflects a change in the group's status with respect to other groups. In other words, T-scores provide information on *status compared with children's peers*, while the IRT scale scores and proficiency scores represent *status with respect to achievement on a particular criterion set of assessment items*. The T-scores only provide an indicator of the extent to which an individual or a subgroup ranks higher or lower than the national average and how much this relative ranking changes over time.

The standardized scores (T-scores) reported in the database are transformations of the IRT theta (ability) estimates, rescaled to a mean of 50 and standard deviation of 10 using cross-sectional sample weights for each wave of data. For example, a fall-kindergarten reading T-score of 45 (C1R3RTSC) represents a reading achievement level that is one-half of a standard deviation lower than the mean for the fall-kindergarten population represented by the assessed sample of ECLS-K participants. If the same child had a reading T-score of 50 in fifth grade (C6R3RTSC) this would indicate that the child has made up his or her initial deficit and is reading at a level comparable to the national average. T-scores for earlier rounds have been re-estimated using the ability estimates based on the whole longitudinal item pools. Since the T-scores represent status with respect to a peer group rather than with respect to a criterion set of items, the expansion of the item pool should result in only slight changes in the previously-reported T-score estimates. In making T-score comparisons across rounds, the re-estimated scores should be used.

See table 3-5 for variable names, descriptions, and ranges for the standardized T-scores. Weighted means and standard deviations for the kindergarten through third-grade scores in this table deviate slightly from the mean 50.0, standard deviation 10.0 metric because of sample attrition.

Table 3-5. Direct cognitive assessment: standardized scores: School year 2003–04

Variable	Description	Range of values	Weighted mean	Standard deviation
C1R3RTSC	C1 RC3 Reading T-Score	0 - 96	50.4	9.8
C2R3RTSC	C2 RC3 Reading T-Score	0 - 96	50.6	9.6
C3R3RTSC	C3 RC3 Reading T-Score	0 - 96	50.3	9.4
C4R3RTSC	C4 RC3 Reading T-Score	0 - 96	50.2	9.7
C5R3RTSC	C5 RC3 Reading T-Score	0 - 96	50.1	10.0
C6R3RTSC	C6 RC3 Reading T-Score	0 - 96	50.3	9.7
C1R3MTSC	C1 RC3 Mathematics T-Score	0 - 96	50.3	10.3
C2R3MTSC	C2 RC3 Mathematics T-Score	0 - 96	50.2	10.0
C3R3MTSC	C3 RC3 Mathematics T-Score	0 - 96	50.3	9.7
C4R3MTSC	C4 RC3 Mathematics T-Score	0 - 96	50.4	9.4
C5R3MTSC	C5 RC3 Mathematics T-Score	0 - 96	50.2	10.0
C6R3MTSC	C6 RC3 Mathematics T-Score	0 - 96	50.3	9.9
C5R1STSC	C5 RC1 Science T-Score	0 - 96	50.1	10.1
C6R1STSC	C6 RC1 Science T-Score	0 - 96	50.3	9.5

NOTE: Table estimates based on C1_6SC0 panel weight. Table estimates may differ from those reported in earlier user's manuals, the *ECLS-K Psychometric Report for Kindergarten Through First Grade* (NCES 2002–05) (Rock and Pollack 2002b) and the *ECLS-K Psychometric Report for the Third Grade* (NCES 2005-062) (Pollack et al. 2005) because of re-estimation of scores on a longitudinal scale that includes fifth grade, and because of sample attrition. See chapter 7, section 7.4 for variable naming conventions.
SOURCE: U.S. Department of Education, National Center for Education Statistics, Early Childhood Longitudinal Study, Kindergarten Class of 1998–99 (ECLS-K), spring 2004.

3.1.3 Item Cluster Scores

Several item cluster scores are reported for the reading and science assessments. These are simple counts of the number right on small subsets of items linked to particular skills. These clusters of items are also included in the broad-range scores described above. Because they are based on very few assessment items, their reliabilities are relatively low. See section 3.1.7 for reliability statistics. Cluster scores were not constructed for the mathematics assessment. In both reading and science, there were sets of items that were qualitatively different from the rest of the test. In reading, the conventions of print and decoding items represented skills that were different from the main emphasis, growth toward reading comprehension. The science clusters represented curriculum content in three different subject areas, that might be taught independently in any sequence. However, there were no such sets of mathematics items whose content or typical teaching sequence might set them apart from the main body of the assessment.

3.1.3.1 Reading

The K-1 reading assessment contained three questions assessing children's familiarity with conventions of print. The score for these questions was obtained by counting the number of correct answers (zero to three) for the following three items, administered while the child was looking at an illustrated story.

- Indicating that reading goes from left to right;
- Going to the beginning of the next line after a line ends; and
- Finding the end of the story.

These items were part of the reading score calculations in the direct cognitive assessment but did not necessarily fit into a hierarchical pattern of skill mastery. For example, some children scored high on print familiarity but could not recognize letters, while others had the reverse pattern. These items were not included in the third- and fifth-grade reading forms because nearly all children had mastered them by the end of first grade. The print familiarity scores for the four kindergarten and first-grade rounds are based on the same tasks and may be compared with each other.

A score based on four relatively difficult decoding items was reported for the third- and fifth-grade assessments. These were words that were unlikely to be in most children's everyday vocabulary but could be sounded out phonetically. All four words were present in the third-grade reading assessment but only three of the four were in the fifth-grade forms. In order to make the fifth-grade decoding score comparable to the third-grade score for longitudinal comparisons, an estimate of the probability of a correct answer on the missing item was obtained, based on overall performance, for each fifth-grade test taker.

See table 3-6 for variable names, descriptions, ranges, weighted means, and standard deviations for the reading cluster scores: print familiarity and decoding score. The scores in table 3-6 for rounds one through five can be found in the K-1 and third-grade cross-sectional data files, and in the longitudinal fifth-grade file, but not in the fifth-grade cross-sectional file.

Table 3-6. Direct cognitive assessment: reading cluster scores: School year 2003–04

Variable	Description	Range of values	Weighted mean	Standard deviation
C1R3RPRN	C1 RC3 Print Familiarity	0 – 3	1.8	1.1
C2R3RPRN	C2 RC3 Print Familiarity	0 – 3	2.3	0.9
C3R3RPRN	C3 RC3 Print Familiarity	0 – 3	2.6	0.8
C4R3RPRN	C4 RC3 Print Familiarity	0 – 3	2.8	0.6
C5R3RDEC	C5 RC3 Decoding Score Gr3	0 – 4	1.1	1.3
C6R3RDEC	C6 RC3 Decoding Score Gr5	0 – 4	2.1	1.4

NOTE: Table estimates based on C1_6SC0 panel weight. Table estimates may differ from those reported on earlier user files because of sample attrition. See chapter 7, section 7.4 for variable naming conventions.
SOURCE: U.S. Department of Education, National Center for Education Statistics, Early Childhood Longitudinal Study, Kindergarten Class of 1998–99 (ECLS-K), spring 2004.

3.1.3.2 Science

The 21 routing test items of the round 6 fifth-grade science assessment tapped a range of basic concepts, with seven questions each in life science, physical science, and earth science:

- Life Science: a sample of concepts related to anatomy/health, animal characteristics/ behavior, and botany/ecology;
- Physical Science: a sample of concepts related to states of matter, sound, physical characteristics, and the scientific method; and
- Earth Science: a sample of concepts related to the solar system, earth, soil, minerals, and weather.

The 21 fifth-grade items included the 15 items that had been used for 5-item cluster scores in third grade, plus two additional, more difficult, items in each area. Number-right scores for each of these item clusters are reported. Scores for the 7-item clusters are reported for fifth-graders only, since the harder items were not present in the third-grade assessment. Scores based on the 5-item third-grade subsets are reported for both rounds and may be used for comparison purposes. For example, C5LIFESC and C6LIFESC are the number right on the same set of five Life Science items at two different times, third grade (round 5) and fifth grade (round 6). C6LIFES5 is the number right for the fifth-grade set of 7 items, which includes the 5 items tested in third grade plus 2 additional items. However, the item clusters are not designed to function as subscores, representative of the whole science domain. The scale scores and standardized scores, which are based on a much larger sampling of content, are more appropriate to use for measurement of status and gain.

The item clusters were not selected to have comparable levels of difficulty in the different content areas. For example, the fifth-graders' mean of 4.8 for the round 6 life science cluster compared with 4.3 for earth science does not mean in any sense that children were doing better or learning more relative to the domain curriculum in life science compared with earth science. These clusters simply sample a small set of questions of varying difficulty and content within each domain.

See table 3-7 for variable names, descriptions, ranges, weighted means, and standard deviations for the science cluster scores.

Table 3-7. Direct cognitive assessment: science cluster scores: School year 2003–04

Variable	Description	Range of values	Weighted mean	Standard deviation
C5LIFESC	C5 Life Science Gr3 Item Set	0 – 5	3.0	1.4
C5PHYSSC	C5 Physical Science Gr3 Item Set	0 – 5	2.7	1.4
C5EARTSC	C5 Earth Science Gr3 Item Set	0 – 5	2.6	1.3
C6LIFESC	C6 Life Science Gr3 Item Set	0 – 5	3.7	1.3
C6PHYSSC	C6 Physical Science Gr3 Item Set	0 – 5	3.4	1.2
C6EARTSC	C6 Earth Science Gr3 Item Set	0 – 5	3.5	1.3
C6LIFES5	C6 Life Science Gr5 Item Set	0 – 7	4.7	1.7
C6PHYSS5	C6 Physical Science Gr5 Item Set	0 – 7	4.1	1.5
C6EARTS5	C6 Earth Science Gr5 Item Set	0 – 7	4.3	1.7

NOTE: Table estimates based on C56CW0 panel weight. Table estimates may differ from those reported on earlier user files because of sample attrition. See chapter 7, section 7.4 for variable naming conventions.
SOURCE: U.S. Department of Education, National Center for Education Statistics, Early Childhood Longitudinal Study, Kindergarten Class of 1998–99 (ECLS-K), spring 2004.

3.1.4 Proficiency Levels

Proficiency levels provide a means of distinguishing status or gain in specific skills within a content area from the overall achievement measured by the IRT scale scores and T-scores. Clusters of four assessment questions having similar content and difficulty were included at nine points along the score scales of the reading and mathematics assessments. Clusters of four items provided a more reliable assessment of proficiency than did single items because of the possibility of guessing; it is very unlikely that a student who had not mastered a particular skill would be able to guess enough answers correctly to pass a four-item cluster. The following reading and mathematics proficiency levels were identified in the reading and mathematics assessments for kindergarten through fifth grade. No proficiency scores were computed for the science assessment because the questions did not follow a hierarchical pattern.

3.1.4.1 Reading

- **Level 1: Letter recognition:** identifying upper- and lower-case letters by name;
- **Level 2: Beginning sounds:** associating letters with sounds at the beginning of words;
- **Level 3: Ending sounds:** associating letters with sounds at the end of words;
- **Level 4: Sight words:** recognizing common "sight" words;
- **Level 5: Comprehension of words in context:** reading words in context;
- **Level 6: Literal inference:** making inferences using cues that are directly stated with key words in text (for example, recognizing the comparison being made in a simile);
- **Level 7: Extrapolation:** identifying clues used to make inferences, and using background knowledge combined with cues in a sentence to understand use of homonyms;
- **Level 8: Evaluation:** demonstrating understanding of author's craft (how does the author let you know…), and making connections between a problem in the narrative and similar life problems; and
- **Level 9: Evaluating non-fiction:** critically evaluating, comparing and contrasting, and understanding the effect of features of expository and biographical texts.

3.1.4.2 Mathematics

- **Level 1: Number and shape:** identifying some one-digit numerals, recognizing geometric shapes, and one-to-one counting of up to ten objects;
- **Level 2: Relative size:** reading all single-digit numerals, counting beyond ten, recognizing a sequence of patterns, and using nonstandard units of length to compare objects;
- **Level 3: Ordinality, sequence:** reading two-digit numerals, recognizing the next number in a sequence, identifying the ordinal position of an object, and solving a simple word problem;
- **Level 4: Addition/subtraction:** solving simple addition and subtraction problems;
- **Level 5: Multiplication/division:** solving simple multiplication and division problems and recognizing more complex number patterns;
- **Level 6: Place value:** demonstrating understanding of place value in integers to the hundreds place;
- **Level 7: Rate and measurement:** using knowledge of measurement and rate to solve word problems;
- **Level 8: Fractions:** demonstrating understanding of the concept of fractional parts; and
- **Level 9: Area and volume:** solving word problems involving area and volume, including change of units of measurement.

The proficiency levels were assumed to follow a Guttman model, that is, a student passing a particular skill level was expected to have mastered all lower levels; a failure should be consistent with nonmastery at higher levels. Only a very small percentage of students in kindergarten through fifth grade had response patterns that did not follow the Guttman model, that is, a failing score at a lower level followed by a pass on a more difficult item cluster. Overall, including all six rounds of data collection, less than 7 percent of reading response patterns, and about 3 percent of mathematics assessment results, failed to follow the expected hierarchical pattern. This does not necessarily indicate a different order of learning for these children; since most of the proficiency-level items were multiple choice, many of these reversals may be due to children guessing.

Two types of scores are reported with respect to the proficiency levels: a single indicator of highest level mastered, and a set of IRT-based probability scores, one for each proficiency level. More information on each of these types of scores is provided below. As for the other IRT-based scores (scale scores and T-scores), re-estimated values for earlier rounds should be used when making comparisons of proficiency levels across rounds.

3.1.4.3 Highest Proficiency Level Mastered

Mastery of a proficiency level was defined as answering correctly at least three of the four questions in a cluster. This definition results in a very low probability of guessing enough right answers by chance, generally less than 2 percent. At least two incorrect or “don’t know” responses indicated lack of mastery of a cluster. Questions that were answered with an explicit “I don’t know” were treated as wrong, while omitted items were not counted. Since the ECLS-K direct cognitive child assessment was a two-stage design (where not all children were administered all items), and since more advanced assessment instruments were administered in third and fifth grades, children’s data did not include all of the assessment items necessary to determine pass/fail for every proficiency level at each round of data collection. The missing information was not missing at random; it depended in part on children being routed to second stage assessment forms of varying difficulty within each round, and in part on the range of difficulty of the assessments at the different grade levels. In order to avoid bias due to the non-randomness of the missing proficiency level scores, imputation procedures were undertaken to fill in the missing information.

Pass or fail for each proficiency level was based on actual counts of correct or incorrect responses, if they were present. If too few items were administered or answered to determine mastery of a level, a pass/fail score was assigned based on the remaining proficiency scores only if they indicated a pattern that was unambiguous. That is, a “fail” was inferred for a missing level if there were easier cluster(s) that had been failed *and* no higher cluster passed; or a “pass” was assumed if harder cluster(s) were passed *and* no easier one failed. In the case of ambiguous patterns (e.g., pass, missing, fail, where the missing level could legitimately be either a pass or a fail), an additional imputation step was undertaken that relied on information from the child’s performance on all of the domain items answered in that round of data collection. IRT-based estimates of the probability of a correct answer were computed for each missing assessment item. Then a random number was generated and compared with the computed probability. A right answer was imputed if the random number was less than or equal to the probability; a wrong answer if the random number was greater than the probability. At a low level of ability (and low probability of a correct answer), at least some test takers could be expected to give a correct answer, even if only by guessing. Conversely, some children with a high probability of a right answer (based on their other responses) might get any given item wrong. The imputation procedure employed took this into account, and thus preserved variance better than simply rounding the probability or setting an arbitrary cut point for imputation of right/wrong answers. These imputed responses were then aggregated in the same manner as actual responses to determine mastery at each of the missing levels. About 67 percent of the “highest level” scores in reading and 80 percent in mathematics were determined on the basis of item response data alone; the rest utilized IRT-based probabilities for some or all of the missing items. Scores were not imputed for missing levels that included a reversal (e.g., fail, blank, pass) because no resolution of the missing data could result in a consistent hierarchical pattern.

Scores in the data file represented the highest level of proficiency mastered by each child at each round of data collection, whether this determination was made by actual item responses alone or by a combination of item responses and imputation methods. The highest proficiency level mastered implies that children demonstrated mastery of all lower levels and non-mastery of all higher levels. A zero score indicates non-mastery of the lowest proficiency level. Scores were excluded only if the actual or imputed mastery level data resulted in a reversal pattern as defined above. The highest proficiency level mastered scores did not necessarily correspond to an interval scale, so in analyzing the data, they should be treated as ordinal.

See table 3-8 for variable names, descriptions, and weighted percentages for the highest proficiency level mastered scores.

Table 3-8. Direct cognitive assessment: highest proficiency level mastered, in percent: School year 2003–04

Variable	Description	Below Level 1	Level 1	Level 2	Level 3	Level 4	Level 5	Level 6	Level 7	Level 8	Level 9
C1R3RPF	C1 RC3 Reading Highest Prof Lvl Mastered	33	36	16	13	1	1	0	0	0	0
C2R3RPF	C2 RC3 Reading Highest Prof Lvl Mastered	6	19	25	37	9	3	1	0	0	0
C3R3RPF	C3 RC3 Reading Highest Prof Lvl Mastered	4	11	17	44	15	6	2	0	0	0
C4R3RPF	C4 RC3 Reading Highest Prof Lvl Mastered	1	2	5	14	34	32	10	2	0	0
C5R3RPF	C5 RC3 Reading Highest Prof Lvl Mastered	0	0	1	1	5	20	26	26	20	1
C6R3RPF	C6 RC3 Reading Highest Prof Lvl Mastered	0	0	0	0	1	9	18	33	34	6
C1R3MPF	C1 RC3 Mathematics Highest Prof Lvl Mastered	8	37	34	18	3	1	0	0	0	0
C2R3MPF	C2 RC3 Mathematics Highest Prof Lvl Mastered	1	15	30	36	15	2	0	0	0	0
C3R3MPF	C3 RC3 Mathematics Highest Prof Lvl Mastered	1	6	22	39	27	5	0	0	0	0
C4R3MPF	C4 RC3 Mathematics Highest Prof Lvl Mastered	0	1	6	22	47	21	3	0	0	0
C5R3MPF	C5 RC3 Mathematics Highest Prof Lvl Mastered	0	0	0	5	19	31	29	14	1	0
C6R3MPF	C6 RC3 Mathematics Highest Prof Lvl Mastered	0	0	0	1	7	18	34	27	12	2

NOTE: Table estimates based on C1_6SC0 panel weight. Table estimates may differ from those reported on earlier user files because of re-estimation of scores on a longitudinal scale that includes fifth grade, and because of sample attrition. See chapter 7, section 7.4 for variable naming conventions.
SOURCE: U.S. Department of Education, National Center for Education Statistics, Early Childhood Longitudinal Study, Kindergarten Class of 1998–99 (ECLS-K), spring 2004.

3.1.4.4 Proficiency Probability Scores

Proficiency probability scores were reported for each of the proficiency levels described above, at each round of data collection. The scores estimate the probability of mastery of each level, and can take on any value from zero to one. An IRT model was employed to calculate the proficiency probability scores, which indicated the probability that a child would have passed a proficiency level, based on the child's whole set of item responses in the content domain. The item clusters were treated as single items for the purpose of IRT calibration, in order to estimate students' probabilities of mastery of each set of skills. The hierarchical nature of the skill sets justified the use of the IRT model in this way.

The proficiency probability scores differed from the highest-level scores in that they could be used to measure gains over time, and from the IRT scale scores in that they targeted specific sets of skills. The proficiency probability scores can be averaged to produce estimates of mastery rates within population subgroups. These continuous measures can provide a close look at individuals' status and change over time. Gains in probability of mastery at each proficiency level allow researchers to study not only the amount of gain in total scale score points but also where along the score scale different children made their largest gains in achievement during a particular time interval. For example, subtracting the reading level 1 probability at time 1 (C1R3RPB1) from the level 1 probability at time 2 (C2R3RPB1) indicates whether a student advanced in mastery of the particular set of level 1 (i.e., letter recognition) skills during this time interval. Thus, students' school experiences can be related to improvements in specific skills.

See tables 3-9 and 3-10 for variable names, descriptions, ranges, weighted means, and standard deviations for the proficiency probability scores in reading and mathematics.

Table 3-9. Fifth-grade direct cognitive assessment: proficiency probability scores—reading: School year 2003–04

Variable	Description	Range of values	Weighted mean	Standard deviation
C1R3RPB1	C1 RC3 Prob1 - Letter Recognition	0 - 1	0.68	0.32
C1R3RPB2	C1 RC3 Prob2 - Beginning Sounds	0 - 1	0.29	0.33
C1R3RPB3	C1 RC3 Prob3 - Ending Sounds	0 - 1	0.16	0.25
C1R3RPB4	C1 RC3 Prob4 - Sight Words	0 - 1	0.03	0.13
C1R3RPB5	C1 RC3 Prob5 - Word in Context	0 - 1	0.01	0.08
C1R3RPB6	C1 RC3 Prob6 - Literal Inference	0 - 1	0.00	0.03
C1R3RPB7	C1 RC3 Prob7 - Extrapolation	0 - 1	0.00	0.01
C1R3RPB8	C1 RC3 Prob8 - Evaluation	0 - 1	0.00	0.01
C1R3RPB9	C1 RC3 Prob9 - Evaluating Non-fiction	0 - 1	0.00	0.00
C2R3RPB1	C2 RC3 Prob1 - Letter Recognition	0 - 1	0.93	0.16
C2R3RPB2	C2 RC3 Prob2 - Beginning Sounds	0 - 1	0.70	0.31
C2R3RPB3	C2 RC3 Prob3 - Ending Sounds	0 – 1	0.51	0.34
C2R3RPB4	C2 RC3 Prob4 - Sight Words	0 – 1	0.15	0.25
C2R3RPB5	C2 RC3 Prob5 - Word in Context	0 – 1	0.06	0.15
C2R3RPB6	C2 RC3 Prob6 - Literal Inference	0 – 1	0.01	0.07
C2R3RPB7	C2 RC3 Prob7 - Extrapolation	0 – 1	0.00	0.02
C2R3RPB8	C2 RC3 Prob8 - Evaluation	0 – 1	0.00	0.01
C2R3RPB9	C2 RC3 Prob9 - Evaluating Non-fiction	0 – 1	0.00	0.00
C3R3RPB1	C3 RC3 Prob1 - Letter Recognition	0 – 1	0.97	0.12
C3R3RPB2	C3 RC3 Prob2 - Beginning Sounds	0 – 1	0.82	0.26
C3R3RPB3	C3 RC3 Prob3 - Ending Sounds	0 – 1	0.66	0.32
C3R3RPB4	C3 RC3 Prob4 - Sight Words	0 – 1	0.27	0.32
C3R3RPB5	C3 RC3 Prob5 - Word in Context	0 – 1	0.12	0.21
C4R3RPB4	C4 RC3 Prob4 - Sight Words	0 – 1	0.74	0.33
C4R3RPB5	C4 RC3 Prob5 - Word in Context	0 – 1	0.46	0.32
C4R3RPB6	C4 RC3 Prob6 - Literal Inference	0 – 1	0.16	0.22
C4R3RPB7	C4 RC3 Prob7 - Extrapolation	0 – 1	0.03	0.11
C4R3RPB8	C4 RC3 Prob8 - Evaluation	0 – 1	0.03	0.06
C4R3RPB9	C4 RC3 Prob9 - Evaluating Non-fiction	0 – 1	0.00	0.00
C5R3RPB1	C5 RC3 Prob1 - Letter Recognition	0 – 1	1.00	0.00
C5R3RPB2	C5 RC3 Prob2 - Beginning Sounds	0 – 1	1.00	0.00
C5R3RPB3	C5 RC3 Prob3 - Ending Sounds	0 – 1	1.00	0.01
C5R3RPB4	C5 RC3 Prob4 - Sight Words	0 – 1	0.98	0.08
C5R3RPB5	C5 RC3 Prob5 - Word in Context	0 – 1	0.90	0.17
C5R3RPB6	C5 RC3 Prob6 - Literal Inference	0 – 1	0.68	0.30
C5R3RPB7	C5 RC3 Prob7 - Extrapolation	0 – 1	0.42	0.38
C5R3RPB8	C5 RC3 Prob8 - Evaluation	0 – 1	0.24	0.20
C5R3RPB9	C5 RC3 Prob9 - Evaluating Non-fiction	0 – 1	0.01	0.04
C6R3RPB1	C6 RC3 Prob1 - Letter Recognition	0 – 1	1.00	0.00
C6R3RPB2	C6 RC3 Prob2 - Beginning Sounds	0 – 1	1.00	0.00
C6R3RPB3	C6 RC3 Prob3 - Ending Sounds	0 – 1	1.00	0.00

See notes at end of table.

Table 3-9. Fifth-grade direct cognitive assessment: proficiency probability scores—reading: School year 2003–04—Continued

Variable	Description	Range of values	Weighted mean	Standard deviation
C6R3RPB4	C6 RC3 Prob4 - Sight Words	0 – 1	1.00	0.01
C6R3RPB5	C6 RC3 Prob5 - Word in Context	0 – 1	0.97	0.07
C6R3RPB6	C6 RC3 Prob6 - Literal Inference	0 – 1	0.87	0.20
C6R3RPB7	C6 RC3 Prob7 - Extrapolation	0 – 1	0.71	0.34
C6R3RPB9	C6 RC3 Prob8 - Evaluation	0 – 1	0.44	0.27
C6R3RPB8	C6 RC3 Prob9 - Evaluating Non-fiction	0 – 1	0.07	0.17

NOTE: Table estimates based on C1_6SC0 panel weight. Table estimates may differ from those reported in earlier user's manuals, the *ECLS-K Psychometric Report for Kindergarten Through First Grade* (NCES 2002–05) (Rock and Pollack 2002b), and the *ECLS-K Psychometric Report for the Third Grade* (NCES 2005-062) (Pollack et al. 2005) because of re-estimation of scores on a longitudinal scale that includes fifth grade, and because of sample attrition. See chapter 7, section 7.4 for variable naming conventions.
SOURCE: U.S. Department of Education, National Center for Education Statistics, Early Childhood Longitudinal Study, Kindergarten Class of 1998–99 (ECLS-K), spring 2004.

Table 3-10. Fifth-grade direct cognitive assessment: proficiency probability scores—mathematics: School year 2003–04

Variable	Description	Range of values	Weighted mean	Standard deviation
C1R3MPB1	C1 RC3 Prob1 - Count, Number, Shape	0 - 1	0.91	0.19
C1R3MPB2	C1 RC3 Prob2 - Relative Size	0 - 1	0.54	0.35
C1R3MPB3	C1 RC3 Prob3 - Ordinality, Sequence	0 - 1	0.20	0.30
C1R3MPB4	C1 RC3 Prob4 - Add/Subtract	0 - 1	0.03	0.12
C1R3MPB5	C1 RC3 Prob5 - Multiply/Divide	0 - 1	0.00	0.04
C1R3MPB6	C1 RC3 Prob6 - Place Value	0 - 1	0.00	0.00
C1R3MPB7	C1 RC3 Prob7 - Rate & Measurement	0 - 1	0.00	0.00
C1R3MPB8	C1 RC3 Prob8 - Fractions	0 - 1	0.00	0.00
C1R3MPB9	C1 RC3 Prob9 - Area and Volume	0 - 1	0.00	0.00
C2R3MPB1	C2 RC3 Prob1 - Count, Number, Shape	0 - 1	0.99	0.05
C2R3MPB2	C2 RC3 Prob2 - Relative Size	0 - 1	0.83	0.24
C2R3MPB3	C2 RC3 Prob3 - Ordinality, Sequence	0 - 1	0.52	0.39
C2R3MPB4	C2 RC3 Prob4 - Add/Subtract	0 - 1	0.16	0.25
C2R3MPB5	C2 RC3 Prob5 - Multiply/Divide	0 - 1	0.02	0.08
C2R3MPB6	C2 RC3 Prob6 - Place Value	0 - 1	0.00	0.01
C2R3MPB7	C2 RC3 Prob7 - Rate & Measurement	0 - 1	0.00	0.00
C2R3MPB8	C2 RC3 Prob8 - Fractions	0 - 1	0.00	0.00
C2R3MPB9	C2 RC3 Prob9 - Area and Volume	0 - 1	0.00	0.00
C3R3MPB1	C3 RC3 Prob1 - Count, Number, Shape	0 - 1	1.00	0.03
C3R3MPB2	C3 RC3 Prob2 - Relative Size	0 - 1	0.92	0.17
C3R3MPB3	C3 RC3 Prob3 - Ordinality, Sequence	0 - 1	0.73	0.34
C3R3MPB4	C3 RC3 Prob4 - Add/Subtract	0 - 1	0.31	0.33
C3R3MPB5	C3 RC3 Prob5 - Multiply/Divide	0 - 1	0.04	0.14
C3R3MPB6	C3 RC3 Prob6 - Place Value	0 - 1	0.00	0.03
C3R3MPB7	C3 RC3 Prob7 - Rate & Measurement	0 - 1	0.00	0.00
C3R3MPB8	C3 RC3 Prob8 - Fractions	0 - 1	0.00	0.00
C3R3MPB9	C3 RC3 Prob9 - Area and Volume	0 - 1	0.00	0.00
C4R3MPB1	C4 RC3 Prob1 - Count, Number, Shape	0 - 1	1.00	0.01
C4R3MPB2	C4 RC3 Prob2 - Relative Size	0 - 1	0.99	0.05
C4R3MPB3	C4 RC3 Prob3 - Ordinality, Sequence	0 - 1	0.95	0.16
C4R3MPB4	C4 RC3 Prob4 - Add/Subtract	0 - 1	0.71	0.31
C4R3MPB5	C4 RC3 Prob5 - Multiply/Divide	0 - 1	0.22	0.29
C4R3MPB6	C4 RC3 Prob6 - Place Value	0 - 1	0.03	0.10
C4R3MPB7	C4 RC3 Prob7 - Rate & Measurement	0 - 1	0.00	0.01
C4R3MPB8	C4 RC3 Prob8 - Fractions	0 - 1	0.00	0.00
C4R3MPB9	C4 RC3 Prob9 - Area and Volume	0 - 1	0.00	0.00
C5R3MPB1	C5 RC3 Prob1 - Count, Number, Shape	0 - 1	1.00	0.00
C5R3MPB2	C5 RC3 Prob2 - Relative Size	0 - 1	1.00	0.00
C5R3MPB3	C5 RC3 Prob3 - Ordinality, Sequence	0 - 1	1.00	0.01
C5R3MPB4	C5 RC3 Prob4 - Add/Subtract	0 - 1	0.97	0.09
C5R3MPB5	C5 RC3 Prob5 - Multiply/Divide	0 - 1	0.76	0.32

See notes at end of table.

Table 3-10. Fifth-grade direct cognitive assessment: proficiency probability scores—mathematics: School year 2003–04—Continued

Variable	Description	Range of values	Weighted mean	Standard deviation
C5R3MPB6	C5 RC3 Prob6 - Place Value	0 - 1	0.41	0.39
C5R3MPB7	C5 RC3 Prob7 - Rate & Measurement	0 - 1	0.13	0.23
C5R3MPB8	C5 RC3 Prob8 - Fractions	0 - 1	0.01	0.05
C5R3MPB9	C5 RC3 Prob9 - Area and Volume	0 - 1	0.00	0.01
C6R3MPB1	C6 RC3 Prob1 - Count, Number, Shape	0 - 1	1.00	0.00
C6R3MPB2	C6 RC3 Prob2 - Relative Size	0 - 1	1.00	0.00
C6R3MPB3	C6 RC3 Prob3 - Ordinality, Sequence	0 - 1	1.00	0.00
C6R3MPB4	C6 RC3 Prob4 - Add/Subtract	0 - 1	0.99	0.03
C6R3MPB5	C6 RC3 Prob5 - Multiply/Divide	0 - 1	0.92	0.19
C6R3MPB6	C6 RC3 Prob6 - Place Value	0 - 1	0.74	0.36
C6R3MPB7	C6 RC3 Prob7 - Rate & Measurement	0 - 1	0.43	0.39
C6R3MPB8	C6 RC3 Prob8 - Fractions	0 - 1	0.14	0.30
C6R3MPB9	C6 RC3 Prob9 - Area and Volume	0 - 1	0.02	0.07

NOTE: Table estimates based on C1_6SC0 panel weight. Table estimates may differ from those reported in earlier user's manuals, the *ECLS-K Psychometric Report for Kindergarten Through First Grade* (NCES 2002–05) (Rock and Pollack 2002b), and the *ECLS-K Psychometric Report for the Third Grade* (NCES 2005–062) (Pollack et al. 2005) because of re-estimation of scores on a longitudinal scale that includes fifth grade, and because of sample attrition. See chapter 7, section 7.4 for variable naming conventions.
SOURCE: U.S. Department of Education, National Center for Education Statistics, Early Childhood Longitudinal Study, Kindergarten Class of 1998–99 (ECLS-K), spring 2004.

The following are some examples of interpretation and use of the proficiency probability scores whose means appear in tables 3-9 and 3-10 are the following:

- At entry to kindergarten about 68 percent (mean probability = .68) of children were proficient at letter recognition (C1R3RPB1).

- The largest gains between spring-kindergarten and spring-first grade were made in reading simple sight words, with 15 percent of children having mastered this skill at the end of kindergarten (C2R3RPB4) compared with 74 percent a year later (C4R3RPB4), 98 percent in third grade (C5R3RPB4) and 100 percent in fifth grade (C6R3RPB4).

- There were only small gains in letter recognition after spring-kindergarten, because most children, 93 percent, knew their letters by this time (C2R3RPB1).

- Children's skills in making inferences based on cues directly stated in text (literal inference) increased between first and third grade, from 16 percent (C4R3RPB6) to 68 percent (C5R3RPB6).

- In spring-third grade, most children had not yet demonstrated understanding of the author's craft or making connections between a problem in the narrative and similar life problems. While 24 percent mastered the evaluation level in third grade (C5R3RPB8), 44 percent demonstrated mastery in fifth grade (C6R3RPB8).

- By spring-fifth grade, nearly all children had mastered basic number concepts (i.e., counting, numbers, and shapes; relative size; and ordinality and sequence) and simple arithmetic operations (i.e., adding/subtracting and multiplying/dividing) (C6R3MPB1 - C6R3MPB5).

- Fourteen percent of children understood interpretation and manipulation of simple fractions (C6R3MPB8) by the spring of fifth grade.

- Two percent of fifth-graders could solve word problems involving area and volume (C6R3MPB9).

Comparisons of subgroups may be made by computing the mean probability for each group at a single point in time, or the mean gain for each group from one time to another. See section 3.1.6 for further discussion of measurement of gain.

3.1.5 Choosing the Appropriate Score for Analysis

Each of the types of scores described earlier measures children's achievement from a slightly different perspective. The choice of the most appropriate score for analysis purposes should be driven by the context in which it is to be used:

- A measure of overall achievement versus achievement in specific skills;

- An indicator of status at a single point in time versus growth over time; and

- A criterion-referenced versus norm-referenced interpretation.

3.1.5.1 Item Response Theory-Based Scores

The scores derived from the IRT model (IRT scale scores, T-scores, proficiency probabilities) were based on all of the child's responses to a subject area assessment. That is, the pattern of right and wrong answers, as well as the characteristics of the assessment items themselves, were used to estimate a point on an ability continuum, and this ability estimate, theta, then provided the basis for criterion-referenced and norm-referenced scores. As noted earlier, estimates of gains and comparisons of achievement across rounds that make use of the IRT-based scales should use re-estimated values for the earlier rounds, not values found on earlier user files (see section 3.1.2).

- **The IRT scale scores** are overall, criterion-referenced measures of status at a point in time. They are useful in identifying cross-sectional differences among subgroups in overall achievement level and provide a summary measure of achievement useful for correlational analysis with status variables, such as demographics, school type, or behavioral measures.

  The IRT scale scores may be used as longitudinal measures of overall growth. However, gains made at different points on the scale have qualitatively different interpretations. For example, children who made gains in recognizing letters and letter sounds are learning very different skills than those who are making the jump from reading words to reading sentences, although the gains in number of scale score points may be the same. Comparison of gain in scale score points is most meaningful for groups that started with similar initial status.

- **The standardized scores (T-scores)** are also overall measures of status at a point in time, but they are norm-referenced rather than criterion-referenced. They do not answer the question, "What skills do children have?" but rather "How do they compare with their peers?" The transformation to a familiar metric with a mean of 50 and standard deviation of 10 facilitates comparisons in standard deviation units. T-score means may be used longitudinally to illustrate the increase or decrease in gaps in achievement among subgroups over time. T-scores are not recommended for measuring individual gains over time. The IRT scale scores or proficiency probability scores are used for that purpose.

- **Proficiency probability scores**, derived from the overall IRT model, are criterion-referenced measures of proficiency in specific skills. Because each proficiency score targets a particular set of skills, they are ideal for studying the details of achievement, rather than the single summary measure provided by the IRT scale scores and T-scores. They are useful as longitudinal measures of change because they show not only the extent of gains but also where on the achievement scale the gains are taking place. Thus, they can provide information on differences in skills being learned by different groups, as well as the relationships with processes, both in and out of school, that correlate with learning specific skills. For example, high socioeconomic status (SES) kindergarten children showed very little gain in the lowest reading proficiency level, letter recognition, because they were already proficient in this skill at kindergarten entry. At the same time, low-SES children made big gains in basic skills, but most had not yet made major gains in reading words and sentences by the end of kindergarten. Similarly, the best readers in fifth grade may be working on learning to make evaluative judgments based on reading material, which would show up as large gains in reading levels 8 and 9. Less skilled readers may show their largest gains between third and fifth grades at levels 5 or 6, literal inference and extrapolation. The proficiency level at which the largest change is taking place is likely to be different for children with different initial status, background, and school setting. Changes in proficiency probabilities over time may be used to identify the process variables that are effective in promoting achievement gains in specific skills.

3.1.5.2 Scores Based on Number Right for Subsets of Items (Non-IRT Based Scores)

The routing test number-right and item cluster scores do not depend on the assumptions of the IRT model. They were derived from item responses on specific subsets of assessment items, rather than estimates based on patterns of overall performance; therefore the values of these scores reported in user files for earlier rounds were not re-estimated. Highest proficiency level mastered also, in theory, was derived from item responses, although a relatively small number of IRT-based estimates were substituted for missing data.

- **Routing test number-right scores** for the fifth-grade reading, mathematics, and science assessments are based on 25, 18, and 21 items respectively (15, 17, and 15 items for the corresponding assessments in third grade, and 20, 16, and 12 items for the kindergarten/first grade reading, mathematics, and general knowledge assessments, respectively). They target specific sets of skills and cover a broad range of difficulty. These scores may be of interest to researchers because they are based on a specific set of assessment items, which was the same for all children who took the assessment.

- **Item cluster scores** in reading (e.g., C5R3RDEC: Decoding Score Gr 3) and science (e.g., C6LIFES5: Life Science Gr 5) are based on a count of the number correct for a small set of items. Users may wish to relate these scores to process variables to get a perspective that is somewhat different from that of the hierarchical levels of skills. However, with only three to seven items in each of these item cluster scores, reliabilities tend to be relatively low.

- **Highest proficiency level mastered** is based on the same sets of items as the proficiency probability scores but consist of a series of dichotomous pass/fail scores, reported as a single highest mastery level. The highest proficiency level mastered should be treated as an ordinal variable. Pass/fail on each of the individual levels in the set is based on whether children were able to answer correctly at least three out of four actual items in each cluster. For about one-third of reading scores and 20 percent of mathematics scores, the item data was supplemented with IRT-based estimates so that the "highest level" scores would not have to be reported as missing data. Analysis of missing data that are not missing at random (i.e., the "missingness" is a consequence of the child's skill level or grade level) requires special treatment in order to avoid misleading results. For further discussion of the imputation process and interpretation of these scores, please see the *ECLS-K Psychometric Report for the Fifth Grade* (NCES 2006-036) (Pollack et al. 2005)

3.1.6 Measuring Gains

This section outlines approaches to measuring gains that rely on multiple criterion-referenced points to identify different patterns of student growth. It describes how analysts might use the proficiency probability scores to address policy questions dealing with subgroup differences in achievement growth over time.

Traditional approaches using a total scale score to measure change, without accounting for initial stutus, may yield uninformative if not misleading results. For example, analysis of the gain in total scale score points in reading between fall- and spring-kindergarten shows an average increase of about 10 points. Subgroup analysis shows nearly identical average gains of about the same magnitude for groups broken down by sex, race/ethnicity, SES, and school type, even though the *mean scores* for the subgroups are quite different. Similarly, each of these groups gained about 7 points, on average, on the mathematics scale during the same time, again starting from very different initial status. The *ECLS-K Psychometric Report for Kindergarten Through First Grade* (NCES 2002–05) (Rock and Pollack 2002b) describes this analysis in more detail.

It would be incorrect to conclude that because different subgroups of children are gaining quantitatively the same number of scale score points, they are learning the same things, or that these gains are qualitatively comparable in any sense. The problem is non-equivalence of scale units: children who gain 10 points at the low end of the scale, for example, by mastering letter recognition and letter sounds, are not learning the same things as more advanced children, who are achieving their 10 point gains by learning to read words and sentences.

The use of adaptive assessments increases the reliability of individual assessment scores by removing the sources of floor and ceiling effects. When assessment forms are matched to children's ability levels, all students have an equal chance to gain on the vertical scale, that is, a scale that spans several time points. Depending on how adaptive the measure is, how the scale is constructed, and how even-handed the educational treatment, one may not observe large differences in individual children's amounts of gain in total scale score points. Individual and group differences in the *amount* of gain given a fairly standard treatment (e.g., a year of schooling) can be relatively trivial compared to individual and group differences in *where* the gains take place. It is more likely that one will see substantial subgroup differences in initial status than in gains, suggesting that the gains being made by individuals at different points on the score scale are qualitatively different. Thus analysis of the total IRT scale score without explicitly taking into consideration where the gain takes place tells only part of the story.

The ECLS-K design utilized adaptive assessments to maximize the accuracy of measurement and minimize floor and ceiling effects, and then to develop an IRT-based vertical scale with multiple criterion-referenced points along that scale. These points, the 9 reading and 9 mathematics proficiency levels described in section 3.1.4, model critical stages in the development of skills. Criterion-referenced points serve two purposes at the individual level: (1) they provide information about changes in each child's mastery or proficiency at *each* level, and (2) they provide information about *where* on the scale the child's gain is taking place. This provides analysts with two options for analyzing achievement gains and relating them to background and process variables. First, gains in probability of proficiency at any level may be aggregated by subgroup, and/or correlated with other variables. Second, the location of maximum gain may be identified for each child by comparing the gains in probability for all of the levels, and focusing on the skills the child is acquiring during a particular time interval.

The probabilities of proficiency at any level may be averaged to estimate the proportion of children mastering the skills marked by that level. For example, the spring-first grade mean for mathematics level 5, "Multiply/Divide," was 0.22, analogous to 22 percent of the first-grade population demonstrating mastery of this set of items. The mean probability at the end of third grade, 0.76, is equivalent to a population mastery rate of 76 percent, with a mastery rate of 92 percent by the end of fifth grade. While most children were making their largest gains between first and third grades at level 5, a small number of children were advancing their skills in solving word problems based on rate and measurement, level 7, and others were still catching up with simple addition and subtraction, level 4. The mastery rate for level 7 rose from near zero at the end of first grade to about 13 percent at the end of third grade, while level 4 mastery advanced from 71 to 97 percent. These proportions, and the average gains in the proportions for this particular skill, would very likely be quite different for subgroups of children defined by various demographic and school-process categories. Similarly, gains at each level between time 1 and time 2 may be computed for individual children and treated as outcome variables in multivariate models that include background and process measures.

Another approach entails computing differences in probabilities of proficiency between time 1 and time 2 for all of the proficiency levels. The largest difference marks the mastery level where the largest gain for a given child is taking place: the "locus of maximum gain." The locus of maximum gain is likely to vary for different subgroups of children categorized according to variables of interest.

Once having identified mutually exclusive groups of children according to the proximity of their gains to each of the critical points on the developmental scale, one can treat the different types of gains as qualitatively different dichotomous outcome measures to be explained by background and process variables. For an example of an analysis using this approach, see section 8.3 of the *ECLS-K Psychometric Report for Kindergarten through First Grade* (NCES 2002-05) (Rock and Pollack 2002).

Each different analytical approach provides a different perspective with respect to understanding student growth. While comparisons of scale score means may be used to capture information about children at a single point in time, analysis of gain in probability of proficiency is more likely to provide useful information about the contribution of background and process variables to gains in achievement over time. Examples of these approaches can be found in Rock and Pollack (2002a).

Another important issue to be considered in analyzing achievement scores and gains is assessment timing: children's age at first assessment, assessment dates, and the time interval between successive assessments. This issue is most relevant in the early years, kindergarten and first grade. Assessment dates ranged from September to November for fall data collections, and from March to June for spring rounds. At kindergarten entry, boys, on average, tend to be older than girls. Children assessed in November of their kindergarten year may be expected to have an advantage over children assessed in the first days or weeks of school. Substantial differences in intervals between assessments may also affect analysis of gain scores. Children assessed in September and June of kindergarten or first grade have more time to learn skills than children assessed in November and March. These differences in intervals may have a relatively small impact on analysis results for long time intervals, such as measuring gains from spring-third grade to spring-fifth grade, but may be more important within grade, especially fall- to spring-kindergarten. In designing an analysis plan, it is important to consider whether and how differences in ages, assessment dates and intervals may affect the results, to look at relationships between these factors and other variables of interest, and to compensate for differences if necessary. Walston and West (2004) address the issue in their report on full-day and half-day kindergarten.

3.1.7 Reliability

Reliability statistics assess consistency of measurement, in other words, the extent to which a set of test items is related to each other and to the score scale as a whole. For tests of equal length, reliability estimates can be expected to be higher for sets of items that are closely related to the underlying construct than for tests with more diversity of content. Conversely, for tests with similar levels of diversity in content, reliabilities tend to be higher for longer tests compared with shorter tests. In general, the most diverse subject, science, had lower reliability coefficients than reading and mathematics. Reliabilities for scores using the greatest number of test items, the IRT ability estimates that are based on all items taken by each child, were highest. Reliabilities for scores based on the fewest items, the item cluster scores in reading and science, were lowest. Reliability statistics appropriate for each type of score were computed for each subject area for each round of data collection.

For the IRT-based scores, the reliability of the overall ability estimate, theta, is based on the variance of repeated estimates of theta compared with total sample variance. These reliabilities, ranging from 87 percent to 96 percent, apply to all of the scores derived from the theta estimate, namely, the IRT scale scores, T-scores, and proficiency probabilities. Alpha coefficients for the routing test number correct ranged from 79 percent to 88 percent for the fifth-grade assessment forms. The third-grade reading alpha is somewhat lower than for kindergarten/first grade and fifth grade, due at least in part to the third-grade routing test having fewer items (15) than the 20-25 items in the other years. Conversely, the alpha coefficient for the mathematics routing test was slightly higher in third and fifth grades, due at least in part to an increase in the number of mathematics routing items, from 16 in the K-1 form to 17 in third grade to 18 in fifth grade. Split-half reliabilities were computed for the item cluster scores in reading and science. These reliabilities tended to be higher for the reading clusters (.60 to .70) than for the science scores (.41 to .64). The difference in internal consistency statistics is due to the reading items being essentially replications of the same or similar tasks, while the science items had a greater diversity of content.

It was not possible to apply standard measures of reliability to the "highest proficiency mastered" score, for the following reasons. The score is not a set of items replicating the same or similar tasks, so an internal consistency measure such as split-half reliability or alpha coefficient cannot be computed. Nor can the reliability be evaluated based on the variance of repeated estimates of overall ability that was appropriate for the IRT-based scores.

The definition of reliability–consistency of measurement under different circumstances–suggested an appropriate way to assess the reliability of the "highest proficiency level mastered" score. The score denoting the highest level mastered reduces the series of pass/fail scores on the hierarchical set of proficiency levels to a single score. For example, a student demonstrating mastery of the first five reading levels but not the remaining four would be said to have a "highest proficiency mastered" score of five. The question to be answered by a reliability estimate is how likely it would be that the same highest level score would be obtained under other circumstances. In this case, the other circumstances available are not a parallel set of items, but two different methods of arriving at the score. A student's highest level mastered could be determined on the basis of actual item response data alone for about 67 percent of the reading and 80 percent of the mathematics scores. Alternatively, IRT ability estimates and item parameters could be used to generate pass/fail scores, and the composite highest level scores, for these same students. The percent of cases for which these two different methodologies result in identical or adjacent "highest level mastered" scores can be considered to be a reliability estimate. The high level of exact-plus-adjacent agreement between the methods indicates that the IRT approach supports the use of the highest level score sufficiently well for use in aggregate statistics.

Tables 3-11 through 3-14 present the reliability statistics for all of the assessment scores.

Table 3-11. Reliability of item response theory-based scores: IRT scale scores, T-scores, proficiency probabilities, by round of data collection and domain: School years 1998–1999, 1999–2000, 2001–02, and 2003–04

Domain	Fall-kindergarten	Spring-kindergarten	Fall-first grade	Spring-first grade	Spring-third grade	Spring-fifth grade
Reading	.91	.93	.95	.96	.93	.93
Mathematics	.89	.91	.92	.92	.94	.94
Science	†	†	†	†	.88	.87

† Not applicable.
NOTE: Approximately 90 percent of the children interviewed were in fifth grade during the 2003-2004 school year, 9 percent were in fourth grade, and less than 1 percent were in third or other grades. Table estimates may differ from those reported in earlier user's manuals, the *ECLS-K Psychometric Report for Kindergarten Through First Grade* (NCES 2002–05) (Rock and Pollack 2002b), and the *ECLS-K Psychometric Report for the Third Grade* (NCES 2005–062) Pollack et al. 2005) because of re-estimation of scores on a longitudinal scale that includes fifth grade.
SOURCE: U.S. Department of Education, National Center for Education Statistics, Early Childhood Longitudinal Study, Kindergarten Class of 1998–99 (ECLS-K), fall 1998, spring 1999, fall 1999, spring 2000, spring 2002, and spring 2004.

Table 3-12. Reliability of routing test number correct (alpha coefficient), by round of data collection and domain: School years 1998–1999, 1999–2000, 2001–02, and 2003–04

Domain	Fall-kindergarten	Spring-kindergarten	Fall-first grade	Spring-first grade	Spring-third grade	Spring-fifth grade
Reading	.86	.88	.88	.86	.75	.88
Mathematics	.78	.81	.83	.80	.86	.88
Science	†	†	†	†	.75	.79

† Not applicable.
NOTE: Approximately 90 percent of the children interviewed were in fifth grade during the 2003-2004 school year, 9 percent were in fourth grade, and less than 1 percent were in third or other grades.
SOURCE: U.S. Department of Education, National Center for Education Statistics, Early Childhood Longitudinal Study, Kindergarten Class of 1998–99 (ECLS-K), fall 1998, spring 1999, fall 1999, spring 2000, spring 2002, and spring 2004.

Table 3-13. Split-half reliability of item-cluster-based scores, by round of data collection and cluster: School years 1998–1999, 1999–2000, 2001–02, and 2003–04

Cluster score	Fall-kindergarten	Spring-kindergarten	Fall-first grade	Spring-first grade	Spring-third grade	Spring-fifth grade
Print Familiarity	.70	.68	.68	.60	†	†
Decoding Score	†	†	†	†	.67	†
Life Science Gr.3	†	†	†	†	.59	.59
Phys. Science Gr.3	†	†	†	†	.49	.41
Earth Science Gr.3	†	†	†	†	.46	.52
Life Science Gr.5	†	†	†	†	†	.64
Phys. Science Gr.5	†	†	†	†	†	.43
Earth Science Gr.5	†	†	†	†	†	.62

† Not applicable.
NOTE: Approximately 90 percent of the children interviewed were in fifth grade during the 2003–04 school year, 9 percent were in fourth grade, and less than 1 percent were in third or other grades.
SOURCE: U.S. Department of Education, National Center for Education Statistics, Early Childhood Longitudinal Study, Kindergarten Class of 1998–99 (ECLS-K), fall 1998, spring 1999, fall 1999, spring 2000, spring 2002, and spring 2004.

Table 3-14. Percent agreement of highest proficiency level mastered score, by round of data collection: School years 1998–1999, 1999–2000, 2001–02, and 2003–04

Domain	Fall-kindergarten	Spring-kindergarten	Fall-first grade	Spring-first grade	Spring-third grade	Spring-fifth grade
Reading						
Exact Agreement	63	54	55	55	50	51
Exact + Off by 1	96	94	94	95	95	95
Mathematics						
Exact Agreement	54	51	52	57	56	55
Exact + Off by 1	97	95	96	97	97	97

NOTE: Approximately 90 percent of the children interviewed were in fifth grade during the 2003-2004 school year, 9 percent were in fourth grade, and less than 1 percent were in third or other grades. Table estimates may differ from those reported in earlier user's manuals, the *ECLS-K Psychometric Report for Kindergarten Through First Grade* (NCES 2002–05) (Rock and Pollack 2002b), and the *ECLS-K Psychometric Report for the Third Grade* (NCES 2005–062) (Pollack et al. 2005) because of re-estimation of scores on a longitudinal scale that includes fifth grade.
SOURCE: U.S. Department of Education, National Center for Education Statistics, Early Childhood Longitudinal Study, Kindergarten Class of 1998–99 (ECLS-K), fall 1998, spring 1999, fall 1999, spring 2000, spring 2002, and spring 2004.

3.1.8 Validity

Evidence for the validity of the direct cognitive assessments was derived from several sources. A review of national and state performance standards, comparison with state and commercial assessments, the judgments of curriculum experts and teachers all provided input to test specifications. In addition, comparing the reading and mathematics field-test item pool scores with those obtained from an established instrument provided validity information.

The ECLS-K test specifications were derived from a variety of sources. For the third- through fifth-grade assessments, national and state performance standards in each of the domains were examined. The scope and sequence of materials from state assessments, as well as from major publishers, were also considered. The resulting ECLS-K fourth-grade frameworks are similar to the NAEP fourth-grade frameworks, with some differences due to ECLS-K formatting and administration constraints. The fourth-grade frameworks were modified for third and fifth grades (and for the earlier K-1 forms). An expert panel of early elementary school educators, including curriculum specialists in the subject areas and teachers at the targeted grade levels from different regions of the country, examined the pool of items and the recommended allocations. The assessment specifications indicated target percentages for content strands within each of the subject areas. These percentages were matched as closely as possible in developing the field-test assessment item pool as well as in selecting items for the fifth-grade assessment forms. Some compromises in matching target percentages were necessary to satisfy constraints related to other issues, including linking to K-1 and third-grade scales, avoiding floor and ceiling effects, and field-test item performance. This was especially true for the reading assessment, whose structure, i.e., several questions based on each reading passage, placed an additional constraint on the selection of items to match content strands. Experts in each of the subject areas then reviewed the proposed fifth-grade forms for appropriateness of content and relevance to the assessment framework.

An additional method of evaluating the construct validity of the reading and mathematics assessments was addressed by the inclusion of the Woodcock-McGrew-Werder Mini-Battery of Achievement (MBA) in the spring 2002 field test of fifth-grade items. Selected field-test forms that included reading sections also included the MBA reading test, while the MBA mathematics test was administered along with field-test mathematics forms. Correlations were computed for the MBA scores with the theta estimates based on ECLS-K field-test responses. Test scores can be related to other measures only to the extent that they are consistent within themselves. Generally, a correlation between two variables cannot exceed the square root of the reliability of either variable. Reliabilities for the MBA were computed both with not-administered and omitted items treated as missing, and with these items treated as incorrect. The correlations of MBA with ECLS-K measures were quite close to the square roots of the reliabilities, indicating that the two assessments were measuring closely related skills. The correlations are presented in table 3-15.

Table 3-15. Validity coefficients for reading and mathematics field test item pools: School year 2003–04

Computation	Reading	Mathematics
Reliability of MBA (computed both ways)	.73 and .77	.61 and .68
Square root of reliability	.85 and .88	.78 and .82
Correlation of MBA x ECLS-K grade 5 field test item pool	.73	.80

SOURCE: U.S. Department of Education, National Center for Education Statistics, Early Childhood Longitudinal Study, Kindergarten Class of 1998–99 (ECLS-K), spring 2004. First reliability statistic is computed with not-administered and omitted items treated as missing; second statistic treats these items as incorrect.

3.2 Indirect Cognitive Assessment

The Academic Rating Scale (ARS) was developed for the ECLS-K to measure teachers' evaluations of students' academic achievement in three domains: language and literacy (reading and writing), science, and mathematical thinking. In earlier grades, teachers also rated students' achievement in a fourth domain: social studies. Teachers rated the child's skills, knowledge, and behaviors on a scale from "Not Yet" to "Proficient" (see table 3-16). If a skill, knowledge, or behavior had not been introduced into the classroom yet, the teacher coded that item as N/A (not applicable). In fifth grade, many schools are departmentalized so different teachers may be rating the student on science and mathematical thinking. Students were rated on either their language and literacy and mathematical thinking, or their language and literacy and science. The differences between the direct and indirect cognitive assessments, and the scores available, are described here. For a discussion of the content areas of the ARS, see chapter 2, section 2.3.1.

Table 3-16. Academic Rating Scale response scale: School year 2003–04

Value	Response	Description
1	Not yet:	Child has not yet demonstrated skill, knowledge, or behavior.
2	Beginning:	Child is just beginning to demonstrate skill, knowledge, or behavior but does so very inconsistently.
3	In progress:	Child demonstrates skill, knowledge, or behavior with some regularity but varies in level of competence.
4	Intermediate:	Child demonstrates skill, knowledge, or behavior with increasing regularity and average competence but is not completely proficient.
5	Proficient:	Child demonstrates skill, knowledge, or behavior competently and consistently.
N/A	Not applicable:	Skill, knowledge, or behavior has not been introduced in classroom setting.

SOURCE: U.S. Department of Education, National Center for Education Statistics, Early Childhood Longitudinal Study, Kindergarten Class of 1998–99 (ECLS-K), spring 2004.

3.2.1 Comparison to Direct Cognitive Assessment

The ARS was designed both to overlap and to augment the information gathered through the direct cognitive assessment battery. Although the direct and indirect instruments measure children's skills and behaviors within the same broad curricular domains with some intended overlap, several of the constructs they were designed to measure differ in significant ways. Most importantly, the ARS includes items designed to measure both the process and products of children's learning in school, whereas the direct cognitive battery is more limited. Because of time and space limitations, the direct cognitive battery is less able to measure the process of children's thinking, including the strategies they use to read, solve mathematical problems, or investigate a scientific phenomenon. The ARS language and literacy questionnaire collects information on children's written composition, an area also not assessed on the direct measure.

The criterion-referenced indirect measures on the ARS are targeted to the specific grade level of the student and draw upon the daily observations made by teachers of the students in their class.

3.2.2 Rasch Scores Available for the Academic Rating Scale

A Rasch analysis was used to create measures of the reported performance of students on a hierarchy of skills, knowledge, and behavior. The Rasch Rating Scale model uses the pattern of ratings on items to determine an estimate of the difficulty of each item and to place each student on an interval scale set with a minimum score of one and a maximum score of five. The Rasch analysis showed that the reliability of the estimates of child ability was very high for all domains of the ARS (see table 3-17).

Table 3-17. Person separation reliability statistics for the Rasch-based score, by category: School year 2003–04

Category	Grade 5
ARS Language and Literacy	.95
ARS Mathematical Thinking	.92
ARS Science	.94

NOTE: Person separation reliability is a measure of internal consistency and is analogous to the KR-20 and Cronbach's alpha. Person separation is the ratio of the adjusted standard deviation to the root mean standard error: ((S. D. of Measure)^2 – (RMSE)^2) / RMSE. Person separation reliability is the square of this separation statistic divided by one plus the separation squared (Linacre and Wright, 2000).
SOURCE: U.S. Department of Education, National Center for Education Statistics, Early Childhood Longitudinal Study, Kindergarten Class of 1998–99 (ECLS-K), spring 2004.

As mentioned, the ARS scores are scaled to have a low value of one and a high value of five to correspond to the 5-point rating scale that teachers used in rating children on these items. The item difficulties and student scores are placed on a common scale. Students have a high probability of receiving a high rating on items whose difficulty is below their scale score, and a lower probability of receiving a high rating on items above their scale score. Therefore, the scores children receive on the ARS subscales should not be interpreted as mean scores, but as the child's relative probability of success with the items. Students who received maximum ratings on all the items or minimum ratings on all the items are assigned an estimated score.

The variable names, descriptions, value ranges, weighted means, and standard deviations for the fifth-grade (T6) ARS scores are shown in table 3-18. The description for each variable in the tables begins with a "T," indicating that it is a teacher questionnaire child-level variable. The items and the metric for the fifth-grade ARS are different from the ARS ratings in earlier rounds of data collection, so the scores are not directly comparable to those for kindergarten, first, and third grades. The students' scores are calculated relative to the item difficulty. With different items used across the grades and separate calibrations performed, the size of the metric differs from one grade to another.

On the ARS, teachers indicated “not applicable” when the knowledge, skill, or behavior had not been introduced to the classroom. Because some children might have already had this skill (from home or other opportunities for learning), the “not applicable” ratings were treated as missing data and the child’s score was estimated based on the items on which the child was rated. Although the Rasch program estimates scores for all children based on the information provided, the file includes only the scores of children who had more than 60 percent of the items in a scale rated. In other words, if 40 percent or more of the items in a scale were not rated, then the score was set to missing. Fewer than 1 percent of literacy, fewer than 2 percent of the mathematics scores, and fewer than 6 percent of science scores, failed to meet the completeness criterion.

Table 3-18. Spring-fifth grade Academic Rating Scale: variable names, descriptions, ranges, weighted means, and standard deviations: School year 2003–04

Variable name	Description	Range of values	Weighted mean	Standard deviation
T6ARSLIT	T6 Literacy ARS Score	1 - 5	3.35	0.84
T6ARSMAT	T6 Mathematics ARS Score	1 - 5	3.36	0.71
T6ARSSCI	T6 Science ARS Score	1 - 5	3.27	0.88

NOTE: Table estimates based on C6CW0 weight. See chapter 7, section 7.4 for variable naming conventions.
SOURCE: U.S. Department of Education, National Center for Education Statistics, Early Childhood Longitudinal Study, Kindergarten Class of 1998–99 (ECLS-K), spring 2004.

Tables 3-19 to 3-21 provide the estimates of difficulty for each of the items. Higher values mean that teachers rated fewer students as proficient on those items. Students would have a greater than 50 percent probability of receiving ratings of “5” on items below their ability level. Tables are provided for fifth-grade items.

Table 3-19. Spring-fifth grade Academic Rating Scale language and literacy item difficulties (arranged in order of difficulty): School year 2003–04

Item difficulty	Item number and abbreviated content
2.76	Q2. Understands and interprets a story or other text read aloud
2.77	Q4. Reads fluently
2.86	Q1. Conveys ideas clearly when speaking
2.96	Q5. Reads and comprehends expository text
3.04	Q6. Composes multi-paragraph stories/reports with an understandable beginning, middle, and end
3.06	Q3. Uses various strategies to gain information
3.07	Q8. Makes mechanical corrections when reviewing a rough draft,
3.22	Q7. Rereads and reflects on writing, making changes to clarify or elaborate

SOURCE: U.S. Department of Education, National Center for Education Statistics, Early Childhood Longitudinal Study, Kindergarten Class of 1998–99 (ECLS-K), spring 2004.

Table 3-20. Spring-fifth grade Academic Rating Scale mathematical thinking item difficulties (arranged in order of difficulty): School year 2003–04

Item difficulty	Item number and abbreviated content
2.27	Q1. Subtracts numbers that require regrouping
2.69	Q6. Shows understanding of place value
2.83	Q9. Divides multi-digit problems with remainders in the quotient
2.99	Q7. Makes reasonable estimates of quantities and checks answers
3.04	Q5. Uses measuring tools accurately
3.04	Q8. Uses strategies to multiply and divide
3.14	Q4. Recognizes properties of shapes such as area, perimeter, and volume
3.15	Q2. Reduces fractions to lowest denominator
3.19	Q10. Demonstrates algebraic thinking
3.29	Q3. Demonstrates money management skills

SOURCE: U.S. Department of Education, National Center for Education Statistics, Early Childhood Longitudinal Study, Kindergarten Class of 1998–99 (ECLS-K), spring 2004.

Table 3-21. Spring-fifth grade Academic Rating Scale science item difficulties (arranged in order of difficulty): School year 2003–04

Item difficulty	Item number and abbreviated content
2.72	Q3. Classifies and compares living and non-living things in different ways
2.82	Q7. Demonstrates understanding of life science concepts
2.91	Q5. Applies scientific principles to experiences of daily living
2.91	Q1. Makes logical predictions when conducting scientific investigations
2.95	Q4. Forms explanations and conclusions based on observation and investigation
3.03	Q2. Communicates scientific information
3.09	Q6. Demonstrates understanding of physical science concepts

SOURCE: U.S. Department of Education, National Center for Education Statistics, Early Childhood Longitudinal Study, Kindergarten Class of 1998–99 (ECLS-K), spring 2004.

The ARS scale was designed to provide information on children's abilities at a given point in time, not necessarily over time. In addition, although some item stems are similar to those used in the teacher questionnaires in previous grades, the actual items include performance criteria that increase in difficulty from one time to the next. Moreover, the ARS scores are placed on different metrics relative to the item difficulty in a given grade. Therefore, change scores cannot be calculated between time points. However, covariance models may be used to compare teacher's ratings of performance in different grades. Before using these variables in such analyses, the distribution of the samples should be assessed to determine if the assumption of normal distribution is met.

Tables 3-22 to 3-24 provide standard errors (SE) for each of the Rasch scores for fifth grade. The "Score" column is the sum of the raw score ratings. "Measure" is the Rasch-based score. The column labeled “SE” is the corresponding standard error of measurement for those scores. These standard errors can be used in analytic models to correct for the heteroskedasticity of scores.

Table 3-22. Spring-fifth grade Academic Rating Scale language and literacy standard errors: School year 2003–04

Score	Measure	SE	Score	Measure	SE	Score	Measure	SE
8	1.00E	.42	19	2.41	.14	30	3.58	.15
9	1.30	.24	20	2.51	.15	31	3.69	.16
10	1.49	.19	21	2.60	.15	32	3.81	.16
11	1.63	.16	22	2.71	.15	33	3.92	.16
12	1.74	.15	23	2.82	.16	34	4.03	.16
13	1.84	.15	24	2.94	.16	35	4.14	.16
14	1.94	.15	25	3.05	.16	36	4.25	.16
15	2.03	.15	26	3.17	.16	37	4.37	.17
16	2.13	.15	27	3.27	.15	38	4.50	.19
17	2.22	.15	28	3.38	.15	39	4.70	.24
18	2.32	.14	29	3.48	.15	40	5.00E	.42

NOTE: E=estimated extreme score.
SOURCE: U.S. Department of Education, National Center for Education Statistics, Early Childhood Longitudinal Study, Kindergarten Class of 1998–99 (ECLS-K), spring 2004.

Table 3-23. Spring-fifth grade Academic Rating Scale mathematical thinking standard errors: School year 2003–04

Score	Measure	SE	Score	Measure	SE	Score	Measure	SE
10	1.00E	.54	24	2.57	.12	38	3.44	.14
11	1.38	.31	25	2.62	.12	39	3.51	.15
12	1.62	.23	26	2.68	.13	40	3.59	.15
13	1.77	.20	27	2.73	.13	41	3.66	.15
14	1.89	.18	28	2.79	.13	42	3.74	.15
15	1.99	.16	29	2.84	.13	43	3.83	.16
16	2.08	.15	30	2.90	.13	44	3.92	.16
17	2.15	.14	31	2.96	.13	45	4.91	.17
18	2.22	.14	32	3.03	.14	46	4.12	.18
19	2.29	.13	33	3.09	.14	47	4.24	.20
20	2.35	.13	34	3.16	.14	48	4.39	.23
21	2.40	.13	35	3.23	.14	49	4.63	.30
22	2.46	.13	36	3.29	.14	50	5.00E	.53
23	2.51	.12	37	3.36	.14	†	†	†

† Not applicable
NOTE: E=estimated extreme score.
SOURCE: U.S. Department of Education, National Center for Education Statistics, Early Childhood Longitudinal Study, Kindergarten Class of 1998–99 (ECLS-K), spring 2004.

Table 3-24. Spring-fifth grade Academic Rating Scale science standard errors: School year 2003–04

Score	Measure	SE	Score	Measure	SE	Score	Measure	SE
7	1.00E	.43	17	2.37	.15	27	3.60	.19
8	1.30	.25	18	2.48	.16	28	3.76	.20
9	1.51	.19	19	2.59	.16	29	3.93	.19
10	1.64	.17	20	2.71	.17	30	4.07	.18
11	1.76	.16	21	2.84	.17	31	4.21	.17
12	1.87	.15	22	2.97	.17	32	4.34	.18
13	1.97	.15	23	3.09	.17	33	4.49	.19
14	2.07	.15	24	3.21	.17	34	4.69	.25
15	2.17	.15	25	3.33	.17	35	5.00E	.43
16	2.27	.15	26	3.46	.17	†	†	†

† Not applicable
NOTE: E=estimated extreme score.
SOURCE: U.S. Department of Education, National Center for Education Statistics, Early Childhood Longitudinal Study, Kindergarten Class of 1998–99 (ECLS-K), spring 2004.

The teacher ratings do not represent a systematic national sample of teachers. Each set of teacher ratings is linked to a sampled child, and teachers were asked to rate as many ECLS-K sample children as they had in class.

3.3 Teacher Social Rating Scale

The teacher Social Rating Scale (SRS) asked fifth-grade teachers to report how often students exhibited certain social skills and behaviors. Teachers rated individual students as part of a self-administered questionnaire. (In the kindergarten and first-grade rounds of data collection, SRS questions had been asked of both teachers and parents.) Teachers used a frequency scale (see table 3-25) to report on how often the student demonstrated the behavior described. See chapter 2, sections 2.3 and 2.3.2 for additional information on the teacher SRS instrument.

Table 3-25. Social Rating Scale response scale: School year 2003–04

Value	Response	Description
1	Never	Student never exhibits this behavior.
2	Sometimes	Student exhibits this behavior occasionally or sometimes.
3	Often	Student exhibits this behavior regularly but not all the time.
4	Very often	Student exhibits this behavior most of the time.
N/O	No opportunity	No opportunity to observe this behavior.

SOURCE: U.S. Department of Education, National Center for Education Statistics, Early Childhood Longitudinal Study, Kindergarten Class of 1998–99 (ECLS-K), spring 2004.

Five teacher SRS scales were developed based on responses to the scale. The scale score on each SRS scale is the mean rating on the items included in the scale. Scores were computed only if the student was rated on at least two-thirds of the items in that scale. The five social skill teacher scales are as follows: approaches to learning, self-control, interpersonal skills, externalizing problem behaviors, and internalizing problem behaviors. Although all 26 fifth-grade SRS items were the same as in third grade, and 24 of the 26 were in the kindergarten/first-grade instrument, teachers may place different interpretations on the meaning of the items at different time points. Therefore, these scores would be most appropriately used as covariates rather than as change scores.

Two items were added to the third- and fifth-grade scales due to a high number of maximum scores on the third-grade field test of these items. One item was added to the externalizing problem behavior scale ("child talks during quiet study time"). The second item "child follows classroom rules" was added to the SRS in an attempt to increase variance in the self-control scale. Analysis of the item responses indicated that it contributed strongly to the approaches to learning scale, increasing the variance and reliability of that scale. Thus, this item is included in the approaches to learning scale.

In third grade, examination of the responses suggested a different perception of a student's self-control and interpersonal social abilities. The self-control scale includes items on control of attention as well as control of emotions and behavior in interactions. Third-grade students who were rated higher on self-control were also rated higher on interpersonal skills that involved peers. Thus, in addition to the self-control and interpersonal social abilities scale scores, a peer relations scale score was included. This additional scale combines responses on both the interpersonal and self-control scale items that relate to peers. The two items added in third grade were retained in the fifth-grade instrument and scales, and the same peer relations scale was also computed.

Variable names for the teacher scores, descriptions, ranges, weighted means, and standard deviations for these scales are shown in table 3-26. About 90 percent of the children whose teachers provided social ratings data were in fifth grade during the round 6 data collection, 9 percent were in fourth grade, and nearly all of the others were in third grade. Numbers in the table are for fifth-graders, with scores for children who at round 6 were still in third or fourth grade shown in parentheses. The number of children who had advanced to sixth grade by round 6 was too small to be analyzed separately.

Table 3-26. Teacher Social Rating Scale scores: variable names, descriptions, ranges, weighted means, and standard deviations: School year 2003–04

Variable	Description	Range of values	Weighted mean	Standard deviation
T6LEARN	T6 Approaches to Learning	1 - 4	3.0 (2.7)	0.7 (0.7)
T6CONTRO	T6 Self-Control	1 - 4	3.2 (3.0)	0.6 (0.6)
T6INTERP	T6 Interpersonal	1 - 4	3.1 (2.8)	0.6 (0.7)
T6EXTERN	T6 Externalizing Problem Behaviors	1 - 4	1.7 (1.9)	0.6 (0.7)
T6INTERN	T6 Internalizing Problem Behaviors	1 - 4	1.7 (1.8)	0.6 (0.6)
T6SCINT	T6 Combo of Self-Control & Interpersonal (Peer Relations)	1 - 4	3.1 (2.9)	0.6 (0.6)

NOTE: Table estimates based on C6CW0 weight. Numbers outside of parentheses represent children in fifth grade at the time of assessment. Numbers within parentheses represent third and/or fourth-graders at the time of assessment. See chapter 7, section 7.4 for variable naming conventions.
SOURCE: U.S. Department of Education, National Center for Education Statistics, Early Childhood Longitudinal Study, Kindergarten Class of 1998–99 (ECLS-K), spring 2004.

Mean teacher ratings for fifth-graders were nearly identical to the mean ratings for third-graders on the same scales two years earlier. In both third and fifth grades, children who were below modal grade level scored about one-third to more than one-half standard deviation lower than on-grade-level children on scales measuring positive behaviors, and about one-third of a standard deviation higher on measures of problem behaviors.

The split-half reliabilities for the teacher SRS scales are high (see table 3-27). Reliabilities are nearly identical for fifth-graders in round 6 and for children who were not yet in fifth grade, so the table contains only reliabilities for the whole sample. These reliabilities are also nearly identical to round 5 results.

Table 3-27. Split-half reliability for the teacher Social Rating Scale scores: School year 2003–04

Variable	Description	Split-half reliability
T6LEARN	T6 Approaches to Learning	.91
T6CONTRO	T6 Self-control	.79
T6INTERP	T6 Interpersonal	.88
T6EXTERN	T6 Externalizing Problem Behaviors	.89
T6INTERN	T6 Internalizing Problem Behaviors	.77
T6SCINT	T6 Combo of Self-Control & Interpersonal (Peer Relations)	.92

NOTE: See chapter 7, section 7.4 for variable naming conventions.
SOURCE: U.S. Department of Education, National Center for Education Statistics, Early Childhood Longitudinal Study, Kindergarten Class of 1998–99 (ECLS-K), spring 2004.

Care should be taken when entering these scales into the same analysis due to problems of multicollinearity. The intercorrelations among the five SRS factors (excluding the combined peer relations scale) are high (see table 3-28). The factor intercorrelations with the internalizing problem behaviors are the lowest. The absolute values of correlations among the teacher SRS factors range from .31 to .96 for all round 6 children, with nearly identical patterns on most factors for fifth-graders and for children who were still in third or fourth grade. The exception to this similarity is the Internalizing Problem Behaviors factor, which was more strongly correlated with other factors for fifth-graders compared with third- and fourth-graders (although in both groups it had the lowest intercorrelations).

Table 3-28. Intercorrelations among the teacher Social Rating Scale scores: School year 2003–04

Measures[1]	T6LEARN	T6CONTRO	T6INTERP	T6EXTERN	T6INTERN	T6SCINT
T6LEARN	1.00					
T6CONTRO	.69	1.00				
T6INTERP	.72	.81	1.00			
T6EXTERN	-.60	-.72	-.63	1.00		
T6INTERN	-.40	-.34	-.38	.31	1.00	
T6SCINT	.75	.93	.96	-.70	-.38	1.00

[1] T6LEARN = T6 Approaches to Learning
T6CONTRO = T6 Self-control
T6INTERP = T6 Interpersonal
T6EXTERN = T6 Externalizing Problem Behaviors
T6INTERN = T6 Internalizing Problem Behaviors
T6SCINT = T6 Combo of Self-Control & Interpersonal (Peer Relations)

NOTE: See chapter 7, section 7.4 for variable naming conventions.

SOURCE: U.S. Department of Education, National Center for Education Statistics, Early Childhood Longitudinal Study, Kindergarten Class of 1998–99 (ECLS-K), spring 2004.

3.4 Self-Description Questionnaire

In third grade and again in fifth grade, students rated their perceived competence and interest in reading, mathematics, and all school subjects. They also rated their perceived competence and popularity with peers and reported on problem behaviors with which they might struggle. The "Externalizing Problems" scale included questions about anger and distractibility, while "Internalizing Problems" scale included items on sadness, loneliness, and anxiety. For further description of the Self-Description Questionnaire (SDQ) see chapter 2, section 2.1.1. Students rated whether each item was "not at all true," "a little bit true," "mostly true," or "very true." Five scales were produced from the SDQ items. The scale scores on all SDQ scales represent the mean rating of the items included in the scale. Students who responded to the SDQ answered virtually all of the questions, so treatment of missing data was not an issue. As with most measures of social-emotional behaviors, the distributions on these scales are skewed (negatively skewed for the positive social behavior scales, and positively skewed for the problem behavior scales). The reliability is lower for scales with only six items (see table 3-29). The means and standard deviations for the scales are presented in table 3-30.

Table 3-29. Self-Description Questionnaire scale reliabilities (alpha coefficient): School year 2003–04

Variable	Description	Number of items	Alpha coefficient
C6SDQRDC	C6 SDQ Prcvd Interest[1]/Competence - Reading	8	.90
C6SDQMTC	C6 SDQ Prcvd Interest/Competence - Math	8	.92
C6SDQSBC	C6 SDQ Prcvd Interest/Competence - All Sbj	6	.83
C6SDQPRC	C6 SDQ Prcvd Interest/Competence - Peer Rl	6	.82
C6SDQEXT	C6 SDQ Externalizing Problems	6	.78
C6SDQINT	C6 SDQ Internalizing Problems	8	.79

[1] "Prcvd Interest" = Perceived Interest.
NOTE: See chapter 7, section 7.4 for variable naming conventions.
SOURCE: U.S. Department of Education, National Center for Education Statistics, Early Childhood Longitudinal Study, Kindergarten Class of 1998–99 (ECLS-K), spring 2004.

Table 3-30. Self-Description Questionnaire scale range, mean, and standard deviation (weighted): School year 2003-04

Variable	Description	Range of Values	Weighted mean	Standard Deviation
C6SDQRDC	C6 SDQ Prcvd Interest[1]/Competence - Reading	1 - 4	3.00	.74
C6SDQMTC	C6 SDQ Prcvd Interest/Competence - Math	1 - 4	2.92	.79
C6SDQSBC	C6 SDQ Prcvd Interest/Competence - All Sbj	1 - 4	2.71	.65
C6SDQPRC	C6 SDQ Prcvd Interest/Competence - Peer Rl	1 - 4	2.98	.63
C6SDQEXT	C6 SDQ Externalizing Problems	1 - 4	1.89	.69
C6SDQINT	C6 SDQ Internalizing Problems	1 - 4	2.08	.64

[1] "Prcvd Interest" = Perceived Interest.
NOTE: Table estimates based on C6CW0 weight. See chapter 7, section 7.4 for variable naming conventions.
SOURCE: U.S. Department of Education, National Center for Education Statistics, Early Childhood Longitudinal Study, Kindergarten Class of 1998–99 (ECLS-K), spring 2004.

4. SAMPLE DESIGN AND IMPLEMENTATION

This chapter describes the sample design of the Early Childhood Longitudinal Study, Kindergarten Class of 1998–99 (ECLS-K), and how it was modified and implemented for each round of data collection. An overview of the sample design is given here and described in more detail in the following sections, followed by a discussion of the types of weights needed for analyses, and how they were computed.

The ECLS-K employed a multistage probability sample design to select a nationally representative sample of children attending kindergarten in 1998–99. In the base year the primary sampling units (PSUs) were geographic areas consisting of counties or groups of counties. The second-stage units were schools within sampled PSUs. The third and final stage units were students within schools.

The first-grade data collection targeted base year respondents, where a case was considered responding if there was a completed child assessment or parent interview in fall- or spring-kindergarten. While all base-year respondents were eligible for the spring-first grade data collection, fall-first grade was limited to a 30 percent subsample. The spring student sample was freshened to include current first-graders who had not been enrolled in kindergarten in 1998–99 and, therefore, had no chance of being included in the ECLS-K base year kindergarten sample. For both fall- and spring-first grade, only a subsample of students who had transferred from their kindergarten school was followed.

The third-grade data collection targeted base year respondents and children sampled in first grade through the freshening operation where the spring-first grade sample was freshened to include first-graders who had not been enrolled in kindergarten in 1998–99 and therefore had no chance of being included in the ECLS-K base year kindergarten sample. As in the first-grade data collection where only a subsample of students who had transferred from their kindergarten school was followed, a subsampling of movers was also used in third grade. In third grade, however, the subsampling rate applied to transferred children was slightly higher; children whose home language was non-English (also known as children belonging to the language minority group) who moved for the first time between kindergarten or first grade and third grade, were followed at 100 percent. In other words, children belonging to the language minority group who did not move in first grade but moved in third grade were all followed into their new third-grade schools. The higher subsampling rate allows for the preservation of this group in the sample for analytic reasons. Children not in the language minority group continued to be subsampled for followup if they moved in third grade.

In fifth grade, the sample that was fielded was reduced by excluding certain special groups of children from data collection, and by setting differential sampling rates for movers in different categories. Specifically, children in four groups were not fielded for the fifth-grade survey, irrespective of other subsampling procedures that were implemented. They are children who became ineligible in an earlier round because they died or moved out of the country, children who were subsampled out in previous rounds because they were movers, children whose parents emphatically refused to cooperate (hard refusals), and children eligible for the third-grade data collection for whom there are neither first-grade nor third-grade data. Of the remaining children, those who move from their original schools during fifth grade or earlier were subsampled for followup. Children whose home language is not English (language minority) continued to be a special domain of analytic interest, and were subsampled at higher rates. Children were subsampled at different rates depending on the longitudinal data available for those children.

4.1 Base Year Sample

In the base year, children were selected for the ECLS-K using a multistage probability design. The PSUs were counties or groups of counties selected with probability proportional to size (PPS). The basic PSU measure of size was the number of 5-year-olds, but this was modified to facilitate the oversampling of Asian and Pacific Islanders (APIs) required to meet precision goals. In all, there were 100 PSUs selected for the ECLS-K. The 24 PSUs with the largest measure of size were designated self-representing (SR) and were included in the sample with certainty. The remaining non-SR PSUs were partitioned into 38 strata of roughly equal size. An initial cross-classification of census region with metropolitan statistical area (MSA) status, created eight superstrata. These were further subdivided by percent minority, PSU measure of size (a composite count of 5-year-old children), and 1988 per capita income. From each non-SR stratum, two PSUs were selected PPS without replacement using Durbin’s Method (Durbin 1967).

Table 4-1 summarizes the characteristics of the ECLS-K PSU sample.

Table 4-1. Distribution of the ECLS-K primary sampling unit (PSU) sample by self-representing (SR) status, metropolitan statistical area (MSA) status, and census region: School year 1998–99

SR status	MSA status	Total	Census region			
			Northeast	Midwest	South	West
Total		100	18	25	34	23
SR	MSA	24	6	5	6	7
Non-SR	MSA	52	10	12	18	12
Non-SR	Non-MSA	24	2	8	10	4

SOURCE: U.S. Department of Education, National Center for Education Statistics, Early Childhood Longitudinal Study, Kindergarten Class of 1998–99 (ECLS-K), fall 1998 and spring 1999.

In the second stage, public and private schools offering kindergarten programs were selected. For each PSU, a frame of public and private schools, offering kindergarten programs, was constructed using existing school universe files: the 1995–96 Common Core of Data (CCD; U.S. Department of Education 1995–96) and the 1995–96 Private School Universe Survey (PSS; U.S. Department of Education 1998). The 1995–96 Office of Indian Education Programs Education Directory was consulted in order to complete the list of Bureau of Indian Affairs (BIA) schools in the CCD file. For Department of Defense (DOD) domestic schools, a 1996 list of schools was obtained directly from the DOD. These schools constitute the original frame. A procedure was implemented to create a freshened frame by identifying kindergarten programs that would be operational at the time of ECLS-K's base year data collection, but that were not included in the original frame. These were newly opened schools that were not listed in the CCD and the PSS, as well as schools that were in the CCD and the PSS but did not appear to offer kindergarten programs according to those sources. The selection of schools was systematic, with probability proportional to a weighted measure of size based on the number of kindergartners enrolled. As with the PSU sample, the measure of size was constructed taking into account the desired oversampling of APIs. Public and private schools constituted distinct sampling strata. Within each stratum, schools were sorted to ensure good sample representation across other characteristics. In total, 1,280 schools were sampled from the original frame, and 133 from the freshened frame. Of these, 953 were public schools and 460 were private schools.

The characteristics of the ECLS-K school sample are presented in table 4-2. During recruitment, 136 schools were discovered to be ineligible because they did not have any kindergarten programs in the school. They are not included in table 4-2.

Table 4-2. Number of schools in the ECLS-K base year school sample, by selected school characteristics: School year 1998–99

Characteristic	Total	Sector	
		Public	Private
Total	1,277	914	363
Region			
Northeast	243	161	82
Midwest	298	210	88
South	418	306	112
West	318	237	81
Type of locale			
Large city	245	168	77
Midsize city	248	172	76
Urban fringe of large city	382	265	117
Urban fringe of midsize city	99	78	21
Large town	33	24	9
Small town	112	76	36
Rural	158	131	27
School affiliation			
Public	914	914	†
Catholic	120	†	120
Non-Catholic, religious	149	†	149
Nonreligious, private	94	†	94
School type			
Regular[1]	1,162	893	269
Ungraded	4	1	3
No grade beyond kindergarten	49	19	30
Unknown	62	1	61

† Not applicable.
[1] School offers kindergarten and at least another grade between 1^{st} and 12^{th}.
SOURCE: U.S. Department of Education, National Center for Education Statistics, Early Childhood Longitudinal Study, Kindergarten Class of 1998–99 (ECLS-K), fall 1998 and spring 1999.

The third stage sampling units were children of kindergarten age, selected within each sampled school. The goal of the student sample design was to obtain an approximately self-weighting sample of students and at the same time to achieve a minimum required sample size for APIs who were the only subgroup that needed to be oversampled to meet the study's precision goals. For each sampled school, the field staff obtained a complete list of kindergartners enrolled. Two independent sampling strata were formed within each school, one containing API students and the second, all other students. Within each stratum, students were selected using equal probability systematic sampling, using a higher rate for the API stratum.[1] In general, the target number of children sampled at any one school was 24. Once the sampled children were identified, parent contact information was obtained from the school. The information was used to locate a parent or guardian and gain parental consent for the child assessment and for the parent interview. Table 4-3 presents characteristics of children sampled and eligible for the base year.

During the fall-kindergarten data collection, a census of kindergarten teachers was taken at each school. Each sampled child was linked to his or her kindergarten teacher. In spring-kindergarten, teacher-child linkages were reviewed and updated. If new kindergarten teachers had joined the school, they were added to the census of kindergarten teachers. Special education teachers who taught one or more sampled children were included in the spring-kindergarten data collection. If a sampled child received special education services from such a teacher, the teacher was linked to that child.

While the sample of schools is the same for fall- and spring-kindergarten, the child sample is larger in spring than in fall. In spring-kindergarten, 1,426 additional children were sampled from the schools that refused to participate in fall but were converted into respondents in spring.

For a detailed description of the base year sample, see the *ECLS-K Base Year Public-Use Data Files and Electronic Codebook: User's Manual* (NCES 2001–029r) (Tourangeau, Burke, Lê, et al. 2004).

4.2 Fall-First Grade Subsample

A subsample of ECLS-K base year PSUs was selected for fall-first-grade data collection. All 24 of the SR PSUs were retained. Of the 76 non-self-representing (NSR) PSUs, 38 were retained by sampling one PSU per stratum with equal probability.

[1] See the *ECLS-K Base Year Public-Use Data Files and Electronic Codebook: User's Manual* (NCES 2001–029r) (Tourangeau, Burke, Lê, et al. 2004).

Table 4-3. Number (unweighted) of children in the ECLS-K base year student sample, by selected characteristics: School year 1998–99

Characteristic	Total	Sector: Public	Sector: Private
Total	22,666	17,777	4,889
Region			
Northeast	4,262	3,045	1,217
Midwest	5,628	4,292	1,336
South	7,461	6,179	1,282
West	5,315	4,261	1,054
Type of locale			
Large city	4,550	3,365	1,185
Midsize city	4,728	3,569	1,159
Urban fringe of large city	6,470	4,945	1,525
Urban fringe of midsize city	1,644	1,434	210
Large town	714	577	137
Small town	1,905	1,485	420
Rural	2,655	2,402	253
School affiliation			
Public	17,777	17,777	†
Catholic	2,510	†	2,510
Non-Catholic, religious	1,445	†	1,445
Nonreligious, private	934	†	934
School type			
Regular[1]	21,436	17,390	4,046
Ungraded	56	24	32
No grade beyond kindergarten	663	338	325
Unknown	511	25	486
Child race/ethnicity			
White	11,723	8,533	3,190
Black	3,204	2,800	404
Hispanic, with race	1,749	1,455	294
Hispanic, without race	1,983	1,741	242
Asian	1,355	1,102	253
Pacific Islander	220	199	21
Native American	377	334	43
Multirace	511	416	95
Unknown	1,544	1,197	347

See notes at end of table.

Table 4-3. Number (unweighted) of children in the ECLS-K base year student sample, by selected characteristics: School year 1998–99—Continued

Characteristic	Total	Sector	
		Public	Private
Highest parent level of education			
Less than high school	2,027	1,968	59
High school graduate	5,251	4,703	548
Vocational/technical	1,139	964	175
Some college	5,351	4,182	1,169
College graduate	4,004	2,568	1,436
Masters	1,429	850	579
Ph.D./professional	890	456	434
Unknown	2,575	2,086	489

† Not applicable.

[1] School offers kindergarten and at least another grade between 1st and 12th.

SOURCE: U.S. Department of Education, National Center for Education Statistics, Early Childhood Longitudinal Study, Kindergarten Class of 1998–99 (ECLS-K), fall 1998 and spring 1999.

Base year schools in the 62 fall-first grade sampled PSUs were stratified by frame source (original public, original private, freshened public, and freshened private as described in section 4.1) and arranged in their original selection order. A 30 percent equal probability subsample of schools was drawn in the 24 SR PSUs and a 60 percent subsample of schools was drawn in the 38 NSR PSUs. In total 311 schools that had cooperated in either fall- or spring-kindergarten were selected. The characteristics of the base year cooperating schools selected for fall-first grade data collection are presented in table 4-4.

Table 4-4. Number of base year cooperating schools selected for fall-first grade, by selected school characteristics: School year 1999–2000

Characteristic	Total	Sector	
		Public	Private
Total	311	228	83
Region			
Northeast	57	39	18
Midwest	83	59	24
South	99	77	22
West	72	53	19
Type of locale			
Large city	62	42	20
Midsize city	59	45	14
Urban fringe of large city	86	61	25
Urban fringe of midsize city	18	14	4
Large town	15	12	3
Small town	28	19	9
Rural	43	35	8
School affiliation			
Public	228	228	†
Catholic	29	†	29
Non-Catholic, religious	33	†	33
Nonreligious, private	21	†	21
School type			
Regular[1]	292	222	70
Ungraded	1	1	0
No grade beyond kindergarten	18	5	13

† Not applicable.
[1] School offers kindergarten and at least another grade between 1st and 12th.
SOURCE: U.S. Department of Education, National Center for Education Statistics, Early Childhood Longitudinal Study, Kindergarten Class of 1998–99 (ECLS-K), fall 1999 and spring 2000.

The fall-first grade data collection consisted of the direct child assessment and the parent interview. Data collection was attempted for every eligible child found still attending the school in which he or she had been sampled during kindergarten and a subset of eligible children who had transferred from the school in which he or she was originally sampled. "Eligible" is defined as a base year respondent (i.e., a child who had either a fall- or spring-kindergarten child assessment or parent interview or was excluded from assessment because of a disability or because the child belonged in the language minority, not Spanish group). To contain the costs of data collection, a random 50 percent of children were flagged to be followed for fall-first grade data collection in the event that they had transferred.

Except for children who were repeating kindergarten, all base year children sampled in schools with a high grade of kindergarten are de facto movers. Since many of these movers may move *en masse* to the same first-grade school, steps were taken to follow these children at a higher rate. Using the information collected during spring-kindergarten, a list of destination schools was compiled for each such school. The destination school having the most movers was designated as primary, unless no such school had more than three movers. Children who moved *en masse* into a primary destination school in fall-first grade were treated as "nonmovers" and were not subsampled (that is, they continued to be followed and were part of the ECLS-K sample). In this way, movers are defined differently in this chapter (statistical movers) than in chapter 5 (operation movers).

As discussed above, a random 50 percent of children were subsampled to be followed if they moved out of the kindergarten school. Prior to sampling, children were stratified into groups of nonmovers, movers with information identifying their new schools, and movers without such identifying information. Sampling was done with equal probability within subsampling strata using the same sampling rate of 0.5 in each substratum. A flag was created for each child indicating whether the child had been sampled to be followed.

Table 4-5 shows the characteristics of the children subsampled and eligible for fall-first grade. Region, locale, school affiliation, and school type describe the school the child attended in kindergarten.

4.3 Spring-First Grade Sample

The ECLS-K spring-first grade data collection targeted all base year respondents (i.e., respondent in fall- or spring-kindergarten). In addition, the spring student sample was freshened to include current first-graders who had not been enrolled in kindergarten in 1998–99 and, therefore, had no chance of being included in the ECLS-K base year kindergarten sample. While all students still enrolled in their base year schools were recontacted, only a 50 percent subsample of base year sampled students who had transferred from their kindergarten school was followed for data collection.

Table 4-5. Number (unweighted) of children subsampled and eligible for fall-first grade, by selected characteristics: School year 1999–2000

Characteristic	Total	Sector	
		Public	Private
Total	5,650	4,446	1,204
Region			
Northeast	1,000	759	241
Midwest	1,416	1,068	348
South	1,873	1,557	316
West	1,361	1,062	299
Type of locale			
Large city	1,154	816	338
Midsize city	1,109	874	235
Urban fringe of large city	1,558	1,205	353
Urban fringe of midsize city	320	276	44
Large town	306	246	60
Small town	518	390	128
Rural	685	639	46
School affiliation			
Public	4,446	4,446	†
Catholic	535	†	535
Non-Catholic, religious	254	†	254
Nonreligious, private	415	†	415
School type			
Regular[1]	5,374	4,338	1,036
Ungraded	24	24	0
No grade beyond kindergarten	138	84	54
Unknown	114	0	114
Child's race/ethnicity			
White	3,131	2,288	843
Black	849	718	131
Hispanic, with race	419	345	74
Hispanic, without race	522	475	47
Asian	305	243	62
Pacific Islander	99	97	2
Native American	137	132	5
Multirace	163	127	36
Unknown	25	21	4

See notes at end of table.

Table 4-5. Number (unweighted) of children subsampled and eligible for fall-first grade, by selected characteristics: School year 1999–2000—Continued

Characteristic	Total	Sector	
		Public	Private
Highest parent level of education			
Less than high school	530	521	9
High school graduate	1,252	1,124	128
Vocational/technical	335	285	50
Some college	1,419	1,119	300
College graduate	1,038	680	358
Masters	398	241	157
Ph.D./professional	255	125	130
Unknown	423	351	72

† Not applicable.

[1] School offers kindergarten and at least another grade between 1st and 12th.

NOTE: School characteristics (i.e., region, locale, school affiliation, and school type) describe the school the child attended in kindergarten.

SOURCE: U.S. Department of Education, National Center for Education Statistics, Early Childhood Longitudinal Study, Kindergarten Class of 1998–99 (ECLS-K), fall 1999 and spring 2000.

4.3.1 Subsampling Movers

In spring-first grade all children in a random 50 percent subsample of base year schools were flagged to be followed for data collection if they transferred from their base year school. (This is in contrast to fall-first grade, where a random 50 percent of children in each of the 30 percent of schools subsampled were flagged). In order to maximize the amount of longitudinal data, care was taken during spring-first grade sampling to ensure that any child who had been flagged to be followed in fall-first grade would continue to be so.

In selecting the spring-first grade 50 percent subsample of schools where movers would be flagged for followup, the three primary strata were SR PSUs, NSR PSUs that had been selected for fall-first grade, and NSR PSUs that had not been selected for fall-first grade. Within these major strata, schools were grouped by frame source (original public, original private, freshened public, and freshened private as described in section 4.1). Finally, within each frame source, schools were stratified by whether the school participated in the base year study, and were then arranged in original selection order. Schools that had been part of the 30 percent fall-first grade sample were automatically retained. Then equal probability sampling methods were employed to augment the sample to the desired 50 percent. The net result of these procedures was that every base year selected school had on average a 50 percent chance of having its ECLS-K transfer students followed during spring-first grade, and any transfer student who had been followed in fall-first grade would still be followed in spring-first grade.

Table 4-6 shows the characteristics of the eligible children in the spring-first grade sample, excluding freshened students. Region, locale, school affiliation, and school type describe the school at which the child attended kindergarten.

Table 4-6. Number (unweighted) of eligible children in spring-first grade sample excluding freshened students, by selected characteristics: School year 1999–2000

Characteristic	Total	Sector	
		Public	Private
Total	18,084	14,248	3,836
Region			
Northeast	3,339	2,434	905
Midwest	4,578	3,474	1,104
South	6,050	5,029	1,021
West	4,117	3,311	806
Type of locale			
Large city	3,459	2,575	884
Midsize city	3,761	2,797	964
Urban fringe of large city	5,140	3,991	1,149
Urban fringe of midsize city	1,288	1,126	162
Large town	576	466	110
Small town	1,578	1,215	363
Rural	2,282	2,078	204
School affiliation			
Public	14,248	14,248	†
Catholic	2,091	†	2,091
Non-Catholic, religious	1,139	†	1,139
Nonreligious, private	606	†	606
School type			
Regular[1]	17,277	13,971	3,306
Ungraded	40	24	16
No grade beyond kindergarten	420	235	185
Unknown	347	18	329

See notes at end of table.

Table 4-6. Number (unweighted) of children in spring-first grade sample excluding freshened students, by selected characteristics: School year 1999–2000—Continued

Characteristic	Total	Sector	
		Public	Private
Child's race/ethnicity			
White	10,208	7,472	2,736
Black	2,597	2,289	308
Hispanic, with race	1,460	1,220	240
Hispanic, without race	1,648	1,456	192
Asian	1,149	939	210
Pacific Islander	202	186	16
Native American	332	294	38
Multirace	434	347	87
Unknown	54	45	9
Highest parent level of education			
Less than high school	1,529	1,491	38
High school graduate	3,779	3,356	423
Vocational/technical	1,078	926	152
Some college	4,211	3,313	898
College graduate	3,348	2,194	1,154
Masters	1,191	719	472
Ph.D./professional	749	395	354
Unknown	2,199	1,854	345

† Not applicable.

[1] School offers kindergarten and at least another grade between 1st and 12th.

NOTE: School characteristics (i.e., region, locale, school affiliation, and school type) describe the school the child attended in kindergarten.

SOURCE: U.S. Department of Education, National Center for Education Statistics, Early Childhood Longitudinal Study, Kindergarten Class of 1998–99 (ECLS-K), fall 1999 and spring 2000.

4.3.2 Student Freshening

The spring-first grade student freshening used a half-open interval sampling procedure (Kish 1965). The procedure was implemented in the same 50 percent subsample of ECLS-K base year schools where transfer students were flagged for followup. Each of these schools was asked to prepare an alphabetized roster of students enrolled in first grade and the names of ECLS-K kindergarten-sampled students were identified on this list. Beginning with the name of the first kindergarten-sampled child, school records were checked to see whether the student directly below in the sorted list attended kindergarten in the United States in fall 1998. If not, (1) that child was considered to be part of the freshened sample and (2) the record search procedure was repeated for the next listed child, and so forth.

When the record search revealed that a child had been enrolled in kindergarten the previous year, that child was not considered part of the freshened sample and the procedure was begun all over again with the second base year sampled student name, and so on. Note: the student roster was "circularized" (i.e., the first name on the roster was considered to follow the last name on the roster in the implementation of the procedure). Student freshening brought 165 first graders into the ECLS-K sample, which increased the weighted survey estimate of the number of first-graders in the United States by about 2.6 percent.

The student freshening procedure was not entirely free of bias. A first grader would have no chance of being in the ECLS-K first grade sample if he or she was enrolled in a school where neither the child nor any of his or her classmates had attended kindergarten in the United States in the fall of 1998. This would be a rare circumstance and is not thought to be an important source of bias. A more significant source of potential bias is nonresponse. One source of nonresponse inherent to the freshening plan was that the procedure only involved students who had not transferred from the school in which they had been sampled during the base year. A more detailed discussion of freshened student nonresponse can be found in section 5.7.2 of the *ECLS-K User's Manual for the ECLS-K First Grade Public-Use Data Files and Electronic Codebook* (NCES 2002–135) (Tourangeau, Burke, et al. 2002).

4.4 Spring-Third Grade Sample

The sample of children for spring-third grade consists of all children who were base year respondents and children who were brought into the sample in spring-first grade through the sample freshening procedure described in section 4.3.2. Sample freshening was not implemented in third grade, hence no new students entered the sample.

While all students still enrolled in their base year schools were recontacted, slightly more than 50 percent of the base year sampled students who had transferred from their kindergarten school were followed for data collection. This subsample of students was the same 50 percent subsample of base year movers flagged for following in spring-first grade, with the addition of movers whose home language was not English (language minority students). The two special sampling procedures implemented in spring-third grade are described below.

4.4.1 Subsampling Movers

In spring-first grade, all children in a random 50 percent subsample of base year schools were flagged to be followed for data collection if they transferred from their base year school at any point in the future. In order to maximize the amount of longitudinal data, care was taken during spring-first grade sampling to ensure that any child who had been flagged to be followed in fall-first grade would continue to be followed. The spring-first grade sampling procedure for movers is described in section 4.3.1. In spring-third grade, children who were followed in spring-first grade were retained in the sample (i.e., the mover followup still targeted the same 50 percent subsample of children in the base year schools). In addition, language minority children who moved between first and third grade were followed with certainty as described below.

4.4.2 Language Minority Children

In addition to the subsample of movers to be followed described above, children whose home language was not English and who moved between spring-first grade and spring-third grade were all retained rather than being subsampled at the 50 percent rate. Operationally, this means that children whose home language was not English who were not flagged for followup in the previous round had their flags switched from “not to be followed” to “to be followed.” This mover flag was set in first grade to specify whether a child was to be followed if he or she moved from the kindergarten school at any point in the future. This only affects language minority children who had not moved out of the original sample schools before third grade. If they had moved before third grade, then their flags were not switched and they continued not to be followed. This modification to the mover followup procedure provides a larger sample of children whose home language is not English. The mover followup activities that originally targeted a 50 percent subsample of children in base year schools resulted in a 54 percent subsample with the addition of language minority children.

Table 4-7 shows the characteristics of eligible children in the spring-third grade sample, excluding freshened students. Region, locale, school affiliation, and school type describe the school at which the child attended kindergarten.

Table 4-7. Number (unweighted) of eligible children in spring-third grade sample excluding freshened students, by selected characteristics: School year 2001–02

Characteristic	Total	Sector	
		Public	Private
Total	16,670	13,166	3,504
Region			
Northeast	3,102	2,274	828
Midwest	4,208	3,187	1,021
South	5,522	4,607	915
West	3,838	3,098	740
Type of locale			
Large city	3,150	2,344	806
Midsize city	3,385	2,536	849
Urban fringe of large city	4,747	3,705	1,042
Urban fringe of midsize city	1,194	1,033	161
Large town	536	428	108
Small town	1,491	1,149	342
Rural	2,167	1,971	196
School affiliation			
Public	13,166	13,166	†
Catholic	1,924	†	1,924
Non-Catholic, religious	1,036	†	1,036
Nonreligious, private	544	†	544
School type			
Regular[1]	15,930	12,901	3,029
Ungraded	34	23	11
No grade beyond kindergarten	391	222	169
Unknown	315	20	295
Child's race/ethnicity			
White	9,348	6,853	2,495
Black	2,238	1,977	261
Hispanic, with race	1,450	1,222	228
Hispanic, without race	1,547	1,367	180
Asian	1,115	911	204
Pacific Islander	196	180	16
Native American	305	273	32
Multirace	432	351	81
Unknown	39	32	7

See notes at end of table.

Table 4-7. Number (unweighted) of children in spring-third grade sample excluding freshened students, by selected characteristics: School year 2001–02—Continued

Characteristic	Total	Sector	
		Public	Private
Highest parent level of education			
Less than high school	1,586	1,543	43
High school graduate	3,536	3,196	340
Vocational/technical	935	801	134
Some college	4,500	3,621	879
College graduate	3,517	2,352	1,165
Masters	1,324	825	499
Ph.D./professional	813	429	384
Unknown	459	399	60
Home language			
Not English	4,409	3,676	733
English	12,261	9,490	2,771

† Not applicable.
[1] School offers kindergarten and at least another grade between 1st and 12th.
NOTE: School characteristics (i.e., region, locale, school affiliation, and school type) describe the school the child attended in kindergarten.
SOURCE: U.S. Department of Education, National Center for Education Statistics, Early Childhood Longitudinal Study, Kindergarten Class of 1998–99 (ECLS-K), spring 2002.

For a detailed description of the third-grade sample, see the *User's Manual for the ECLS-K Third Grade Public-Use Data File and Electronic Code Book* (NCES 2004–001) (Tourangeau, Brick, Lê, et al. 2004).

4.5 Spring-Fifth Grade Sample

In fifth grade, four groups of children were not followed, irrespective of other subsampling procedures that were implemented. They are (1) children who became ineligible in an earlier round (because they died or moved out of the country), (2) children who were subsampled out in previous rounds because they moved out of the original schools and were not subsampled to be followed, (3) children whose parents emphatically refused to cooperate (hard refusals) in any of the data collection rounds since spring-kindergarten, and (4) children eligible for the third-grade data collection for whom there are neither first-grade nor third-grade data. Among the 21,357 children who were eligible for the study after the base year, 5,214 were excluded from the fifth-grade survey, and they are distributed as shown in table 4-8.

Table 4-8. Number of children eligible after the base year but excluded from the fifth-grade data collection: School year 2003–2004

Characteristics[1]	Total	Mover subsampled out in first or third grade[2]	Ineligible in first or third grade	Hard refusal	Eligible for third-grade sample, with no first- or third-grade data
Total	5,214	4,117	122	571	404
School affiliation					
Public	4,000	3,129	98	433	340
Catholic	485	405	7	52	21
Non-Catholic, religious	361	270	9	61	21
Nonreligious, private	352	313	7	19	13
Unknown	16	0	1	6	9
Urbanicity					
City	2,436	1,960	68	218	190
Suburb and town	2,388	1,869	45	300	174
Rural	381	288	5	51	37
Unknown	9	0	4	2	3
Race/ethnicity					
White	2,794	2,272	36	327	159
Black	1,061	867	12	88	94
Hispanic	811	584	47	82	98
Asian/Pacific Islander	313	225	20	46	22
Other	201	158	5	16	22
Unknown	34	11	2	12	9
Language minority					
Not English	1,000	684	84	124	108
English	4,214	3,433	38	447	296
SES quintile					
First (lowest)	975	772	29	75	99
Second	982	811	20	81	70
Third	874	707	14	89	64
Fourth	933	791	17	84	41
Fifth (highest)	948	793	36	82	37
Unknown	502	243	6	160	93

[1] Characteristics are from the most recent data available for the child (e.g., if a child was not subsampled in third grade and had data from first grade, then the characteristics of the child come from first grade).
[2] These are statistical movers, not operation movers as discussed in chapter 5.
SOURCE: U.S. Department of Education, National Center for Education Statistics, Early Childhood Longitudinal Study, Kindergarten Class of 1998–99 (ECLS-K), fall 1998, spring 1999, fall 1999, spring 2000, spring 2002, and spring 2004.

Of the remaining children, those who moved from their original schools during fifth grade or earlier were subsampled for followup. In order to contain the cost of data collection, the rate of subsampling was lower in fifth grade than it had been in previous years. The subsampling rates maximize the amount of longitudinal data available for key analytic groups. Children whose home language is not English (language minority) continued to be a special domain of analytic interest, and were subsampled at higher rates. Children were subsampled at different rates depending on the longitudinal data available for those children.

For base year respondents, the sampling rates for following movers are as follows:

- 0.33 for non-language minority (LM) movers with full longitudinal data;
- 0.25 for non-LM movers with third-grade but not first-grade data;
- 0.15 for non-LM movers with first-grade but not third-grade data;
- 0.75 for LM movers with full longitudinal data;
- 0.50 for LM movers with third-grade but not first-grade data; and
- 0.25 for LM movers with first-grade but not third-grade data.

For subsampling freshened children (i.e., children sampled in first grade) who are movers in fifth grade (or earlier) the rates are:

- 0.33 for non-LM movers with full longitudinal data;
- 0.15 for non-LM movers with third-grade but not first-grade data;
- 0.15 for non-LM movers with first-grade but not third-grade data;
- 0.75 for LM movers with full longitudinal data;
- 0.25 for LM movers with third-grade but not first-grade data; and
- 0.25 for LM movers with first-grade but not third-grade data.

These rates are different than those used in third grade where movers were subsampled uniformly at a rate of 0.5, and language minority children were followed at 100 percent (unless they were already subsampled out in first grade). The mover followup activities that originally targeted a 50 percent subsample of children in base year schools resulted in a 54 percent subsample with the addition of language minority children in third grade. For fifth grade, these mover followup activities targeted a 42 percent subsample of movers where were eligible to be fielded in fifth grade and resulted in a 41 percent subsample.

Table 4-9 shows the characteristics of eligible children in the spring-fifth grade sample, excluding freshened students. Region, locale, school affiliation, and school type describe the school at which the child attended kindergarten.

A new feature of the fifth-grade sample is the subsampling of children for the administration of the mathematics or science questionnaires. While all children retained for the fifth-grade data collection had child-level questionnaires filled out by their reading teachers, half were subsampled to have child-level questionnaires filled out by their mathematics teachers and the other half had child-level questionnaires filled out by their science teachers.

4.6 Sample Attrition

In a longitudinal study, sample attrition due to nonresponse and change in eligibility status is expected. The sample of respondents decreases with each round of data collection. In the case of the ECLS-K, a combination of field and sampling procedures was applied that caused the sample to increase after the fall-kindergarten data collection, but then decrease in each subsequent round.

Table 4-9. Number (unweighted) of eligible children in spring-fifth grade sample excluding freshened students, by selected characteristics: School year 2003–04

Characteristic	Total	Sector	
		Public	Private
Total	12,029	9,567	2,462
Region			
Northeast	2,254	1,705	549
Midwest	3,124	2,354	770
South	3,849	3,237	612
West	2,802	2,271	531
Type of locale			
Large city	2,208	1,631	577
Midsize city	2,370	1,698	672
Urban fringe of large city	3,419	2,764	655
Urban fringe of midsize city	833	739	94
Large town	373	295	78
Small town	1,140	884	256
Rural	1,686	1,556	130
School affiliation			
Public	9,567	9,567	†
Catholic	1,477	†	1,477
Non-Catholic, religious	700	†	700
Nonreligious, private	285	†	285
School type			
Regular[1]	11,611	9,404	2,207
Ungraded	26	17	9
No grade beyond kindergarten	203	141	62
Unknown	189	5	184
Child's race/ethnicity			
White	6,846	5,075	1,771
Black	1,365	1,229	136
Hispanic, with race	1,103	934	169
Hispanic, without race	1,161	1,027	134
Asian	852	703	149
Pacific Islander	156	142	14
Native American	228	204	24
Multirace	290	229	61
Unknown	28	24	4

See notes at end of table.

Table 4-9. Number (unweighted) of eligible children in spring-fifth grade sample excluding freshened students, by selected characteristics: School year 2003–04—Continued

Characteristic	Total	Sector	
		Public	Private
Highest parent level of education			
Less than high school	1,013	992	21
High school graduate	2,481	2,261	220
Vocational/technical	673	590	83
Some college	3,362	2,736	626
College graduate	2,693	1,862	831
Masters	1,076	700	376
Ph.D./professional	667	366	301
Unknown	64	60	4
Home language			
Not English	3,485	2,908	577
English	8,544	6,659	1,885

† Not applicable.
[1] School offers kindergarten and at least another grade between 1st and 12th.
NOTE: School characteristics (i.e., region, locale, school affiliation, and school type) describe the school the child attended in kindergarten.
SOURCE: U.S. Department of Education, National Center for Education Statistics, Early Childhood Longitudinal Study, Kindergarten Class of 1998–99 (ECSL-K), spring 2004.

The first procedure was the school-level refusal conversion in spring-kindergarten, resulting in a number of schools that agreed to participate in the study after having refused to do so in the previous round. From these schools, 1,426 children were sampled and added to the initial sample of 21,387 kindergarten children. The second procedure was sample freshening in spring-first grade as described in section 4.3.2. This brought in 165 eligible children to add to the sample of 21,192 base year respondents who remained eligible after the base year. A base year responding child was defined as one with at least one direct cognitive test score in fall- or spring-kindergarten or whose parent responded to the family structure section of the parent instrument in fall- or spring-kindergarten. The third procedure—applied in first, third, and fifth grades—required that a subsample of children who moved out of their original sample schools not be followed into their new schools, as described in sections 4.3.1 and 4.4.1, resulting in a decrease in the sample. The fourth and last procedure, applied in fifth grade only, is the exclusion from the data collection of children who were difficult to field, as described in section 4.5, also resulting in a significant decrease in the sample.

Table 4-10 shows the sample size for each round of data collection of the ECLS-K, and the response status of the children in each round. Tables 4-11 and 4-12 show the same children separately by the original sample school affiliation (public/private).

Table 4-10. Number (unweighted) of children in the ECLS-K sample, by response status and data collection round: School years 1998–99, 1999–2000, 2001–02, and 2003–04

Data collection round	Unweighted sample size	Response status				
		Ineligibles	Unknown eligibility	Non-followed movers	Nonrespondents	Respondents
Fall-kindergarten	21,387	31	†	†	1,672	19,684
Spring-kindergarten	22,813[1]	147	†	†	2,088	20,578
Fall-first grade	6,507	39	37	781	226	5,424
Spring-first grade	21,357[2]	56	202	2,850	925	17,324
Spring-third grade	21,357	122	289	4,117	1,524	15,305
Spring-fifth grade	16,143[3]	39	210	3,765	309	11,820

† Not applicable.

[1] 1,426 children were sampled from refusal-converted schools.

[2] 21,192 children remained eligible after the base year. In addition, 165 children were sampled via the sample freshening procedure.

[3] 5,214 children were excluded from the fifth-grade data collection. They were children who became ineligible in an earlier round, movers not subsampled to be followed in previous rounds, and hard-to-field cases such as hard refusals, and children with neither first-grade nor third-grade data.

NOTE: Response status is defined in terms of completed child assessment OR completed family structure data of the parent interview. Children who died or moved out of the country are classified as ineligible. Children who moved and were subsampled for followup but could not be located were treated as belonging to the unknown eligibility category. A portion of children who moved was subsampled out and not followed into their new schools. The numbers of children in this table are different than in tables 4-3 to 4-7 and table 4-9 since the earlier tables only include eligible children.

SOURCE: U.S. Department of Education, National Center for Education Statistics, Early Childhood Longitudinal Study, Kindergarten Class of 1998–99 (ECLS-K), fall 1998, spring 1999, fall 1999, spring 2000, spring 2002, and spring 2004.

Table 4-11. Number (unweighted) of public school children in the ECLS-K sample, by response status and data collection round: School years 1998–99, 1999–2000, 2001–02, and 2003–04

Data collection round	Unweighted sample size	Response status				
		Ineligibles	Unknown eligibility	Non-followed movers	Nonrespondents	Respondents
Fall-kindergarten	17,003	23	†	†	1,324	15,656
Spring-kindergarten	17,894[1]	117	†	†	1,676	16,101
Fall-first grade	5,118	35	36	601	173	4,273
Spring-first grade	16,784[2]	45	181	2,164	733	13,661
Spring-third grade	16,784	99	250	3,129	1,236	12,070
Spring-fifth grade	12,771[3]	37	190	2,889	243	9,412

† Not applicable.

[1] 891 public school children were sampled from refusal-converted schools.

[2] 16,638 public school children remained eligible after the base year. In addition, 146 public school children were sampled via the sample freshening procedure.

[3] 4,013 children in original sample public schools were excluded from the fifth-grade data collection. They were children who became ineligible in an earlier round, movers not subsampled to be followed in previous rounds, and hard-to-field cases such as hard refusals, and children with neither first-grade nor third-grade data.

NOTE: Response status is defined in terms of completed child assessment OR completed family structure data of the parent interview. Children who died or moved out of the country were classified as ineligible. Children who moved and were subsampled for followup but could not be located were treated as belonging to the unknown eligibility category. A portion of children who moved was subsampled out and not followed into their new schools. The numbers of children in this table are different than in tables 4-3 to 4-7 and table 4-9 since the earlier tables only include eligible children.

SOURCE: U.S. Department of Education, National Center for Education Statistics, Early Childhood Longitudinal Study, Kindergarten Class of 1998–99 (ECLS-K), fall 1998, spring 1999, fall 1999, spring 2000, spring 2002, and spring 2004.

Table 4-12. Number (unweighted) of private school children in the ECLS-K sample, by response status and data collection round: School years 1998–99, 1999–2000, 2001–02, and 2003–04

Data collection round	Unweighted sample size	Response status				
		Ineligibles	Unknown eligibility	Non-followed movers	Nonrespondents	Respondents
Fall-kindergarten	4,384	8	†	†	348	4,028
Spring-kindergarten	4,919[1]	30	†	†	412	4,477
Fall-first grade	1,389	4	1	180	53	1,151
Spring-first grade	4,573[2]	11	21	686	192	3,663
Spring-third grade	4,573	23	39	988	288	3,235
Spring-fifth grade	3,372[3]	2	20	876	66	2,408

† Not applicable.
[1] 535 private school children were sampled from refusal-converted schools.
[2] 4,554 private school children remained eligible after the base year. In addition, 19 private school children were sampled via the sample freshening procedure.
[3] 1,201 children from original private schools were excluded from the fifth-grade data collection. They were children who became ineligible in an earlier round, movers not subsampled to be followed in previous rounds, and hard-to-field cases such as hard refusals, and children with neither first-grade nor third-grade data.
NOTE: Response status is defined in terms of completed child assessment OR completed family structure data of the parent interview. Children who died or moved out of the country were classified as ineligible. Children who moved and were subsampled for followup but could not be located were treated as belonging to the unknown eligibility category. A portion of children who moved was subsampled out and not followed into their new schools. The numbers of children in this table are different than in tables 4-3 to 4-7 and table 4-9 since the earlier tables only include eligible children.
SOURCE: U.S. Department of Education, National Center for Education Statistics, Early Childhood Longitudinal Study, Kindergarten Class of 1998–99 (ECLS-K), fall 1998, spring 1999, fall 1999, spring 2000, spring 2002, and spring 2004.

The number of children who participated in all four years of the ECLS-K data collection (base year, first grade, third grade, and fifth grade) is 10,590 (8,506 in original public schools and 2,084 in original private schools). This represents 50 percent of the base year respondents or 46 percent of children sampled for the base year.

4.7 Calculation and Use of Sample Weights

As in previous years, the ECLS-K data were weighted to compensate for differential probabilities of selection at each sampling stage and to adjust for the effects of nonresponse. In the ECLS-K base year, weights were computed at the child, school and teacher levels. Estimates using the base year weights are representative of all kindergarten children, all schools with kindergarten programs and all kindergarten teachers. After the base year, only child-level weights were computed. The use of these weights is essential to produce estimates that are representative of the cohort of children who were in kindergarten in 1998–99 or in first grade in 1999–2000. Since the sample was not freshened after the first-grade year with third or fifth-graders who did not have a chance to be sampled in kindergarten or first grade (as was done in first grade), estimates from the ECLS-K third- and fifth-grade data are representative of the population cohort rather than all third-graders in 2001–02 or all fifth-graders in 2003–04. The estimated number of third-graders from the third-grade ECLS-K data collection is approximately 86 percent of all third-graders. From the fifth-grade ECLS-K data collection, the estimated number of fifth-graders is approximately 83 percent of all fifth-graders. While the vast majority of children in third grade in the 2001–02 school year and in fifth grade in the 2003–04 school year are members of the cohort, third-graders who repeated second or third grade, fifth-graders who repeated third or fourth grade, and recent immigrants are not covered. Data were collected from teachers and schools to provide important contextual information about the environment for the sampled children. The teachers and schools are not representative of third-grade teachers and schools in 2001–02, nor of fifth-grade teachers and schools in 2003–04. For this reason, the weights produced from the study after the kindergarten year are for making statements about children, including statements about the teachers and schools of those children.

Several sets of weights were computed for fifth grade. As in previous years, there are several survey instruments administered to sampled children and their parents, teachers and schools: cognitive and physical assessments for children; self-description child questionnaire (third and fifth grade only), parent instruments; several types of teacher instruments completed by reading, mathematics, science and special education teachers; and school instruments. The stages of base year sampling in conjunction with differential nonresponse at each stage and the diversity of survey instruments require that multiple fifth-grade cross-sectional sampling weights be computed for use in analyzing the fifth-grade ECLS-K data. Several combinations of kindergarten through fifth-grade longitudinal weights were also computed. Details on these longitudinal weights are available in chapter 9. This section describes the different types of fifth-grade cross-sectional weights, how they were calculated, how they should be used, and their statistical characteristics.

4.7.1 Types of Cross-Sectional Sample Weights

Five sets of cross-sectional weights were computed for children in the fifth-grade sample. These weights are defined as follows:

- C6CW0 is nonzero if the child has completed the assessment data or the child was excluded from direct assessment due to a disability.
- C6PW0 is nonzero if the child has completed the parent interview data.

- C6CPTR0 is nonzero if the child has completed the assessment data (or excluded from direct assessment due to a disability) and parent interview data and teacher-level data either from the reading teacher and/or the mathematics/science teacher.

- C6CPTM0 is nonzero if the child was sampled to have a child-level questionnaire completed by the mathematics teacher and the child has completed the assessment data (or excluded from direct assessment due to a disability) and parent interview data and teacher-level data either from the reading teacher or the mathematics teacher.

- C6CPTS0 is nonzero if the child was sampled to have a child-level questionnaire completed by the science teacher and the child has completed the assessment data (or excluded from direct assessment due to a disability) and parent interview data and teacher-level data either from the reading teacher or the science teacher.

If the child has only subject-specific child-level data from the teacher (reading, mathematics, or science) but no data from the teacher-level questionnaire, then the child is considered a nonrespondent for the CPT weights calculated, hence has none of the CPT weights.

In previous rounds, only one child-parent-teacher weight was computed based on the presence of the teacher questionnaire B (teacher-level). With the addition of the subject specific questionnaires filled out by teachers for each child in the ECLS-K sample, and the subsampling of children for the administration of the mathematics and science teacher questionnaires, three child-parent-teacher weights were computed. They are used to analyze direct child assessment data combined with parent interview data and data provided by the subject-specific teacher (child- and/or teacher-level data) in conjunction with school-level data, as described below.

Careful consideration should be given to the choice of a weight for a specific analysis since it depends on the type of data analyzed. Each set of weights is appropriate for a different set of data or combination of sets of data. Exhibit 4-1 summarizes how the different types of cross-sectional weights should be used. Cross-sectional weights are used to provide estimates for the fifth-grade data collection. Details under “to be used for analysis of . . .” provide guidance based on whether the data to be used with the weights were collected through the child assessments, parent interviews, or different types of teacher questionnaire.

Exhibit 4-1. ECLS-K fifth-grade cross-sectional weights: School year 2003–04

Weight	To be used for analysis of ...
C6CW0	Fifth-grade direct child assessment data, alone or in conjunction with any combination of (a) a limited set of child characteristics (e.g., age, sex, race/ethnicity), (b) teacher-level data from any fifth-grade teacher questionnaire without child-level teacher data, or (c) data from the school administrator questionnaire or school facilities checklist.
C6PW0	Fifth-grade parent interview data alone or in combination with (a) fifth-grade child assessment data, (b) data from any fifth-grade teacher questionnaire (teacher-level or child-level), or (c) data from the school administrator questionnaire or school facilities checklist. *Exception*: If data from the parent AND child assessment AND teacher (child- and/or teacher-level) are used together, then either C6CPTR0, C6CPTM0 or C6CPTS0 should be used.
C6CPTR0	Fifth-grade direct child assessment data combined with fifth-grade parent interview data AND fifth-grade teacher-level data with or without child-level data from the reading teacher, alone or in conjunction with data from the school administrator or facilities checklist.
C6CPTM0	Fifth-grade direct child assessment data combined with fifth-grade parent interview data AND fifth-grade child data from mathematics teacher (with or without teacher-level data), alone or in conjunction with data from the school administrator or facilities checklist.
C6CPTS0	Fifth-grade direct child assessment data combined with fifth-grade parent interview data AND fifth-grade child data from science teacher (with or without teacher-level data), alone or in conjunction with data from the school administrator or facilities checklist.

SOURCE: U.S. Department of Education, National Center for Education Statistics, Early Childhood Longitudinal Study, Kindergarten Class of 1998–99 (ECLS-K), spring 2004.

Weight C6CW0 is used to estimate child-level characteristics or assessment scores for fifth grade. Examples of such estimates are the percent of children who are male, the percent of children who are API, the percent of children who are 11-years-old at the beginning of the fifth-grade data collection, and the mean reading score of children in the fifth-grade data collection. These weights exist not only for children who were administered a child assessment but also for children who could not be assessed due to a disability.[2] These children were not administered the ECLS-K direct cognitive battery, but their background characteristics such as age, sex, race/ethnicity, and characteristics of their parents, teachers, classrooms, and schools are available from the parent interviews, the teacher questionnaires, the school administrator questionnaire, and the school facilities checklist. The academic and social rating scores (see chapter 3) from teachers are also available for children with disabilities, regardless of whether they completed the direct child assessment.

[2] In kindergarten and first grade, children who were not proficient in English due to a non-English or non-Spanish home language (LM/not Spanish) also had weights even though they were not administered a child assessment. In third grade and fifth grade, this is no longer applicable, since there were no children not assessed due to English language ability.

C6PW0 is used for child-level estimates associated with data collected through the parent interview. Examples are the percent of children whose mothers are currently employed, the percent of children who are in a particular type of child care, and the percent of children who have a library card. These weights should not be used for estimates solely using direct child assessment data but should be used when analyzing parent and child assessment data together. For example, they should be used when exploring the relationship between home literacy behaviors and children's reading skills.

When analyzing child assessment data in conjunction with teacher data and parent data, one of the three child-parent-teacher weights should be used. C6CPTR0 should be used if only teacher-level data are used without subject-specific child-level data. C6CPTR0 should also be used if teacher-level data are combined with child-level data from the reading teacher questionnaire. However, C6CPTM0 or C6CPTS0 should be used if child-level data from mathematics or science teacher questionnaire (respectively) are included with or without teacher-level data. Weight C6CW0 may be used when analyzing child assessment data in conjunction with teacher-level data alone. In this case, some data may be missing because some teachers did not complete the questionnaire, but these are the most appropriate weights for this type of analysis.

Here are some examples of how the child-parent-teacher weights may be used. C6CPTR0 is used when child direct assessment *and* parent data *and* teacher-level data and/or child-level reading data from teachers are combined in an analysis; for example, in the analysis of the relationship between parent education, teacher education, and children's reading knowledge and skills. If it is the children's mathematics knowledge and skills as reported by the teacher that are analyzed, then C6CPTM0 should be used. Likewise, C6CPTS0 should be used if children's science knowledge and skills as reported by the teacher are combined with direct assessment, parent and teacher-level data. These weights should not be used for estimates using only direct child assessment data or only parent interview data. An example of the use of C6CW0 is in the analysis of the relationship between children's approaches to learning as rated by their teachers and the teacher's type of teaching certification.

Careful consideration should be given to which set of weights is appropriate for the desired analysis. Using the wrong weights will result in more biased or inefficient estimates. For example, if C6CPTR0 were used in an analysis of child and teacher-level data only, then the resulting estimates will be inefficient compared to estimates using C6CW0. The lower parent response causes C6CPTR0 to result in a smaller sample with positive weights. If using C6CPTR0 with child-level data from the questionnaire filled out by the mathematics teacher, then there will be missing mathematics-related data for approximately half of the children. There may be combinations of data for which no weights were specifically developed, but all analyses should incorporate whichever weight that matches most closely.

The distribution of schools by the number of sampled students with nonzero fifth-grade cross-sectional weights (any of C6CW0, C6PW0, C6CPTR0, C6CPTM0, and C6CPTS0) and the mean number of sampled students with nonzero fifth-grade cross-sectional weights per school are useful when considering analyses using hierarchical linear modeling. These are given in table 4-13. In fifth grade, 70 percent of all schools in the sample have five or fewer ECLS-K students with nonzero fifth-grade weights; 94 percent of these schools with small numbers of children are schools to which students transferred (not in tables). For this reason, schools are classified in table 4-13 on the basis of the number of students who had never transferred schools. In other words, table 4-13 shows the clustering of children within the schools originally sampled in the base year and does not include the schools to which students subsequently transferred.

Table 4-13. Distribution of originally sampled schools by number of children with nonzero weights and by type of fifth-grade sample weights: School year 2003–04

Sample	Number of cases					Mean cases per school
	1 – 5	6 – 10	11 – 15	16 – 20	21 – 27	
Spring-fifth grade						
C6CW0	172	290	244	105	8	10
C6PW0	203	306	227	77	6	9
C6CPTR0	202	303	218	75	5	9
C6CPTM0	486	276	5	0	0	5
C6CPTS0	486	281	6	0	0	5

SOURCE: U.S. Department of Education, National Center for Education Statistics, Early Childhood Longitudinal Study, Kindergarten Class of 1998–99 (ECLS-K), spring 2004.

4.7.2 Weighting Procedures

Among the 21,357 children who were eligible for the study after the base year (21,292 base year respondents and 165 children sampled in first grade), the fifth-grade sample excluded 5,214 children as explained in section 4.5. In the weighting procedures, these excluded children are considered ineligible if they became ineligible in an earlier round (because they died or moved out of the country), as movers not subsampled for followup if they were subsampled out in previous rounds because they moved out of the original sample, or of unknown eligibility if they were hard refusal cases or if they had neither first-grade nor third-grade data. Excluded children are properly adjusted for in the weighting procedures.

As in third grade, the fifth-grade subsampling of movers continues to give more weight to children in the language minority group (i.e., movers in this group were subsampled at higher rates than non-language minority movers). Other smaller groups of movers were also subsampled at lower rates, such as selected groups of movers who were sampled in first grade (as compared with base year respondent movers), and movers who did not have full longitudinal data. Differential sampling rates of movers are presented in section 4.5. Another feature of the fifth-grade sample is the subsampling of children for the administration of the mathematics or science questionnaires as discussed in section 4.5. These features of the design are taken into account in the weighting. The weighting procedures were divided into three main stages.

The first stage of weighting was to compute an initial child weight that reflects the following:

- Adjustment of the school base weight for base year school-level nonresponse;
- Adjustment of the child weights for base year child-level nonresponse; and
- Adjustment of the base year child weight for subsampling of schools for freshening in first grade (for children sampled in first grade only).

The procedures used in this first stage are the same as in all rounds of data collection after the base year because the same sample of children (base year respondents and children sampled in first grade) is eligible for subsequent rounds of data collection. The initial child weights were extracted from the first-grade weighting file to be used in fifth grade. The procedures used for computing these weights are described again in section 4.7.3 for completeness.

The second stage of weighting was to adjust the initial child weight computed in the first stage for the following:

- Subsampling of movers; and
- Child-level nonresponse.

For the mathematics and science child-parent-teacher weights, an additional adjustment was necessary (before the second stage adjustment for the subsampling of movers and for nonresponse) to adjust for the subsampling of children for whom mathematics or science teacher data questionnaires were administered.

The third and last stage was to rake the weights adjusted in the second stage to sample-based control totals. Raking is a multivariate poststratification of the weights, explained in section 4.7.4.3.

The computation of the initial child weights is described in section 4.7.3. The subsequent weight adjustments are described in section 4.7.4. Section 4.7.5 describes the different types of weights computed for spring-fifth grade.

In general, in each adjustment to the weight, the adjustment factor is multiplied by the weight in the prior step to get the adjusted weight. This fact is not repeated in the discussions of the weight adjustments in the following sections; only the computation of the adjustment factor is discussed.

4.7.3 Computation of Spring-First Grade Initial Child Weights

As mentioned earlier, the first stage of weighting was to compute an initial child weight that reflects: (1) the adjustment of the school base weight for base year school-level nonresponse (school-level weights), (2) the adjustment of the child weights for base year child-level nonresponse (child-level weights), and (3) the adjustment of the base year child weight for subsampling of schools for freshening in first grade (child-level weights, for children sampled in first grade only). These weights were already computed for spring-first grade. For completeness, they are described below, in section 4.7.3.1 for the school-level weights, and in section 4.7.3.2 for the child-level weights.

4.7.3.1 Base Year Nonresponse-Adjusted School Weights

This weight is the same as that computed for the first-grade data collection. It was computed as the school base weight adjusted for base year school-level nonresponse. The base weight for each school was the inverse of the probability of selecting the PSU (county or group of counties), multiplied by the inverse of the probability of selecting the school within the PSU. For schools selected in the base year through the frame freshening procedure, an additional factor equal to the inverse of the selection probability of the district or diocese was included in the base weight.

A base year responding school was an original sample school with at least one child with a positive C1CW0, C2CW0, C1PW0, or C2PW0 weight. C1CW0 is positive for LM/not Spanish children,

children with disabilities and children with at least one direct cognitive test score in fall-kindergarten. C1PW0 is positive for children whose parents completed the family structure questions of the parent interview in fall-kindergarten. C2CW0 and C2PW0 weights are positive under similar circumstances except for spring-kindergarten. Schools that did not meet this condition are nonrespondents and their weights distributed across responding units (at the school level) in this stage. The base year school weight was adjusted within nonresponse weighting classes created in the base year using the Chi-squared Automatic Interaction Detector (CHAID) and variables with known values for both respondents and nonrespondents. School characteristics used for constructing nonresponse cells were the school affiliation (public, Catholic, non-Catholic religious, or nonreligious private), the school locale (large city, midsize city, suburb of large city, suburb of midsize city, large town, small town, or rural area), the region where the school is located (Northeast, Midwest, South, or West), and the size classification of the school in terms of school enrollment. Once the weighted nonresponse cells were determined, the nonresponse adjustment factors are the reciprocals of the response rates within the selected nonresponse cells.

4.7.3.2 Base Year Child Weights

As mentioned earlier, two groups of children were fielded in spring-third grade: base year respondents, and eligible children who were sampled in first grade as part of the sampling freshening procedure. The base year child weights for the two groups were the same as those computed for the first - grade year. A description of them follows.

Base year child weights for base year respondents. As previously described, a base year respondent was defined as one with at least one direct cognitive test score in fall- or spring-kindergarten or was excluded from assessment because of a disability or because the child belonged in the language minority/not Spanish group, or whose parent responded to the family structure section of the parent instrument in fall- or spring-kindergarten. In terms of weights, a base year respondent is a sampled child with a positive fall- or spring-kindergarten weight (i.e., C1CW0, C2CW0, C1PW0 or C2PW0 weights). The base year child weight is the product of the base year nonresponse-adjusted school weight and the inverse of the within school selection probability of the child, adjusted for child-level nonresponse. The nonresponse weighting classes included school characteristics from the school nonresponse adjustments such as school affiliation, locale, region, school enrollment class, and child characteristics such as age group, sex, and race/ethnicity. These weighting classes are similar to those used for the original child weights in fall- and spring-kindergarten. For a description of the computation of child weights in fall- and spring-kindergarten, see chapter 4, section 4.3.4 of *Early Childhood Longitudinal Study, Kindergarten Class of 1998–99 (ECLS-K) Base Year Public-Use Data Files and Electronic Codebook: User's Manual* (NCES 2001–029r) (Tourangeau, Burke, Lê, et al. 2004).

Base year child weights for eligible children sampled in first grade. Since each child sampled in first grade was directly linked to a child sampled in kindergarten, the first step was to compute a weight for the children who were sampled in kindergarten that reflected the school freshening subsampling and the school freshening nonresponse (some schools refused to provide information needed for freshening). This weight was then assigned to the child sampled in first grade and further adjusted for nonresponse due to not obtaining the data from the sample of freshened children (i.e., children sampled in first grade).

Part 1: School weight adjusted for subsampling of schools for freshening. First the school base year weight adjusted for school nonresponse (as described in section 4.7.3.1) was adjusted for the subsampling of schools for freshening. Student freshening was done in the same 50 percent subsample of schools that were flagged for following movers in spring-first grade. The school freshening subsampling adjustment factor was computed as:

- 0 if the school was not in the set of schools subsampled for freshening[3] and
- The sum of base year nonresponse-adjusted school weights for all schools over the sum of base year nonresponse-adjusted school weights for schools subsampled for freshening, if the school was in the set of schools subsampled for freshening.

This adjustment was done within cells defined by school affiliation and census region.

Part 2: School weight adjusted for freshening nonresponse. The freshening procedure could not be applied in all designated schools because some schools did not provide the information needed for freshening. These schools are considered freshening nonrespondents. The school weight adjusted for freshening subsampling was then adjusted for this type of nonresponse. The school freshening nonresponse adjustment factor was calculated as the sum of weights of the freshening-adjusted school weights for all schools designated for freshening over the sum of weights of the freshening-adjusted school weights for schools who responded to freshening. In both the numerator and denominator of this factor, the school measure of size was incorporated; the school measure of size is relevant because the

[3] These weights, used only to link children sampled in first grade to children sampled in kindergarten, sum up to zero in schools not subsampled for freshening, meaning that there are no children sampled in those schools through freshening.

weights will be used for child-level estimates, not school-level estimates. The nonresponse cells for this adjustment were created using school affiliation and urbanicity.

Part 3: Base year child weight. The school-adjusted weight was multiplied by the inverse of the within school selection probability of the child in the base year to obtain a base year child weight. The base year child weight was then adjusted for base year child nonresponse because children who did not respond in the base year could not be linked to children in first grade in spring 2000. The adjustment factor was computed as the sum of the base year child weights of all base year children over the sum of the base year child weights of base year respondents within each nonresponse cell. The nonresponse cells were created using school characteristics such as school affiliation, locale, region, school enrollment class, and child characteristics such as age group, sex, and race/ethnicity.

Part 4: Base year child weight adjusted for movers. Only children who did not move from their original schools were designated as links to children in the freshening procedure. The children who moved and were followed into their new schools were not identified to participate in the freshening process in their new schools. As a result, all the children who moved were considered nonrespondents to the freshening process. Additionally, nonmovers and movers who were not in first grade were not eligible for freshening (e.g., if a child was in kindergarten in spring 2000, he or she would be linked only to other kindergarten children and thus was not eligible for the freshening of first-graders). Adjustment was necessary to account for these two groups of children and was done in two steps.

In the first step, adjustment was done for movers whose grade was unknown. A portion of the movers was assumed to be in first grade. In the second step, the weights were adjusted for children who were in first grade but who were not identified to participate in the freshening process because they had moved into a new school. For this two-step adjustment, each child was classified as: (a) mover in first grade, (b) mover in another grade, (c) mover with unknown grade, (d) nonmover in first grade, and (e) nonmover in another grade.

The first step adjustment for movers whose grade was unknown was computed as

- 0 if the child was a mover with unknown grade (group c);
- 1 if the child was a nonmover, in first grade or another grade (group d or e); and
- The sum of the nonresponse-adjusted base year child weights (computed in part 3) of all movers (group a, b, or c) over the sum of the nonresponse-adjusted base year child weights of movers with known grade (group a or b), if the child was a mover with known grade (group a or b).

The second step adjustment for movers who could not be used as links for freshening was computed as

- 0 if the child was a first grade mover (group a);
- 1 if the child was in a grade other than first grade (group b or e); and
- The sum of the weights adjusted in the first step of part 4 of all first graders (group a or d) over the sum of the weights adjusted in the first step of part 4 of nonmovers in first grade (group d), if the child was a nonmover in first grade (group d).

This two-step adjustment was done within cells defined by school affiliation and census region.

The weights thus created for children sampled in kindergarten were then linked to the children who were brought into the sample in first grade through sample freshening. In other words, the weight of the child sampled in first grade was defined at this point to be the weight computed for the child sampled in kindergarten that was responsible for bringing the first-grader into the sample.

For the next step in the computation of the spring-first grade child weights, the two groups of children—base year respondents and children sampled in first grade through sample freshening—were put together, and a common variable and label were used to designate the initial child weight. This is the base year child weight as computed above for each group of children.

4.7.4 Computation of Spring-Fifth Grade Child Weights

The initial child weights described in section 4.7.3 were adjusted for movers between the base year and fifth grade and nonresponse in fifth grade, and raked to sampled-based control totals to obtain the final spring-fifth grade child weights.

4.7.4.1 Adjustment for Movers

First, the initial child weights were adjusted to reflect the subsampling of movers. In the ECLS-K, a child could move more than once and at different times. For example, a child could move out of his original sample school because the school did not have grades higher than kindergarten. Then he could move again between first and third grade, first and fifth grade, or third and fifth grade. Once a child was identified as a mover, he stayed a mover unless he moved back to the original sample school. For example, a child who moved between kindergarten and third grade, but stayed in that same school between third and fifth grade, was considered a mover for the fifth grade.

Each mover in the fifth grade had a flag indicating whether he was followed into the new school. These flags were set according to the mover subsampling plan described in section 4.5. Children who were excluded from the fifth-grade data collection because they moved out of the original schools and were subsampled out for followup in previous rounds had their flag set to "not followed." In fifth grade, children were fielded as described in exhibit 4-2.

Exhibit 4-2. Movers and nonmovers by retention status: School year 2003–04

Child moved out of original school		Child subsampled for followup		
Before fifth grade	During fifth grade	Before fifth grade	During fifth grade	Child fielded in fifth grade
No	No	†	†	Yes
No	Yes	†	No	No
No	Yes	†	Yes	Yes
Yes	No, did not move again	No	†	No
Yes	No, did not move again	Yes	No	No
Yes	No, did not move again	Yes	Yes	Yes
Yes	Back in original school	†	†	Yes

† Not applicable.
SOURCE: U.S. Department of Education, National Center for Education Statistics, Early Childhood Longitudinal Study, Kindergarten Class of 1998–99 (ECLS-K), fall 1998, spring 1999, fall 1999, spring 2000, spring 2002, and spring 2004.

The initial child weight described in section 4.7.3.2 was adjusted to reflect the subsampling of movers. The adjustment factor for subsampling movers (who moved before or during fifth grade) was computed as follows:

- 1 if the child was not a mover;
- 0 if the child was a mover and the value of the follow flag was 0 (i.e., not to follow); and
- The sum of initial child weights of children who were movers over the sum of initial child weights of children who were movers and whose follow flags have value 1, if the child was a mover whose follow flag has value 1.

For the third category, the adjustment factor was computed within cells created using the following characteristics: whether children were sampled in kindergarten or first grade, and whether they were language minority children.[4] Twelve children with large weights had their weights trimmed by 40 percent. However, the weights were not redistributed because the total sum of weights was re-established in the raking procedure that came later.

4.7.4.2 Adjustment for Nonresponse

After the adjustment for subsampling movers, the child weights were adjusted for nonresponse. As in spring-first grade and spring-third grade, the nonresponse adjustment was done in two steps. In the first step, the adjustment was for children whose eligibility was not determined (unknown eligibility). A portion of children of unknown eligibility was assumed to be ineligible, equal to the proportion of children of known eligibility who are ineligible. In the second step, the adjustment was for eligible nonrespondents. To carry out these adjustments, each child was classified as (a) an eligible respondent, (b) an eligible nonrespondent, (c) ineligible (out of the country or deceased) or (d) of unknown eligibility (mover who could not be located). The first adjustment factor (for children of unknown eligibility) was computed as

- 0 if the child was of unknown eligibility (group d); and
- The sum of the mover adjusted weights of all children (any group) over the sum of the mover adjusted weights of children who were eligible respondents, eligible

[4] Fewer characteristics were used than in previous years to create cells for mover adjustments. This is due to cells with a small number of records, requiring them to be collapsed in order to avoid large adjustment factors. This resulted in fewer cells, hence fewer characteristics being used.

nonrespondents or ineligible (group a, b, or c), if the child was not of unknown eligibility.

The second adjustment factor (for eligible nonrespondents) was computed as:

- 0 if the child was an eligible nonrespondent (group b); and
- The sum of the weights adjusted in the first step of eligible children (group a or b) over the sum of the weights adjusted in the first step of eligible responding children (group a), if the child was an eligible respondent.

In both steps of the adjustment, separate nonresponse classes were created for movers and nonmovers using various combinations of response status of child assessments and parent interviews in the base year as well as whether children belong to the language minority group, the type of household collected from the parent interviews (all cross-sectional weights except C6CW0), and the school affiliation including whether the child was homeschooled (C6CPTR0, C6CPTM0 and C6CPTS0 only).

4.7.4.3 Raking to Sample-Based Control Totals

To reduce the variability due to the subsampling of schools and movers, the child weights were then raked to sample-based control totals computed using the initial child weights computed in section 4.7.3. The child records included in the file used for computing the control totals are records of base year respondents and records of eligible children sampled in first grade, including records of children who became ineligible in spring-fifth grade. The sum of weights thus calculated is the estimated number of children who were in kindergarten in 1998–99 or first grade in 1999–2000. In the previous steps, the weights of the nonresponding children were distributed to the responding children while the weights of the ineligible children were not affected by this weighting step. The weights of the ineligible children are set to zero at the end of this process because these children are not included in the analysis of the spring-fifth-grade data. The reason for including the ineligible children in the raking step is that these children were included in the sampled-based control totals.

The raking factor was computed separately within raking cells as the sample-based control total for the raking cell over the sum of the nonresponse-adjusted weights for children in the same cell. Raking cells (also known as raking dimensions) were created using school and child characteristics collected in the base year or first-grade year: school affiliation, region, urbanicity, sex, age, race/ethnicity, SES, language minority status, whether sampled in kindergarten or first grade, and if sampled in kindergarten, mover status.

4.7.4.4 Additional Adjustment for Child-Parent-Teacher Cross-Sectional Weights

In all three child-parent-teacher weights described in section 4.7.1, the presence of at least one completed teacher-level questionnaire is the factor that determines whether the child would have a positive child-parent-teacher weight in the two subjects to which they were assigned (i.e., reading and mathematics, or reading and science). A child could have one teacher who taught all subjects, in which case the teacher was asked to fill out both the reading questionnaire and the mathematics questionnaire (if the child was selected for mathematics) or science questionnaire (if the child was selected for science). A child could also have different teachers teaching different subjects, in which case the child may have a reading teacher filling out the reading questionnaire and a mathematics teacher filling out the mathematics questionnaire, and both teachers could have filled out the teacher-level questionnaire. Because of the subsampling, no children have completed both the mathematics and the science questionnaires.

Table 4-14 shows the distribution of children who have direct child assessment data, parent interview data and child-level data from the mathematics teacher by the number of teachers they had who filled out the teacher-level questionnaire. The first column in this table shows the number of teachers that each child had: only one teacher who taught both reading and mathematics, or two teachers, one teaching reading and the other teaching mathematics. The second column shows the type of teacher who filled out the teacher-level questionnaire. If the child had only one teacher, then it was this teacher–identified in the table as the reading teacher–who filled out the teacher-level questionnaire (3,142 cases out of 5,009 or 63 percent). If the child had two teachers, then in the majority of cases, both teachers filled out the teacher-level questionnaire (1,803 cases out of 5,009 or 36 percent). There are very few cases where only one of the two teachers filled out the teacher-level questionnaire. Table 4-15 shows the same information for science. Since C6CPTM0 and C6CPTS0 are used for the analysis of child and parent data with data from mathematics and science teachers, another option to define these weights is to use the presence of child-level data from the mathematics/science teachers. However, tables 4-14 and 4-15 show that by considering the presence of teacher-level data in constructing the child-parent-teacher weights, there are more records with positive weights for analysis (5,017 as shown in table 4-17 compared with 5,009 in table 4-14 for C6CPTM0; and 5,103 as shown in table 4-17 compared with 5,088 in table 4-15 for

C6CPTS0). Using teacher-level data to define the child-parent-teacher weights is also consistent with previous years' practice.

Table 4-14. Number of children with direct child assessment, parent interview and child-level data from mathematics teacher, by number of teachers who filled out teacher-level questionnaire: School year 2003–04

Number of teachers that each child had	Teachers who completed teacher level questionnaire	Number of children with child-parent-mathematics data from the child-level mathematics questionnaire
Total		5,009
1	Reading	3,142
2	Reading	25
2	Math	39
2	Reading and Math	1,803

SOURCE: U.S. Department of Education, National Center for Education Statistics, Early Childhood Longitudinal Study, Kindergarten Class of 1998–99 (ECLS-K), spring 2004.

Table 4-15. Number of children with direct child assessment, parent interview and child-level data from science teacher, by number of teachers who filled out teacher-level questionnaire: School year 2003–04

Number of teachers that each child had	Teachers who completed teacher level questionnaire	Number of children with child-parent-science data from the child-level science questionnaire
Total		5,088
1	Reading	2,999
2	Reading	42
2	Science	35
2	Reading and Science	2,012

SOURCE: U.S. Department of Education, National Center for Education Statistics, Early Childhood Longitudinal Study, Kindergarten Class of 1998–99 (ECLS-K), spring 2004.

An additional adjustment is necessary to adjust for the subsampling of children for whom mathematics or science teacher data questionnaires were administered. For the child-parent-mathematics teacher weight, this adjustment (before adjustment for movers and nonresponse adjustments, described in section 4.7.4.1 and 4.7.4.2, respectively) was computed as:

- 0 if the child was sampled for science rather than mathematics; and

- The sum of the initial child weights of all children over the sum of the initial child weights of children who were sampled for mathematics questionnaires.

Similarly, for the child-parent-science teacher weight, this adjustment was computed as:

- 0 if the child was sampled for mathematics rather than science; and
- The sum of the initial child weights of all children over the sum of the initial child weights of children who were sampled for science questionnaires.

4.7.5 Types of Cross-Sectional Weights and Their Use

The different types of cross-sectional weights are described in section 4.7.1 and their use is summarized in exhibit 4-1. They were all created as described in sections 4.7.4.2 and 4.7.4.3, but the definition of which children were eligible respondents varied for the different weights. The adjustment for movers was done once for all weights, and then the resulting weights were adjusted for nonresponse and raked separately for C6CW0, C6PW0, C6CPTR0, C6CPTM0, and C6CPTS0.

4.7.5.1 Cross-Sectional Weights to Be Used With Direct Child Assessment Data (C6CW0)

In spring-fifth grade, responding children for this type of weight were eligible children who had spring-fifth grade scorable direct child cognitive assessment data, or children with disabilities who, according to specifications in their Individualized Education Plan, could not participate in the assessments. A child was eligible if he or she was a base year respondent or freshened in first grade. Children who transferred to schools and were not flagged to be followed, who moved out of the country or were deceased were ineligible. In spring-fifth grade, responding children were classified using rules similar to those used in spring-first grade and spring-third grade.

Table 4-16 shows the number of children who were not assessed due to the following special situations: children with disabilities, children who had moved out of their original sample schools and were not flagged to be followed, children who had moved and were flagged to be followed but could not be located or moved into a school in a nonsampled county, and children who had moved outside of the country or who were deceased. Of these, only children with disabilities had weights included in the fifth grade data file. Note that the number of children who were nonlocatable and who moved nonsampled

PSUs (thus not assessed) is much smaller in fifth grade than in third grade. This is because hard-to-field children (hard refusals, children with neither first-grade nor third-grade data, and movers previously not followed) were excluded from the fifth-grade data collection as explained in section 4.5.

Table 4-16. Number of children who were not assessed in spring-fifth grade, by special situations: School year 2003–04

Special situation	Number of children: Unweighted	Weighted
Spring-fifth grade		
Children with disabilities[1]	63	29,463
Moved from original sample schools		
Subsampled, not to be followed	7,880	1,477,091
Nonlocatable or moved to nonsampled PSU	676	128,142
To be followed but were ineligible in spring-fifth grade	39	6,607

[1] These children's individualized education plans (IEPs) specifically prohibited assessments.
SOURCE: U.S. Department of Education, National Center for Education Statistics, Early Childhood Longitudinal Study, Kindergarten Class of 1998–99 (ECLS-K), spring 2004.

4.7.5.2 Cross-Sectional Weights To Be Used With Parent Data (C6PW0)

The weight C6PW0 is to be used with parent interview data. In spring-fifth grade, a respondent was defined as a child for whom the family structure section (FSQ) in that child's parent interview for the corresponding round was completed. Note that this weight is at the child level even though the data were collected from the parents; they sum to fifth-grade children, not to the parents of fifth-grade children.

4.7.5.3 Cross-Sectional Weights To Be Used With a Combination of Child Direct Assessment Data and Parent Interview Data and Teacher Data for Children With Reading Teacher Questionnaire (C6CPTR0)

The weight C6CPTR0 is to be used for analysis involving all children with child, parent, and teacher-level data. If child-level data from reading teachers are included in the analysis, then the same weight C6CPTR0 should be used. A respondent for this type of weight was defined as a child who had scorable cognitive assessment data for spring-fifth grade (or excluded from direct assessment due to a disability), whose parent completed the FSQ section of the parent interview for spring-fifth grade, and who had completed teacher-level data either from the reading teacher and/or the mathematics/science teacher.

4.7.5.4 Cross-Sectional Weights To Be Used With a Combination of Child Direct Assessment Data and Parent Interview Data and Teacher Data for Children With Mathematics Teacher Questionnaire (C6CPTM0)

The weight C6CPTM0 is to be used for analysis involving children who were subsampled to have a mathematics teacher questionnaire and who had child, parent, and child-level data from mathematics teachers (with or without teacher-level data). A respondent for this type of weight was defined as a child who had scorable cognitive assessment data for spring-fifth grade (or excluded from direct assessment due to a disability), whose parent completed the FSQ section of the parent interview for spring-fifth grade, and who had completed teacher-level data either from the reading teacher or the mathematics teacher. If there are mathematics data but no teacher-level data, then C6CPTM0 is zero and such a case would not be included in the analysis. See section 4.7.1 for how the child-parent-teacher weights are defined.

4.7.5.5 Cross-Sectional Weights To Be Used With a Combination of Child Direct Assessment Data and Parent Interview Data and Teacher Data for Children With Science Teacher Questionnaire (C6CPTS0)

The weight C6CPTS0 is to be used for analysis involving children who were subsampled to have a science teacher questionnaire and who had child, parent, and child-level data from science teachers (with or without teacher-level data). A respondent for this type of weight was defined as a child who had scorable cognitive assessment data for spring-fifth grade (or excluded from direct assessment due to a disability), whose parent completed the FSQ section of the parent interview for spring-fifth grade, and who had completed teacher-level data either from the reading teacher or the science teacher. If there are science data but no teacher-level data, then C6CPTS0 is zero and such a case would not be included in the analysis. See section 4.7.1 for how the child-parent-teacher weights are defined.

4.7.6 Replicate Weights

For each weight included in the data file, a set of replicate weights was calculated. Replicate weights are used in the jackknife replication method to estimate the standard errors of survey estimates. All adjustments to the full sample weights were repeated for the replicate weights.

For spring-fifth grade, there are 90 replicate weights. Each set of replicate weights has the same prefix in the variable name as the full sample weight. For example, the replicate weights for C6CW0 are C6CW1 through C6CW90. The method used to compute the replicate weights and how they are used to compute the sampling errors of the estimates are described in section 4.8.

4.7.7 Characteristics of Cross-Sectional Sample Weights

The statistical characteristics of the sample weights are presented in table 4-17. For each type of weight, the number of cases with nonzero weights is presented together with the mean weight, the standard deviation, the coefficient of variation (i.e., the standard deviation as a percentage of the mean weight), the minimum weight, the maximum weight, the skewness, the kurtosis, and the sum of weights.

The difference in the estimate of the population of students (sum of weights) between rounds of data collection and between types of weight is due a combination of factors, among them: (1) the number of first-graders or third-graders who became ineligible in fifth grade (due to death, leaving the country, or being a nonsampled mover), and (2) the adjustment of the weights for the children of unknown eligibility.

4.8 Variance Estimation

The precision of the sample estimates derived from a survey can be evaluated by estimating the variances of these estimates. For a complex sample design such as the one employed in the ECLS-K, replication and Taylor Series methods have been developed. These methods take into account the clustered, multistaged characteristics of sampling and the use of differential sampling rates to oversample targeted subpopulations. For the ECLS-K, in which the first-stage self-representing sampling units, (i.e., PSUs) were selected with certainty and the first-stage non-self-representing sampling units were selected with two units per stratum, the paired jackknife replication method (JK2) is recommended. This section describes the JK2 and the Taylor Series estimation methods.

Table 4-17. Characteristics of the fifth-grade cross-sectional child-level weights: School year 2003–04

Sample	Number of cases	Mean	Standard deviation	CV (× 100)	Minimum	Maximum	Skewness	Kurtosis	Sum
C6CW0	11,346	346.92	552.91	159.38	1.91	6556.07	4.36	23.64	3,936,156
C6PW0	10,996	357.86	501.99	140.28	1.80	4909.08	3.54	15.06	3,935,007
C6CPTR0	10,120	388.86	653.95	168.17	1.89	6707.74	4.21	21.04	3,935,285
C6CPTM0	5,017	786.58	1087.08	138.20	6.10	9887.78	4.24	21.85	3,946,286
C6CPTS0	5,103	770.41	1071.77	139.12	4.94	9883.96	4.15	20.55	3,931,397

SOURCE: U.S. Department of Education, National Center for Education Statistics, Early Childhood Longitudinal Study, Kindergarten Class of 1998–99 (ECLS-K), spring 2004.

4.8.1 Paired Jackknife Replication Method

In this method, a survey estimate of interest 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 about 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 (Wolter 1985):

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

where

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

The variance estimates of selected survey items presented in section 4.9.2 were produced using WesVar and JK2 (Westat 2001).

Replicate weights were created to be used in the calculation of variance estimates. Each replicate weight was calculated using the same adjustment steps as the full sample weight but using only the subsample of cases that constitute each replicate. For the original ECLS-K design in the base year, replicate weights were created taking into account the Durbin method of PSU selection. The Durbin method selects two first-stage units per stratum without replacement, with probability proportional to size and a known joint probability of inclusion (Durbin 1967).

In the ECLS-K PSU sample design, there were 24 SR strata and 38 NSR strata. Among the 38 NSR strata, 11 strata were identified as Durbin strata[5] and were treated as SR strata for variance estimation. The purpose of the Durbin strata is to allow variances to be estimated as if the first-stage units were selected with replacement. This brings the number of SR PSUs to 46 (24 original SR PSUs and 22 Durbin PSUs from the 11 Durbin strata). The remaining 54 NSR PSUs are in 27 NSR strata; thus 27 replicates were formed, each corresponding to one NSR stratum. For the SR strata, 63 replicates were formed. The 90 replicates will yield about 76 degrees of freedom for calculating confidence intervals for many survey estimates.

As stated earlier, the sample of PSUs was divided into 90 replicates or variance strata. The 27 NSR strata formed 27 variance strata of two PSUs each; each PSU formed a variance unit within a variance stratum. All schools within an NSR PSU were assigned to the same variance unit and variance stratum. Sampled schools in the 46 SR PSUs were grouped into 63 variance strata. In the SR PSUs, schools were directly sampled and constituted PSUs. Public schools were sampled from within PSU while private schools were pooled into one sampling stratum and selected systematically (except in the SR PSUs identified through the Durbin method where private schools were treated as if they were sampled from within PSU). Schools were sorted by sampling stratum, school affiliation (from the original sample or newly selected as part of freshening), type of frame (for new schools only), and their original order of selection (within stratum). From this sorted list, they were grouped into pairs within each sampling stratum; the last pair in the stratum may be a triplet if the number of schools in the stratum is odd. This operation resulted in a number of ordered preliminary variance strata of two or three units each. The first ordered 63 strata were then numbered sequentially from 1 to 63; the next ordered 63 strata were similarly numbered, and so on until the list was exhausted, thus forming the desired 63 variance strata.

---

[5] For a description of the Durbin method, see *Early Childhood Longitudinal Study, Kindergarten Class of 1998–99 (ECLS-K) Third Grade Methodology Report* (NCES 2005–018) (Tourangeau, Brick, Byrne, et al. 2004).

In strata with two units, a unit being a PSU in the case of NSR PSUs and a school in the case of SR PSUs, the base weight of the first unit was doubled to form the replicate weight, while the base weight of the second unit was multiplied by zero. In strata with three units, two variance strata were created: in the first variance stratum, the base weight of two of the three units was multiplied by 1.5 to form the replicate weight and the base weight of the last unit was multiplied by zero; in the second variance stratum, the base weight of a different group of two units was multiplied by 1.5, and the base weight of the third unit was multiplied by zero. Multiplying the base weight in a unit by zero is equivalent to dropping one unit as required by the jackknife method. All adjustments to the full sample weights were repeated for the replicate weights. For each full sample weight, there are 90 replicate weights with the same weight prefix.

A child sampled in first grade through the freshening process was assigned to the same replicate as the originally sampled child to whom the child was linked. When the child sampled in first grade was assigned a full sample weight (see section 4.7.3.2), he or she was assigned the replicate weights in the same manner.

To reflect the variability of the control totals in the sample-based raking, a set of replicate control totals was created. Each replicate was then raked to the corresponding replicate-based control totals. This resulted in each replicate retaining the variability associated with the original sample estimates of the control totals.

The replicate weights can be used with software such as WesVar (http://www.westat.com/wesvar/), SUDAAN (*SUDAAN Language Manual, Release 9.0* [Research Triangle Institute 2004 or http://www.rti.org/sudaan/], and AM (http://am.air.org).

4.8.2 Taylor Series Method

The Taylor Series method produces a linear approximation of the survey estimate of interest; then the variance of the linear approximation can be estimated by standard variance formulas (Wolter 1985). The stratum and first-stage unit (i.e., PSU) identifiers needed to use the Taylor Series method were assigned, taking care to ensure that there were at least two responding units in each stratum. A stratum that did not have at least two responding units was combined with an adjacent stratum. For the ECLS-K, the method of stratifying first-stage units was the same for each type of cross-sectional weight. For each type of weight, the sample size was examined, and then strata were combined when the sample size was not adequate. The sequential numbering of strata and first-stage units was done separately for each weight. Consequently, there is a different set of stratum and first-stage unit identifiers for each set of weights.

Stratum and first-stage unit identifiers are provided as part of the ECLS-K data file and can be used with software such as SUDAAN, Stata, SAS, SPSS, or AM. They are described in table 4-18.

Table 4-18. ECLS-K Taylor Series stratum and first-stage unit identifiers: School year 2003-04

Variable name	Description
C6TCWSTR	Sampling stratum—spring-fifth grade C-weights
C6TCWPSU	First-stage sampling unit within stratum—spring-fifth grade C-weights
C6TPWSTR	Sampling stratum—spring-fifth grade P-weights
C6TPWPSU	First-stage sampling unit within stratum—spring-fifth grade P-weights
C6CPTRST	Sampling stratum—spring-fifth grade CPTR-weights
C6CPTRPS	First-stage sampling unit within stratum—spring-fifth grade CPTR-weights
C6CPTMST	Sampling stratum—spring-fifth grade CPTM-weights
C6CPTMPS	First-stage sampling unit within stratum—spring-fifth grade CPTM-weights
C6CPTSST	Sampling stratum—spring-fifth grade CPTS-weights
C6CPTSPS	First-stage sampling unit within stratum—spring-fifth grade CPTS-weights

SOURCE: U.S. Department of Education, National Center for Education Statistics, Early Childhood Longitudinal Study, Kindergarten Class of 1998–99 (ECLS-K), spring 2004.

4.8.3 Specifications for Computing Standard Errors

Specifications for computing standard errors (SEs) are given in table 4-19. For each type of analysis described in the table, users can choose the replication method or the Taylor Series method for computing SEs.

For the replication method, the full sample weight, the replicate weights, and the method of replication are required parameters. All analyses of the ECLS-K data should be done using JK2. As an example, to compute spring-fifth grade child-level estimates (e.g., mean reading scores) and their SEs, users need to specify CHILDID in the ID box of the WesVar data file screen, C6CW0 as the full sample weight, C6CW1 to C6CW90 as the replicate weights, and JK2 as the method of replication.

Table 4-19. Specifications for computing standard errors, spring-fifth grade: School year 2003–04

		Computing standard errors					Approximating sampling errors
		Replication method (WesVar, SUDAAN or AM)			Taylor Series method (SUDAAN, Stata, SAS, SPSS or AM)		DEFT (Average root design effect)
Type of analysis	Full sample weight	ID	Replicate weights	Jackknife method	Sample design[1]	Nesting variables	
Spring-fifth grade							2.039
cross-sectional	C6CW0	CHILDID	C6CW1 – C6CW90	JK2	WR	C6TCWSTR C6TCWPSU	
	C6PW0	CHILDID	C6PW1 – C6PW90	JK2	WR	C6TPWSTR C6TPWPSU	
	C6CPTR0	CHILDID	C6CPTR1 – C6CPTR90	JK2	WR	C6CPTRST C6CPTRPS	
	C6CPTM0	CHILDID	C6CPTM1-C6CPTM90	JK2	WR	C6CPTMST C6CPTMPS	
	C6CPTS0	CHILDID	C6CPTS1-C6CPTS90	JK2	WR	C6CPTSST C6CPTSPS	

[1] WR = with replacement, specified only if using SUDAAN. WR is the only option available if using SAS, Stata, SPSS, or AM.

SOURCE: U.S. Department of Education, National Center for Education Statistics, Early Childhood Longitudinal Study, Kindergarten Class of 1998–99 (ECLS-K), spring 2004.

For the Taylor Series method using SUDAAN, Stata, SAS, SPSS, or AM, the full sample weight, the sample design, the nesting stratum and PSU variables are required. For the same example above, the full sample weight (C6CW0), the stratum variable (C6TCWSTR), and the PSU variable (C6TCWPSU) must be specified. The "with replacement" sample design option, WR, must also be specified if using SUDAAN.

The last column in table 4-19 gives the average root design effect that can be used to approximate the SEs for each type of analysis. For a discussion of the use of design effects, see section 4.9.1.

4.9 Design Effects

An important analytic device is to compare the statistical efficiency of survey estimates from a complex sample survey such as the ECLS-K, with what would have been obtained in a hypothetical and usually impractical simple random sample (SRS) of the same size. In a stratified clustered design like the ECLS-K, 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. The basic measure of the relative efficiency of the sample is the *design effect*, 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 root design effect, *DEFT*, is defined as:

$$DEFT = \sqrt{DEFF} = \frac{SE_{DESIGN}}{SE_{SRS}},$$

where *SE* is the standard error of the estimate.

4.9.1 Use of Design Effects

Methods of computing SEs for the ECLS-K are jackknife replication and Taylor Series linearization. If statistical analyses are conducted using software packages that assume the data were collected using simple random sampling, the SEs will be calculated under this assumption and should be corrected using *DEFT*.[6] The SE of an estimate under the actual sample design can be approximated 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: $$Var_{SRS} = \frac{1}{n} \frac{\sum_{1}^{n} w_i \left(x_i - \overline{x_w}\right)^2}{\sum_{1}^{n} w_i} = SE_{SRS}^2,$$

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

$$\overline{x}_w = \frac{\sum_{1}^{n} w_i x_i}{\sum_{1}^{n} w_i}.$$

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.

[6] Common procedures in SAS, SPSS and Stata assume simple random sample. Use the SVY procedure (SAS), the Complex Samples module (SPSS), or the SURVEY command (Stata) to account for complex samples.

In both cases of means and proportions, the SE assuming SRS should be multiplied by *DEFT* to get the approximate standard error of the estimate under the actual design.

4.9.2 Median Design Effects for the ECLS-K

In the ECLS-K, a large number of data items were collected from students, parents, teachers, and schools. Each item has its own design effect that can be estimated from the survey data. Typically, standard errors and design effects are presented for selected items from the study to allow analysts to see the range of standard errors and design effects that can be expected. Another way to produce design effects for analysts' use is to produce median design effects for the same set of selected items, at the overall level and for selected subgroups.

Table 4-20 shows estimates, SEs, and design effects for 52 means and proportions that were selected from the ECLS-K fifth-grade child, parent, child-level teacher, and school data. It is from this set of selected items that median design effects were computed for subgroups and presented in table 4-21.

For each survey item, Table 4-20 presents the number of cases for which data are nonmissing, the estimate, the standard error taking into account the actual sample design (Design SE), the standard error assuming SRS (SRS SE), the root design effect (DEFT), and the design effect (DEFF). Standard errors (Design SE) were produced in WesVar using JK2 based on the actual ECLS-K complex design. For each survey item, the variable name as it appears in the ECLS-K fifth-grade Electronic Codebook (ECB) is also provided in the table. 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-K, and chapter 7, which has a detailed discussion of the other variables.

The survey items were selected so that there was a mix of items from the direct child assessment, the parent interview, and the subject specific child-level teacher questionnaire. They include the different scores from the direct child assessment, the scores from the self-described child questionnaire, the social rating scores as provided by teachers, characteristics of the parents, and characteristics of the students as reported by the parents and teachers. For a small number of estimates, the data were subset to cases where the estimate is applicable; for example, the proportion of children who have access to the Internet is only for children in households with a computer.

Table 4-20. ECLS-K standard errors and design effects by selected child and parent variables, for the full sample—child, parent and child-level teacher questionnaire data: School year 2003–04

Survey item	Variable name	Number of cases	Estimate	Design SE[1]	SRS SE[2]	DEFT[3]	DEFF[4]
Child scores (mean)							
Reading scale score	C6R3RSCL	11,265	136.71	0.544	0.229	2.379	5.661
Mathematics scale score	C6R3MSCL	11,274	111.22	0.568	0.211	2.693	7.253
Science scale score	C6R1SSCL	11,270	56.11	0.354	0.140	2.524	6.372
Self-described : Externalizing problems	C6SDQEXT	11,279	1.89	0.014	0.006	2.164	4.683
Self-described : Internalizing problems	C6SDQINT	11,279	2.08	0.013	0.006	2.193	4.810
Self-described : Competence in math	C6SDQMTC	11,279	2.92	0.013	0.007	1.735	3.011
Self-described : Competence in peer relation	C6SDQPRC	11,279	2.98	0.011	0.006	1.866	3.483
Self-described : Competence in reading	C6SDQRDC	11,279	3.00	0.013	0.007	1.879	3.529
Self-described : Competence in all subjects	C6SDQSBC	11,279	2.71	0.012	0.006	2.007	4.030
Approaches to learning-Teacher	T6LEARN	10,752	2.99	0.013	0.007	1.966	3.865
Self-control-Teacher	T6CONTRO	10,648	3.19	0.013	0.006	2.176	4.735
Interpersonal-Teacher	T6INTERP	10,526	3.02	0.013	0.006	2.054	4.220
Externalizing problems-Teacher	T6EXTERN	10,690	1.71	0.012	0.006	2.093	4.380
Internalizing problems-Teacher	T6INTERN	10,574	1.68	0.013	0.005	2.429	5.902
Child and parent characteristics from parent interview (percent)							
Lived in single parent family	P6HFAMIL	10,996	27.70	0.817	0.427	1.914	3.663
Lived in two-parent family	P6HFAMIL	10,996	69.77	0.884	0.438	2.019	4.077
Mom worked 35 hours+/week	P6HMEMP	8,175	67.36	0.867	0.519	1.671	2.793
Primary care is center-based	P6PRIMNW	3,572	28.87	1.403	0.758	1.850	3.422
Primary care is home-based	P6PRIMNW	3,572	71.13	1.403	0.758	1.850	3.422
Parents had high school or less	W5PARED	10,996	31.73	0.813	0.444	1.831	3.353
Household income category below median	W5INCCAT	10,996	49.15	1.034	0.476	2.170	4.708
Parent attended PTA	P6ATTENP	10,980	39.46	1.200	0.467	2.572	6.615
Visited library	P6LIBRAR	10,968	49.22	0.978	0.477	2.049	4.200
Used computer 1-2 times per week	P6HOMECM						
	P6COMPWK	9,299	35.21	0.994	0.495	fs2.008	4.031
Had internet access	P6HOMECM						
	P6INTACC	9,089	88.16	0.510	0.339	1.504	2.262
Used computer 1-2 times per week for homework	P6HOMECM						
	P6CMPEDU	9,080	55.91	0.819	0.521	1.571	2.469
Had family rule for TV	P6TVHOME						
	P6TVRULE	10,919	89.03	0.567	0.299	1.894	3.589
Had someone help with reading homework	P6HELPR	10,835	97.68	0.293	0.145	2.025	4.102
Talked to child about day at school every day	P6OFTTLK	10,952	82.49	0.676	0.363	1.862	3.466
Talked to child about smoking 3+ times/year	P6TLKSMK	10,953	72.99	0.705	0.424	1.663	2.765
Talked to child about alcohol 3+ times/year	P6TLKALC	10,950	65.45	0.813	0.454	1.789	3.202
Took away privilege when child angry	P6HITPRV	10,829	69.30	1.193	0.443	2.693	7.250
Self-reported in very good health	P6HEALTH	10,695	88.14	0.660	0.313	2.110	4.453
Household received food stamp in last 12 months	P6FSTAMP	10,897	16.65	0.809	0.357	2.268	5.145
Child characteristics from teacher questionnaire (percent)							
Child was in fifth grade	T6GLVL	11,346	85.96	0.936	0.326	2.869	8.233
Participated fully in grade-level assessment	G6ASSMT	10,390	86.51	0.959	0.335	2.862	8.190
Parents attended regularly-scheduled conferences	G6REGCON	10,272	83.98	0.803	0.362	2.219	4.923

See notes at end of table.

Table 4-20. ECLS-K standard errors and design effects by selected child and parent variables, for the full sample—child, parent and child-level teacher questionnaire data: School year 2003–04—Continued

Survey item	Variable name	Number of cases	Estimate	Design SE[1]	SRS SE[2]	DEFT[3]	DEFF[4]
Child characteristics from teacher questionnaire (percent)-continued							
Child usually worked to best ability in reading	G6ABIL	10,756	57.03	1.041	0.477	2.181	4.758
Child was average in language skills	G6RTLANG	10,741	71.54	0.814	0.435	1.870	3.496
Child was in reading class entire school year	G6LNGTM	10,760	82.93	0.797	0.363	2.198	4.831
Child usually worked to best ability in math	M6ABIL	4,960	55.61	1.120	0.706	1.587	2.519
Child was average in mathematics skills	M6RTMTN	4,956	72.53	1.176	0.634	1.855	3.442
Child was in mathematics class entire school year	M6LNGTM	4,950	82.52	0.970	0.540	1.797	3.229
Child usually worked to best ability in science	N6ABIL	4,993	54.30	1.095	0.705	1.554	2.414
Child was average in science studies	N6RTSKIL	4,997	75.57	0.899	0.607	1.480	2.189
Child was in science class entire school year	N6LNGTM	4,999	82.90	1.000	0.532	1.878	3.528
Child characteristics (mean)							
Age of child in months	R6AGE	11,281	134.86	0.105	0.045	2.343	5.490
Child's BMI	C6BMI	11,067	20.68	0.076	0.045	1.680	2.824
Child's household size	P6HTOTAL	10,996	4.55	0.026	0.013	1.929	3.721
Number of children <18 in child's HH	P6LESS18	10,996	2.53	0.028	0.012	2.293	5.258
Number of siblings in HH	P6NUMSIB	10,996	1.57	0.022	0.011	2.013	4.053
Number of hours watched TV after dinner	P6TVAFDH	10,909	1.09	0.016	0.008	1.974	3.895
Median						2.008	4.031
Mean						2.039	4.268
Standard deviation						0.332	1.432
Coefficient of variation						0.163	0.336
Minimum						1.480	2.189
Maximum						2.869	8.233

[1] Design SE is the standard error under the ECLS-K sample design. For an explanation of this statistic, see section 4.9.
[2] SRS SE is the standard error assuming simple random sample. For an explanation of this statistic, see section 4.9.
[3] DEFT is the root design effect. For an explanation of DEFT, see section 4.9.
[4] DEFF is the design effect. For an explanation of DEFF, see section 4.9.
SOURCE: U.S. Department of Education, National Center for Education Statistics, Early Childhood Longitudinal Study, Kindergarten Class of 1998–99 (ECLS-K), spring 2004.

Table 4-21 presents the median design effects from the same 52 survey items for subgroups based on school affiliation, child's sex and race/ethnicity, geographic region, level of urbanicity, and the socioeconomic scale (SES quintiles) of the parents.

Table 4-21. ECLS-K median design effects for subgroups: School year 2003–04

Subgroups	Spring-fifth grade	
	DEFT[1]	DEFF[2]
All students	2.008	4.031
School affiliation[3]		
Public	1.899	3.605
Private	2.269	5.145
Catholic private	2.433	5.921
Other private	2.032	4.129
Sex		
Male	1.893	3.582
Female	2.025	4.100
Race/ethnicity		
White	1.969	3.879
Black	1.741	3.031
Hispanic	1.576	2.484
Asian	1.779	3.165
Pacific Islander	1.390	1.933
American Indian	1.327	1.761
Other	1.659	2.752
Region		
Northeast	2.056	4.225
Midwest	2.153	4.637
South	1.965	3.862
West	1.825	3.330
Urbanicity		
Central city	1.946	3.785
Urban fringe and large town	1.927	3.712
Small town and rural area	1.976	3.903
SES quintiles		
First (lowest)	1.723	2.967
Second	1.806	3.259
Third	1.816	3.296
Fourth	1.853	3.434
Fifth (highest)	1.930	3.723

[1] DEFT is the root design effect. For an explanation of DEFT, see section 4.9.
[2] DEFF is the design effect. For an explanation of DEFF, see section 4.9.
[3] The categories of school affiliation in this table do not match categories of school affiliation in other tables in this chapter. This is to allow users to compare median DEFT and DEFF in fifth grade with those in previous years.
NOTE: Each median is based on 52 items.
SOURCE: U.S. Department of Education, National Center for Education Statistics, Early Childhood Longitudinal Study, Kindergarten Class of 1998–99 (ECLS-K), spring 2004.

In spring-fifth grade, as in first and third grades, design effects are not computed for items from the teacher-level and school administrator's questionnaires since there are no teacher or school weights computed for any of the ECLS-K years after kindergarten. Although SEs and design effects may also be calculated for the teacher and school administrator's questionnaires at the child level, they are quite large compared to those typically found for the ECLS-K data. Design effects for teacher and school items are large because the intraclass correlation is 100 percent for children in the same school and very high for children in the same class; children attending the same school have the same school data, and children in the same class have the same teacher data.

5. DATA COLLECTION METHODS AND RESPONSE RATES

The following sections discuss the data collection procedures and response rates in the fifth-grade data collection phase of the Early Childhood Longitudinal Study, Kindergarten Class of 1998–99 (ECLS-K). Section 5.1 gives an overview of the data collection methods. Detailed information is provided on study training procedures (section 5.2), preassessment school contacts (section 5.3), spring-fifth-grade data collection (section 5.5), and quality control procedures (section 5.6). Spring-fifth grade completion rates are presented and discussed in section 5.7.

5.1 Overview of Data Collection Methods

The ECLS-K fifth-grade data collection was conducted in the fall and spring of the 2003–2004 school year. Fall data collection included contacting sampled schools to set appointments to conduct the child assessments in the spring of the school year, verifying the parent consent procedures, linking children to teachers, identifying children who had withdrawn from the school, and obtaining location information about their new schools. Spring data collection instruments included the direct child assessments, parent interviews, teacher and school questionnaires, student record abstract, and facilities checklist. The activities to locate children and gain cooperation of the schools into which they had transferred began in fall data collection and continued in spring data collection. The content and timeline of the fifth-grade data collections are shown in exhibit 5-1.

The mode of data collection was computer-assisted personal interviewing (CAPI) for the child assessments; telephone and in-person computer-assisted interviewing (CAI) was used to conduct the parent interview; self-administered questionnaires were used to gather information from teachers, school administrators, and student records. The facilities checklist was completed by field staff.

Exhibit 5-1. Timeline of fifth-grade data collection

Fifth grade		
2003	2004	
Fall	Winter	Spring
Advance school contact		Child assessments
Tracing sampled households		Parent interviews conducted
		Teacher information collected
		School and school administrator data, facilities checklist
		Student record data
		Tracing children who move
		Activities to gain transfer school cooperation

SOURCE: U.S. Department of Education, National Center for Education Statistics, Early Childhood Longitudinal Study, Kindergarten Class of 1998–99 (ECLS-K), spring 2004.

5.2 Field Staff Training

Several in-person training sessions were conducted to prepare staff for the fifth-grade data collection. In the fall of 2003, supervisors were trained to contact original schools and recruit transfer schools. In the spring of 2004, two trainings were held: one for field supervisors and one for assessors. Field supervisors managed all the data collection activities within their assigned work areas, supervising the assessors and conducting child assessments and parent interviews. Assessors conducted the child assessments and parent interviews. Twenty interviewers were assigned to complete only parent interviews during spring data collection. The following sections discuss each specific type of training.

5.2.1 Advance Contact and Recruitment Training

Field supervisors were trained for 3 days in September 2003 to contact original sampled schools and transfer schools to set up the data collection in the spring. A total of 39 field supervisors and 2 field managers completed training. Topics included an overview of study activities to date, verifying parent consent procedures, identifying and locating children who had moved from the schools they attended in the third grade, identifying the teachers of ECLS-K children and linking them to those children, and exercises on scheduling schools efficiently within a work area.

As in the third-grade training, advance contact and recruitment training was conducted using the automated Field Management System (FMS). The FMS was used throughout the data collection period to enter information about the sampled children, parents, teachers, and schools and to monitor production on all data collection activities. The field supervisors entered information into the FMS during training presentations, thus acquiring hands-on experience with the FMS and all field procedures prior to beginning data collection. The field supervisors completed role plays and exercises that involved entering information into the FMS.

5.2.2 Spring-Fifth Grade Training

Field supervisors, interviewers, and assessors were trained for the spring-fifth grade data collection in one session in February 2004. Prior to the February in-person training session, supervisors and assessors completed 8 hours of home study training on the study design, field procedures, and computer keyboard skills.

Field Supervisor Training. Field supervisor training preceded the assessor training and lasted for three days. The topics covered in the field supervisor training session included reviewing materials from the fall school recruitment, role plays to practice contacting school coordinators, identifying and locating children who had moved from their third-grade schools, identifying the regular and special education teachers of ECLS-K children and linking them to those children, distributing and following up on teacher questionnaires and school administrator questionnaires, completing the facilities checklist, and conducting quality control observations. Field supervisors were also trained to use the FMS, and the field supervisors entered information into the FMS during training presentations. Eighty-one (81) field supervisors completed training.

Assessor Training. The assessor training sessions included an overview of study activities to date, interactive lectures based on the direct child assessments and the parent interview, practice parent interviews in pairs using role-play scripts, practice direct child assessments using role-play scripts, direct child assessment precertification exercises on each form of the direct child assessments, techniques for parent refusal avoidance, and strategies for building rapport with children. A major goal of the assessor training was to train field staff in the proper procedures for conducting the direct child assessments. This included following standardized procedures for administration of all assessment items and for giving children neutral praise. The sessions provided trainees with hands-on experience with all the direct child assessment materials and procedures and the CAI programs prior to data collection. Interactive lectures and role plays were also used to train field staff in administering the parent interviews. Trainees practiced entering information into the CAI system on laptop computers during training presentations on conducting the direct child assessments and parent interview. Assessor training lasted for five days. Field supervisors were also trained to perform all assessor activities. Two hundred sixty-two (262) assessors and 81 field supervisors completed training. (Twenty trainees were assigned to complete only parent interviews during the spring data collection. They attended the first day-and-a-half of training.)

5.2.2.1 Certification of the Child Assessors

In order to ensure that the supervisors and assessors who completed training administered the direct child assessments in a standardized manner, 323 field staff assigned to conduct child assessments completed certification exercises. Certification was composed of written exercises on each level form of each of the assessment domains (e.g., the red form of reading which corresponds to a low difficulty level) and an observation of each trainee administering the assessment to children specifically recruited for the training sessions.

Written Certification Exercises. Each level form of an assessment domain was reviewed in detail during an interactive lecture. This was followed by independent review and individual practice in administering the assessment domain. After the individual practice, written exercises were distributed.

The written exercises were used to ensure that each trainee understood the coding rules for selected open-ended questions with particularly complex scoring rubrics. Each exercise included certain assessment items from the level form that was just discussed, with an assortment of possible responses. The trainees were instructed to score each response as either correct or incorrect. The exercises were then scored by the co-trainer during the next training session. Trainees who did not achieve a passing score were asked to attend a training session in the evening to review the items. These trainees then re-took the same exercises that they had previously failed to pass.

Most trainees passed the written exercises on the first attempt. All of the trainees who had to re-take the exercises after the remedial evening session achieved a passing score. Less than a quarter of the trainees (77 trainees or 24 percent) did not pass at least one element of the reading certification exercises on the first attempt. The mathematics and science certification exercises were considerably easier for trainees; only two trainees were required to repeat any mathematics exercise and 55 trainees (17 percent) were required to repeat the science yellow certification exercise. This variability was due to the complexity of the grade 5 reading scoring rubrics and the unfamiliarity of the exercises themselves (reading exercises were distributed first, with mathematics and science exercises on later days). Once additional training was given, all of the trainees passed the exercises on the second attempt. Refer to the *ECLS-K Fifth-Grade Methodology Report* (NCES 2006–037) (Tourangeau, Lê, and Nord 2005) for additional detail.

Assessment Certification. In the final stage of the certification process, the trainees were observed conducting a direct child assessment with children brought on site to the training session. Training staff who were already certified on the assessment observed trainees as they administered parts (e.g., routing test and a level test) of the assessment to fifth grade-aged children. They rated the trainees on skills such as rapport with the child, avoidance of coaching or use of inappropriate probing, following proper administration procedures, and pacing. While the trainee administered the assessment, an observer certified on the assessment simultaneously coded the child's answers to preselected open-ended questions. After the assessment was completed, the observer brought up a screen in the CAPI program that displayed the assessor's coding of the open-ended questions. The answers recorded by the assessor were compared with those recorded by the observer. Discrepancies in any of the recorded answers were included in the assessor's overall score on a certification form.

Table 5-1 presents the results of the training certification. There were 242 assessors and 81 field supervisors for a total of 323 trainees who were certified; 20 assessors were only trained to complete parent interviews. Trainees who scored 85 percent or above were certified qualified to administer the child assessments. Trainees who scored between 70 and 84 percent were required to complete remedial training and an additional certification in the field before beginning assessments.

Table 5-1. Results of training certification, fifth grade: School year 2003–04

Trainees	Number	Percent
Total	323	100.0
Score on certification form		
85 percent or above	320	99.1
70–84 percent	3	0.9
Below 70 percent	0	0.0

SOURCE: U.S. Department of Education, National Center for Education Statistics, Early Childhood Longitudinal Study, Kindergarten Class of 1998–99 (ECLS-K), spring 2004.

The majority of the trainees (99 percent) scored above 85 percent on the certification form. Only 1 percent scored between 70 and 84 percent. None of the trainees failed to meet the 70 percent threshold on the assessment certification form. All trainees who needed remedial training were certified to administer the child assessments, after they conducted a second assessment on a fifth grade-aged child who was not part of the ECLS-K sample.

5.3 Fall Preassessment School Contact

Beginning in September 2003, all participating ECLS-K schools (i.e., schools that had participated in third grade), were contacted by telephone to prepare for the spring data collection. When children were identified as having transferred to another school, the child's new school (and district, if necessary) was recruited.

5.3.1 Advance Mailings

In September 2003, an advance package was mailed via Federal Express to all participating ECLS-K schools asking them to prepare for the fall preassessment telephone call. The schools were asked to identify a school staff coordinator to serve as a liaison with the study. (In returning schools, this person was usually the coordinator from previous rounds of data collection.) The advance package contained study findings from first grade and an overview of fifth-grade data collection activities. The school coordinators were asked to complete an information form about the ECLS-K sampled children prior to the telephone call.

5.3.2 Preassessment Contact

The fall preassessment contact was made by telephone between September and December 2003. The fall preassessment school contact was successful in meeting two important goals: (1) contacting original sampled schools to set up the spring assessment and (2) identifying children who had withdrawn from their spring-third grade school and had entered their fifth-grade transfer school. Schools were determined to be ineligible for fifth-grade data collection if no sampled children were currently enrolled. Original sampled schools became ineligible if fourth grade was the highest grade in the school or if the school had closed, that is, was no longer operational. More transfer schools were determined to be ineligible as children transferred out of them into other schools. During the preassessment contact, the field supervisor contacted the school coordinator to schedule the dates of the assessment visits, identify ECLS-K sampled children who were no longer enrolled at the school, collect locating information for those children, identify each enrolled child's reading and mathematics or science teachers and special education teacher, review parental consent status, obtain information on special accommodations[1] during assessment for the enrolled sampled children, and answer any questions that the school coordinator might have.

Identifying ECLS-K Sampled Children Who Withdrew from the School. Field supervisors asked the school coordinators to identify ECLS-K children who had transferred out of the school. If the school records indicated where the children had transferred, then the field supervisors asked the school coordinator to provide the names, addresses, and telephone numbers of these transfer schools. Of those children who had transferred, only a subset was followed to their new school. (See section 4.3.1 in chapter 4 for more detail on how mover children were subsampled.) If the new school belonged to a district that was new to the study, the district was contacted and recruited before any contact was made with the school. If the district was already cooperating, the new school was contacted and recruited directly.

Reviewing Information about ECLS-K Sampled Children. Field supervisors collected information from the school coordinators about the ECLS-K sample children still enrolled in the school, including the child's current grade, the name and classroom for the child's reading teacher, mathematics or science teacher and whether or not the child had an Individualized Education Plan (IEP). If the child had an IEP, then the name and classroom of the child's special education teacher was collected, along

[1] Accommodations included in the data collection protocol were special setting accommodations, scheduling/timing accommodations, presence of a health care aide, or use of an assistive device.

with whether the child required any accommodations to participate in the direct cognitive assessment. The accommodations to the fifth-grade direct cognitive assessment were the same as those for the kindergarten, first-grade and third-grade, direct cognitive assessments. Field supervisors contacted the teachers of the ECLS-K children as necessary for any of this information.

Reviewing Parent Consent. Because parental consent was obtained in the base year and obtained again in the third-grade year, field supervisors did not raise parental consent issues with the school coordinator unless the school district required it. If the school was a transfer school, then field supervisors asked the school coordinator whether parental consent was required. If the schools required consent to be obtained again or changed the type of consent that was required (e.g., from implicit to explicit), parent letters and consent forms were either mailed to the school for distribution to parents or directly to parents from Westat, based on the schools' preference. Parents were requested to return signed consent forms to the school coordinator.

Contacting Families of Homeschooled Children. As part of the fall preassessment contact, children who were homeschooled in previous rounds were identified. The status of homeschooled children who were identified in rounds 1 through 5 was verified with their parents and updated as necessary. In addition, some homeschooled children were identified by the schools during the fall preassessment contact. Their status was also verified with their parents during data collection. Parents of these children were contacted from September through December 2003 to determine if the child was still homeschooled or had enrolled in a school. If the child had enrolled in a school, the new school was contacted and recruited into the study. Parents of children who were still schooled at home were notified about the next round of data collection in the spring.

Identifying the Key Child in Classrooms with Multiple Study Children. In grade 5, the design of the child-level teacher questionnaire was changed to include collecting data about the child's reading class and mathematics or science class. In previous rounds, children had been taught primarily in intact classrooms and teachers only reported classroom level information once for the classroom. Due to the design change, the teacher-child links were broadened to include the domain (reading, mathematics, or science) as well as information to identify the reading, mathematics, or science classroom. In order to reduce data collection burden for teachers who were linked to multiple sample children in the same class, a "Key Domain Child" was identified for each separate subject and class that each teacher taught. The teachers would be asked to report classroom level information only once in the questionnaire for the key domain child and child-level information for all sampled children in their class. Field supervisors collected the teacher-child-domain-classroom link information about each child and entered the information into the FMS. The information was used to generate the hardcopy teacher questionnaires (see section 5.5.1 for more information). Refer to the *ECLS-K Fifth-Grade Methodology Report* (NCES 2006–037) (Tourangeau, Lê, and Nord 2005) for additional detail on the Key Child concept.

5.4 Tracing Activities during the Fifth-Grade Data Collection

In order to ensure that as many of the sampled children as possible were contacted for fifth-grade data collection, locating efforts were undertaken in the summer of 2003. In June 2003, the entire household database was submitted to search vendors to obtain a current address and telephone number. Between June and August 2003, staff in Westat's Telephone Research Center (TRC) traced children who could not be located during previous rounds of data collection. TRC staff also used the Internet, telephone directories, and other means to locate these children and their households. When children and/or households were found, the new school and contacting information was entered into the computer database, for fielding in the spring. Table 5-2 presents the results of this effort. See section 5.5.4 for more details about children who transferred schools in fifth grade.

Table 5-2. Results of the Telephone Research Center's locating efforts, fifth-grade data collection: School year 2003–04

Result	Number	Percent
Total cases worked	829	100.0
Located and entered into database	305	36.8
Unlocatable	519	62.6
Final refusal	5	0.6

NOTE: "Unlocatable' means that the children and their households could not be found using the available tracing and locating strategies, "final refusal" means that the child's family indicated that they did not want to participate.
SOURCE: U.S. Department of Education, National Center for Education Statistics, Early Childhood Longitudinal Study, Kindergarten Class of 1998–99 (ECLS-K), spring 2004.

In mid-September 2003, all eligible households were mailed a letter asking the sampled child's parent/guardian to record their current address and the child's current school information on an enclosed postcard. The TRC began calling households that did not return a postcard in mid-October to obtain current information before spring data collection. By the end of December, approximately 75 percent of the households had responded.

5.5 Spring-Fifth Grade Data Collection

All children who were assessed during the base year or for whom a parent interview was completed in the base year were eligible to be assessed in the spring-fifth grade data collection, with four exceptions: They are (1) children who became ineligible in an earlier round (because they died or moved out of the country), (2) children who were subsampled out in previous rounds because they moved out of the original schools and were not subsampled to be followed, (3) children whose parents emphatically refused to cooperate (hard refusals) in any of the data collection rounds since spring-kindergarten, and (4) children in the third-grade sample for whom there are neither first-grade nor third-grade data. Eligibility for the study was not dependent on the child's current grade, that is, children were eligible whether they had been promoted to fifth grade or had been retained in fourth grade.

As in previous rounds of data collection, the field staff were organized into work areas, each with a data collection team consisting of one field supervisor and two or more assessors. The data collection teams were responsible for all data collection activities in their work areas; they conducted the direct child assessments and the parent interviews, collected all school and teacher questionnaire and completed checklists. The majority of field staff members in fifth grade were continuing from previous rounds of data collection; a few new staff were hired in areas where no experienced ECLS-K staff resided.

5.5.1 Preassessment School Contact

Based on the information collected in the fall of 2003, packets of hard-copy teacher and school administrator questionnaires and instructions were assembled and mailed to schools beginning in January 2004, along with letters confirming the scheduled visits to the school. Teachers and school administrators were asked to complete the questionnaires for pickup on assessment day. In February 2004, letters were also mailed to parents reminding them of the spring-fifth grade data collection activities.

Field supervisors conducted most preassessment activities by telephone starting in February 2004. The preassessment activities for these schools were similar to those conducted in previous rounds of data collection and included confirming the assessment date and receipt of the hard-copy questionnaires and arranging for space to conduct the assessments.

5.5.2 Conducting the Direct Child Assessments

The direct child assessments were conducted from February through June 2004, the same time of year as in prior spring data collections. Over three-quarters of the child assessments were completed in April, 21 percent were completed in May and 1 percent were completed in June. In year-round schools, assessment teams made multiple visits to the school, visiting when each track was in session, to assess the sampled children.

The direct child assessments were usually conducted in a school classroom or library. Before conducting the assessments, field supervisors and assessors set up the room for the assessments. They followed procedures for meeting the children that had been agreed upon during the preassessment contact with the school. Each child was signed out of his or her classroom prior to the assessments and signed back into the classroom upon the conclusion of the assessments. In scheduling schools in the fall, attempts were made to schedule the direct child assessments at about the same point in time between the beginning and the end of the school year, to increase the likelihood that exposure to instruction would be about the same for all children. The fifth-grade direct child assessments averaged 97 minutes.

Table 5-3 displays the number of completed child assessments for each round of data collection, including spring-fifth grade. All of the assessments were completed in English. The majority (83 percent) of the assessments were completed in original schools. About one-sixth of the assessments (17 percent) were completed in transfer schools.

Accommodations and Exclusions. Less than 1 percent of participating children in fifth grade required accommodations or were excluded from the direct child assessments. Children were excluded from the direct assessments if they had a disability (e.g., blindness or deafness), that could not be accommodated by the ECLS-K direct assessments, or if their Individualized Education Plan prevented their participation in assessments or required an accommodation not offered by the ECLS-K assessments. Accommodations offered by the ECLS-K assessments were as follows: alternative setting, scheduling, or timing; health care aide present; or the use of a personal assistive device. Table 5-4 presents the number of children excluded from or requiring an accommodation to the direct child assessment procedures in the spring of fifth grade.

Table 5-3. Completed child assessments by round of data collection: School years 1998–99, 1999–2000, 2001–02, and 2003–04

	Fall-kindergarten		Spring-kindergarten		Fall-first grade		Spring-first grade		Spring-third grade		Spring-fifth grade	
Characteristic	Number	Percent	Number	Percent	Number	Percent	Number	Percent	Number	Percent	Number	Percent
Child assessments completed	19,147	100.0	19,987	100.0	5,297	100.0	16,622	100.0	14,502	100.0	11,368	100.0
In English, no accommodation[1]	17,019	88.9	18,342	91.8	4,848	91.5	15,460	93.0	13,565	93.5	10,813	95.1
In Spanish	1,008	5.3	724	3.6	176	3.3	286	1.7	†	†	†	†
In other language	410	2.1	229	1.1	33	0.6	37	0.2	†	†	†	†
With accommodation[2]	515	2.7	579	2.9	195	3.7	761	4.6	814	5.6	465	4.1
Excluded	88	0.5	70	0.4	28	0.5	47	0.3	74	0.5	62	0.5
Partial complete	107	0.6	43	0.2	17	0.3	31	0.2	49	0.3	28	0.2
Original sampled school	19,147	100.0	19,463	97.4	4,867	91.9	14,830	89.2	10,820	74.6	9,439	83.0
Transfer school	0	0.0	524	2.6	430	8.1	1,792	10.8	3,682	25.5	1,929	17.0

†Not applicable.

[2] The term *accommodation* in this table is the field operational definition of accommodation, which includes the wearing of glasses and hearing aids. These types of aids were systematically tracked to ensure that every child had the same chance at a successful assessment. With this information, assessors could prompt a child (e.g., to get her glasses before being assessed).

NOTE: This table reflects final production numbers prior to statistical adjustment. This table does not include children who were subsampled out in fall- and spring-first grade and spring-third grade (see section 5.5.4.) These numbers should not be used to estimate student mobility.

SOURCE: U.S. Department of Education, National Center for Education Statistics, Early Childhood Longitudinal Study, Kindergarten Class of 1998–99 (ECLS-K), fall 1998, spring 1999, fall 1999, spring 2000, spring 2002, and spring 2004.

Table 5-4. Number of children excluded from or accommodated in the spring-fifth grade assessments: School year 2003–04

Category	Number of children
Exclusions	
Excluded for disability	63
Accommodation[1]	
Alternative setting accommodation	50
Scheduling/timing accommodation	64
Health care aide present	12
Personal assistive device	9

[1] The term *accommodation* in this table includes only those accommodations offered during the assessment such as an alternative setting.
SOURCE: U.S. Department of Education, National Center for Education Statistics, Early Childhood Longitudinal Study, Kindergarten Class of 1998–99 (ECLS-K), spring 2004.

5.5.3 Conducting the Parent Interview

Parent interview procedures mirrored those of previous rounds of data collection. The parent interview was administered, primarily by telephone interview using CAI, between February and June 2004. Slightly over 50 percent of the parent interviews were completed in February and March, 43 percent were completed in April and May, and 6 percent were completed in June. The parent interview averaged 43 minutes. As in previous rounds of data collection, the parent interview was conducted in person if the respondent did not have a telephone. Table 5-5 contains the number of parent interviews per round, including spring-fifth grade. In fifth grade, only 2.7 percent of all completed parent interviews were conducted in person; 8.1 percent of all completed parent interviews were conducted in a language other than English; 95.1 percent of the latter were conducted in Spanish.

Table 5-5. Number and percent of completed parent interviews by data collection mode, language, and wave of data collection: School years 1998–99, 1999–2000, 2001–02, and 2003–04

	Fall-kindergarten		Spring-kindergarten		Fall-first grade		Spring-first grade		Spring-third grade		Spring-fifth grade	
	Number	Percent	Number	Percent	Number	Percent	Number	Percent	Number	Percent	Number	Percent
Parent interviews completed	17,997	100.0	18,907	100.0	5,073	100.0	15,576	100.0	13,504	100.0	10,940	100.0
In person	618	3.4	619	3.3	211	4.2	456	2.9	319	2.4	295	2.7
By phone	17,379	196.6	18,288	96.7	4,862	95.8	15,120	97.1	13,185	97.6	10,645	97.3
Language of parent interviews												
English	17,379	96.6	17,482	92.5	4,717	93.0	14,319	91.9	12,416	91.9	9,444	90.9
Spanish	618	3.4	1,321	7.0	351	6.9	1,071	6.9	932	6.9	846	7.7
Other language	0	0.0	81	0.4	0	0.0	75	0.5	41	0.3	39	0.4
Partial complete	0	0.0	23	0.1	5	0.1	111	0.7	115	0.9	111	1.0

NOTE: This table completes final production numbers prior to statistical adjustment.
SOURCE: U.S. Department of Education, National Center for Education Statistics, Early Childhood Longitudinal Study, Kindergarten Class of 1998–99 (ECLS-K), fall 1998, spring 1999, fall 1999, spring 2000, spring 2002, and spring 2004.

5.5.4 Collecting Data for Children Who Had Withdrawn From Their Previous Round School

While contacting schools, field supervisors asked school coordinators to identify children who had withdrawn from the school since the spring of third grade. School staff were asked whether they knew the name and address of the school to which the child transferred, as well as any new information about the child's household address. For the children who had moved from their spring-third grade school and were not part of the sample to be followed, information was collected only from the school personnel and not parents. For children who had withdrawn from their spring-third grade school and were identified to be followed (i.e., were part of the sample of movers), supervisors also consulted parents and other contacts for information on the children's new school. This information was entered into the FMS and processed at Westat for data collection.

Table 5-6 presents the status of the children who were identified as movers in fifth grade; 12,717 children were identified as having transferred from the school in which they were enrolled during the spring of base year, first grade, or third grade. Of the 12,717 mover children in spring-fifth grade, 4,187 (32.9 percent) were in scope (i.e., children selected to be followed) and followed. The remaining 8,530 mover children were out-of-scope and were not followed; no child assessments or parent interviews were conducted for these children.

Parent interviews were attempted for all in-scope children. However, different school and assessment data collection strategies were followed for children who had moved, depending on where they had moved to and the status of their new school. School and assessment data collection was attempted for children who had moved and were flagged as "follow" in spring-fifth grade in the following ways:

- Data collected for children moving into cooperating base year sampled schools included the child assessments in the school, school administrator questionnaire, regular and/or special education teacher questionnaires, facilities checklist, and student record abstract forms; and

- Data collected for children moving into nonsampled schools in base year cooperating districts included the child assessments in the school, school administrator questionnaires, regular and/or special education teacher questionnaires, and student record abstract forms, if school permission was obtained. If school permission was not obtained, the assessments were conducted in the home and no school or teacher data were collected.

Table 5-6. Number of children who moved in spring-fifth grade by completion category: School year 2003–04

Category	Spring-fifth grade Number of children	Percent
Total movers[1]	12,717	100.0
Out-of-scope[2]	8,530	67.1
Did not follow[3]	7,880	92.4
Moved to outside of U.S.[3]	153	1.8
Deceased[3]	7	0.1
Excluded from spring-fifth grade[4]	490	5.7
In-scope and followed[2]	4,187	32.9
Completed assessment[5]	3,299	78.8
Unlocatable[5]	281	6.7
Nonsampled primary sample unit[5]	395	9.4
Assessment refused[5]	149	3.6
Not assessed/absent[5]	63	1.5

[1] The movers described in this table are defined as "operations movers" rather than "statistical movers" since cooperation must be secured from the transfer schools in order for data collection to proceed.
[2] Percent based on total movers.
[3] Percent based on out-of-scope children.
[4] In fifth grade, four groups of children were excluded, irrespective of other subsampling procedures that were implemented. They were (1) children who had become ineligible in an earlier round (because they had died or moved out of the country); (2) children who were subsampled out in previous rounds because they had moved out of the original schools and were not subsampled to be followed; (3) children whose parents emphatically refused to cooperate (hard refusals) in any of the data collections rounds since spring-kindergarten ; and (4) children in the third-grade sample for whom there were neither first-grade nor third-grade data.
[5] Percent based on in-scope children.
SOURCE: U.S. Department of Education, National Center for Education Statistics, Early Childhood Longitudinal Study, Kindergarten Class of 1998–99 (ECLS-K), spring 2004.

- For children moving into transfer schools that refused, schools in sampled districts that refused, or originally sampled schools that were ineligible when sampled because they did not have kindergarten classes, the direct child assessments were conducted in the home. No school or teacher data were collected;

- For children moving into schools in nonsampled districts or dioceses:

  - If the school was within the primary sampling unit (PSU), data collected included the child assessments in the school, school administrator questionnaire, regular and/or special education teacher questionnaires, facilities checklist, and student record abstract forms, if school permission was obtained. If school permission was not obtained, the assessments were conducted in the home and no school or teacher data were collected; and

  - If the school was outside the PSU, no child, school, or teacher data were collected.

- For children who were not enrolled in school in the spring (including children who were homeschooled), data collected included the child assessments in the home if the child was in the sampled PSU. If the child was outside the sampled PSU, no child assessment or school or teacher data were collected.

Of the children who moved in fifth grade and were selected to be followed, 9.4 percent moved into a school outside the PSU and 6.7 percent of the movers could not be located. Assessments were completed for 78.8 percent of the movers who were followed in the spring-fifth grade data collection.

5.5.5 Teacher and School Data Collection

Data were collected from school administrators, regular classroom teachers, and special education teachers from February through June 2004.

The school and teacher questionnaires were mailed to the school coordinators beginning in January 2004. Using the teacher-child-domain-classroom linkage information collected in the fall, a packet of questionnaires was assembled for each reading, mathematics, science, and special education teacher. The customized teacher questionnaire materials included: a cover letter and a twenty-dollar check attached to the teacher questionnaire; instruction sheets attached to the child-level questionnaires for each separate class; and, a special education instruction sheet attached to the special education questionnaires (if appropriate). Packets were bundled together by school and mailed to the school coordinator for distribution. If the school or teacher and school administrator were not identified in the fall preassessment contact, then the supervisor gathered the relevant information during the spring preassessment call and mailed the packets.

Teachers were asked to complete child-level instruments for the sampled children in their classrooms, and they were reimbursed $7 for each child they rated in reading and mathematics or science. In addition, school staff were asked to complete a student record abstract after the school year closed and were reimbursed $7 for every student record abstract completed. Field supervisors also completed a facilities checklist for each sampled school.

During the field period, field supervisors followed up with school administrators and teachers in visits to the schools to conduct assessments and by telephone to collect completed questionnaires, ensuring that questionnaires were not missing critical information and that completed questionnaires were mailed to Westat.

5.6 Data Collection Quality Control

Continuous quality assurance procedures were employed during all data collection activities, but with a particular focus on the assessments. The procedures were incorporated throughout all stages of the study (e.g., during instrument development, in the staff training program, through assessment certification, and as part of the ongoing staff observations and evaluation activities).

Data collection quality control efforts began with the additional development and testing of redesigned sections of the CAI/CAPI applications and the FMS. As sections of these applications were re-programmed, extensive testing of the entire system was conducted to verify that the systems were working properly from all perspectives. This testing included review by project design staff, statistical staff, and the programmers themselves. Quality control processes continued with the development of field procedures that maximized cooperation and thereby reduced the potential for nonresponse bias.

Quality control activities continued during training and data collection. During assessor training, field staff practiced conducting the parent interview in pairs and administered the direct child assessments with fifth grade-aged children brought to the training site for this purpose. The supervisors and assessors were certified on the child assessments using the Training Certification Form. When the fieldwork began, field supervisors observed each assessor conducting child assessments and made telephone calls to parents to validate the interview. Field managers made telephone calls to the schools to collect information on the school activities for validation purposes.

5.6.1 Child Assessments Observations

Field supervisors conducted on-site observations of the child assessments and completed the child observation form. The quality control plan specified two observations for each of 242 assessors. (Assessors completing only parent interviews were not observed.) The first observation was scheduled to be conducted by the end of March, and the second observation was scheduled to be conducted by the end of April. These procedures were followed for the majority of assessors (97 percent for first observations and 100 percent for second observations), but some assessors were observed only once because they completed their assignments in April before the second observation could be scheduled.

A standardized observation form was used to evaluate the assessor's performance in conducting the child assessments. The assessor was rated in three areas:

1. Rapport building and working with the child—use of neutral praise and the assessor's response to various child behaviors;

2. Cognitive assessment activities—reading questions verbatim, the use of acceptable probes, the use of appropriate hand motions, and the absence of coaching; and

3. Specific assessment activities—correctly coding answers to open-ended questions in the assessments and following administration procedures.

The field supervisors recorded their observations on the form and then reviewed the form with the assessor. The most frequent problems observed were not reading the items verbatim and inappropriate gesturing. Feedback was provided to the assessors on the strengths and weaknesses of their performance and, when necessary, remedial training was provided in areas of weakness. All but one assessor scored at the 85 percent or above level. Table 5-7 presents the result of the observations.

Table 5-7. Results of the child assessments observations, fifth-grade data collection: School year 2003-04

Number of observations[1]	Score on certification form
Total: 466	
465	85 percent or above
1	70–84 percent
0	Below 70 percent

[1] Two hundred and forty-two assessors were to be observed coming out of training; assessors completing only parent interviews were not observed. 239 assessors had initial observations; 227 assessors had second observations. Only 1 assessor failed to pass the observation and was released from the project.
SOURCE: U.S. Department of Education, National Center for Education Statistics, Early Childhood Longitudinal Study, Kindergarten Class of 1998–99 (ECLS-K), spring 2004.

5.6.2 Validation of Parent Interviews

Approximately 10 percent of the respondents who completed parent interviews were selected for a short re-interview conducted by a field supervisor (i.e., a "validation" interview). The first parent interview completed by an assessor was always selected for validation. Over the course of the field period, a running count of an assessor's completed parent interviews was maintained, and each tenth completed parent interview was selected for validation, thus ensuring that 10 percent of each assessor's cases were selected for validation. The parent validation was approximately 5 minutes long and was conducted by telephone. In spring-fifth grade, a total of 1,028 parent interviews were validated with 94 percent reporting the same answers as the original interview. Refer to the *ECLS-K Fifth-Grade Methodology Report* (NCES 2006–037) (Tourangeau, Lê, and Nord 2005) for additional detail.

Field supervisors used a standardized parent validation script to make validation calls to parents. The script covered the following topics:

- Verification of the child's name, date of birth, and sex; and
- Seven questions repeated from the parent interview.

5.6.3 Validation of School Visits

To ensure that assessments proceeded smoothly, a validation call was completed with the school principal in at least two of each supervisor's assigned schools in the spring-fifth grade data collection.

Field managers conducted the school validations by telephone. The first school completed by each team was called to ascertain how well the preassessment and assessment activities went. If the feedback from the school was positive, the fifth school that each team completed was called. If any problems were indicated in the first validation call, immediate action was taken with the field supervisor. The validation feedback was discussed with the supervisor and remedial action was taken, including in-person observation of the supervisor's next school, if necessary. In spring-fifth grade, a total of 162 school visits were validated with no negative reports of the assessment team or study from school staff; all schools reported that the experience was "Very Satisfactory" or "Satisfactory." Refer to the *ECLS-K Fifth-Grade Methodology Report* (NCES 2006–037) (Tourangeau, Lê, and Nord 2005) for additional detail.

Field managers used a standardized script to call the school principals. The script covered the following topics:

- An overall rating of how the assessments went;
- Feedback about the study from the children and teachers;
- Suggestions for improving procedures and making it easier for a school to participate; and
- General comments and suggestions.

5.6.4 Assessor Interrater Reliability

As part of the child assessments observation described in section 5.6.1, field supervisors completed an assessment certification form for each observation they conducted. An important element of this form was the "validation items." With the exception of the reading routing test, all of the assessments included at least one item that both the observer and the assessor scored. The items that were scored by both the assessor and observer had open-ended responses that called for interpretation on the part of the assessor to determine whether a child's response was correct; the reading routing test did not have any items of comparable complexity. By comparing the extent to which assessors and observers agreed on scoring these validation items, a measure of interrater reliability was obtained. Interrater reliability provided a measure of the accuracy of the assessor's scoring compared with the standard, the observer's.

Table 5-8 contains the results of these comparisons. As can be seen, overall interrater reliability was very high throughout all the forms. It was highest for mathematics (98 percent or better depending upon the form) and lowest for reading, with the reading yellow level (the medium reading level) showing the lowest percent agreement (95.7 percent). The reading yellow level path received a relatively large number of observations (232) and also contained a relatively large number of validation items (5) compared with some of the other paths. Thus, there was greater opportunity for disagreement on this path compared with the others. The science blue level (the high science level) also had a relatively higher opportunity for disagreement (212 observations and 4 validation codes) and it, too, exhibited a somewhat lower interrater reliability (96.7 percent) compared with some of the other paths. The reliability, however, even on these more difficult paths, was high and demonstrated that the assessors accurately coded open-ended items.

Table 5-8. Interrater reliability on child assessment validation items: School year 2003–04

Category	Number of Observations	Number of Validation items	Percent agreement: Assessors and observers[1]
Reading	458	13	96.5
Routing	†	0	†
Red	103	3	97.1
Yellow	232	5	95.7
Blue	123	5	97.4
Mathematics	461	10	99.0
Routing	461	2	99.7
Red	162	1	100.0
Yellow	161	3	98.8
Blue	138	4	98.4
Science	463	14	96.8
Routing	460	3	99.1
Red	53	4	98.1
Yellow	198	3	96.6
Blue	212	4	96.7

† Not applicable

1 Percent agreement was calculated as follows: number of validation items observed in which observer agreed with the assessor divided by number of validation items observed.

SOURCE: U.S. Department of Education, National Center for Education Statistics, Early Childhood Longitudinal Study, Kindergarten Class of 1998–99 (ECLS-K), spring 2004.

5.7 Spring-Fifth Grade Completion Rates

Since data were collected from schools, parents, teachers, and children, there were many opportunities for sources to contribute differentially to nonresponse, and this is reflected in the varying completion rates in the tables in this section. These completion rates differ not only by survey instruments, but within each survey instrument they differ also by school and child characteristics.

In this section, fifth-grade completion rates are presented for three groups of children: (1) children sampled in kindergarten, (2) children sampled in first grade through the freshening procedure, and (3) both groups combined. Completion rates for the fifth-grade data collection were computed with the same procedures used for spring-first grade and spring-third grade to allow for comparisons of completion rates for the three years of data collection following the base year. For spring-first grade and spring-third grade, the sample of children is the same: base year respondents (i.e., children who had either a fall- or spring-kindergarten child assessment or parent interview) and children sampled in spring-first grade as part of sample freshening as described in section 4.3.2. For spring-fifth grade, the sample of children was reduced to exclude base year respondents who belonged in the following special groups as described in section 4.5: (1) children who became ineligible in an earlier round (because they died or moved out of the country), (2) children who were subsampled out in previous rounds because they moved out of the original schools and were not subsampled to be followed, (3) children whose parents emphatically refused to cooperate (hard refusals) in any of the data collection rounds since spring-kindergarten, and (4) children eligible for the third-grade sample for whom there are neither first-grade nor third-grade data. Among the 21,357 children who were eligible for the study after the base year, 16,143 were part of the fifth-grade data collection. Weighted completion rates were computed using the base weight (i.e., inverse of selection probabilities) adjusted for movers, but not adjusted for nonresponse.

5.7.1 Children Sampled in Kindergarten

Tables 5-9 to 5-12 present weighted and unweighted child-level completion rates for spring-fifth grade data collection, broken out by school characteristics.[2] These rates pertain to children who were sampled as part of the kindergarten cohort in the base year. For the ECLS-K, a completion rate is a response rate conditioned on the results of an earlier stage of data collection. For the group of children sampled in kindergarten, all completion rates are conditioned on the case having been a base year respondent and retained in the fifth-grade data collection.

In general, completion rates for fifth grade are higher than in third grade. This is due to the exclusion of hard-to-field cases from the fifth-grade collection. Hard-to-field cases are the hard-refusal cases and cases that were nonrespondents in both first and third grades as described in section 4.5. If these cases had not been excluded from the fifth grade, they would most likely be nonrespondents and would bring down the completion rates.

Table 5-9 shows that the completion rates for the child assessment are quite high and uniform across school characteristics. Excluding the “unknown” category, the rates vary from 93.1 percent in non-Catholic private schools to 99.7 percent in schools in large towns. Similarly, the completion rates for the parent interviews were uniform across school characteristics ranging from 87.2 percent for children in schools with 750 or more students and in schools where 50 to 89 percent of level

[2] The categories of school affiliation in the tables in this chapter do not match categories of school affiliation in the tables in chapter 4. This is to allow users to compare completion rates in fifth grade with those in previous years.

children belong to the minority groups, to 94.3 percent for children in small towns (excluding the “unknown” category). The “unknown” category includes children who could not be located and those children who had moved into a nonsampled county. The category “unknown” also includes 35 children who were homeschooled and thus had no information concerning schools.

The “unknown” category aside, both the child assessment and the parent interview completion rates increased between third grade and fifth grade for all school characteristics. The completion rates by mover status are discussed later, but the rates of completing all the instruments are much lower for children who moved than for those who did not move.

Table 5-10 shows that the overall weighted completion rate is 77.1 percent for the school administrator questionnaire and 78.8 percent for the facilities checklist. The rate for school administrator questionnaires is 11 percentage points higher than the corresponding rate in third grade. The rate for facilities checklist is only about 2 percent higher. The completion rates for the school administrator questionnaire range from 87.4 percent for schools with 750 or more students to 100 percent for those in large towns (excluding the “unknown” category). Rates for the facilities checklist range from 90.3 percent for schools in the urban fringe of mid-size cities to 100 percent for schools in large towns. It is worth noting that the completion rates for the school administrator questionnaire are lower for schools with higher percentages of minorities, a phenomenon also observed in previous rounds for the school administrator questionnaire. However, this disparity decreased considerably after the base year, reflecting the success of increased data collection efforts targeted toward these schools.

Table 5-11 shows that the rates for the student records abstract are the lowest of all the instruments, as they were in previous years of the ECLS-K. For fifth grade, this rate is about 70 percent compared with 67 percent in third grade. The “unknown” category aside, the completion rates of the student records abstract range from 71.8 percent in the northeast region to 93.7 percent for children in large towns.

All four of the teacher questionnaires were completed at an overall rate of 78 to 80 percent, much higher than the 62 to 63 range achieved in third grade. The completion rates for the teacher-level questionnaire in table 5-11 are uniform across school characteristics, ranging from 86.9 percent for schools in the northeast and schools with 750 or more students to 99.7 percent for schools in large towns (excluding the “unknown” category). The same uniform rates are found for the subject-specific child-

Table 5-9. Number of completed child-level cases and child-level completion rates for the child assessment and parent interview for children sampled in the base year, by school characteristics: School year 2003–04

School characteristics[1]	Child assessment			Parent interview		
		Completion rates			Completion rates	
	Completes[2]	Weighted	Unweighted	Completes[3]	Weighted	Unweighted
All schools	11,260	84.7	93.6	10,913	89.1	90.7
School affiliation						
Public	9,187	96.2	97.7	8,518	89.9	90.6
Private	2,049	95.2	97.7	1,977	92.0	94.3
Catholic	1,313	97.1	98.3	1,260	94.0	94.3
Other private	736	93.1	96.7	717	89.5	94.2
Unknown	24	5.8	4.6	418	82.5	79.5
Type of locale						
Large city	1,863	97.2	97.6	1,697	89.6	88.9
Mid-size city	1,863	97.6	98.3	1,733	90.8	91.5
Urban fringe of large city	3,286	94.8	96.6	3,085	88.6	90.7
Urban fringe of mid-size city	767	93.5	96.8	725	90.3	91.5
Large town	283	99.7	99.6	270	91.9	95.1
Small town	825	94.1	98.8	781	94.3	93.5
Rural – outside MSA	1,286	97.2	97.9	1,205	88.1	91.7
Rural – inside MSA	922	97.7	99.4	871	93.4	93.9
Unknown	165	17.0	24.6	546	83.4	81.5
School size (total enrollment)						
1 to 299	2,359	96.7	98.2	2,255	93.4	93.8
300 to 499	3,703	96.5	98.0	3,453	89.9	91.4
500 to 749	3,167	97.1	97.6	2,945	90.8	90.8
750 or more	1,963	94.2	97.0	1,798	87.2	88.9
Unknown	68	9.7	11.7	462	82.1	79.8

See notes at end of table.

Table 5-9. Number of completed child-level cases and child-level completion rates for the child assessment and parent interview for children sampled in the base year, by school characteristics: School year 2003–04—Continued

School characteristics[1]	Child assessment			Parent interview		
		Completion rates			Completion rates	
	Completes[2]	Weighted	Unweighted	Completes[3]	Weighted	Unweighted
Percent non-White enrolled						
0 – 10	3,509	97.9	98.3	3,352	93.2	93.9
11 – 49	3,705	94.5	97.2	3,522	90.5	92.4
50 – 89	1,956	96.5	97.4	1,768	87.2	88.0
90 – 100	1,997	97.0	98.0	1,786	88.2	87.7
Unknown	93	11.3	15.4	485	82.4	80.3
Region						
Northeast	2,080	95.5	96.8	1,956	90.5	91.1
Midwest	2,957	97.9	98.8	2,803	93.0	93.6
South	3,614	95.1	97.5	3,334	87.6	90.0
West	2,585	96.5	97.3	2,402	91.2	90.4
Unknown	24	5.8	4.6	418	82.5	79.5

[1] School characteristics are for schools attended by children in the ECLS-K fifth-grade sample and are based on ECLS-K survey data, not data from the sampling frame.
[2] Reading, mathematics, or science assessment was scorable, or child was disabled and could not be assessed.
[3] Family structure portion of parent interview was completed.
SOURCE: U.S. Department of Education, National Center for Education Statistics, Early Childhood Longitudinal Study, Kindergarten Class of 1998–99 (ECLS-K), spring 2004.

Table 5-10. Number of completed child-level cases and child-level completion rates for the school administrator questionnaire and facilities checklist for children sampled in the base year, by school characteristics: School year 2003–04

School characteristics[1]	School administrator questionnaire			Facilities checklist		
		Completion rates			Completion rates	
	Completes[2]	Weighted	Unweighted	Completes[2]	Weighted	Unweighted
All schools	10,937	77.1	89.6	11,154	78.8	91.4
School affiliation						
Public	8,884	90.7	94.5	9,084	92.7	96.6
Private	2,053	93.1	97.9	2,070	95.7	98.7
Catholic	1,323	96.8	99.0	1,328	98.1	99.4
Other private	730	88.8	95.9	742	92.8	97.5
Type of locale						
Large city	1,775	92.2	93.0	1,862	96.4	97.6
Mid-size city	1,825	93.6	96.3	1,844	94.9	97.3
Urban fringe of large city	3,173	87.7	93.2	3.240	90.4	95.2
Urban fringe of mid-size city	762	90.9	96.2	761	90.3	96.1
Large town	284	100.0	100.0	284	100.0	100.0
Small town	823	91.2	98.6	823	91.1	98.6
Rural – outside MSA	1,274	92.0	97.0	1,292	93.6	98.3
Rural – inside MSA	904	93.9	97.4	911	94.6	98.2
Unknown	117	8.2	13.9	137	9.4	16.3
School size (total enrollment)						
1 to 299	2,360	95.1	98.2	2,368	95.5	98.5
300 to 499	3,662	92.0	96.9	3.684	92.6	97.5
500 to 749	3,071	92.3	94.6	3,136	94.3	96.6
750 or more	1,844	87.4	91.2	1,939	91.9	95.8
Unknown	0	0.0	0.0	27	1.7	3.6

See notes at end of table.

Table 5-10. Number of completed child-level cases and child-level completion rates for the school administrator questionnaire and facilities checklist for children sampled in the base year, by school characteristics: School year 2003–04—Continued

School characteristics[1]	School administrator questionnaire			Facilities checklist		
		Completion rates			Completion rates	
	Completes[2]	Weighted	Unweighted	Completes[2]	Weighted	Unweighted
Percent non-White enrolled						
0 – 10	3,502	95.0	98.1	3,515	95.5	98.5
11 – 49	3,633	89.8	95.4	3,674	91.1	96.4
50 – 89	1,864	90.4	92.8	1,920	92.9	95.6
90 – 100	1,916	92.4	94.1	1,996	96.2	98.0
Unknown	22	0.9	2.8	49	2.5	6.3
Region						
Northeast	2,008	87.7	93.5	2,048	90.4	95.3
Midwest	2,920	95.5	97.5	2,938	95.9	98.1
South	3,547	90.7	95.7	3,598	92.1	97.1
West	2,462	89.1	92.7	2,570	93.9	96.8

[1] School characteristics are for schools attended by children in the ECLS-K third-grade sample and are based on ECLS-K survey data, not data from the sampling frame.
[2] A completed questionnaire was defined as one that was not completely left blank.
SOURCE: U.S. Department of Education, National Center for Education Statistics, Early Childhood Longitudinal Study, Kindergarten Class of 1998–99 (ECLS-K), spring 2004.

Table 5-11. Number of completed child-level cases and child-level completion rates for the student records abstract and teacher-level questionnaire for children sampled in the base year, by school characteristics: School year 2003–04

School characteristics[1]	Student records abstract			Teacher-level questionnaire		
		Completion rates			Completion rates	
	Completes[2]	Weighted	Unweighted	Completes[2]	Weighted	Unweighted
All schools	10,015	69.9	82.1	10,872	79.9	90.6
School affiliation						
Public	8,177	82.3	86.9	8,849	90.2	94.1
Private	1,838	84.6	87.6	2,023	93.6	96.5
Catholic	1,209	90.3	90.5	1,313	97.0	98.3
Other private	629	77.9	82.7	710	89.4	93.3
Type of locale						
Large city	1,585	78.9	83.1	1,758	89.8	92.1
Mid-size city	1,688	85.3	89.1	1,819	94.1	96.0
Urban fringe of large city	2,778	76.8	81.6	3,131	87.4	92.0
Urban fringe of mid-size city	698	82.1	88.1	751	90.2	94.8
Large town	272	93.7	95.8	283	99.7	99.6
Small town	758	85.7	90.8	819	91.1	98.1
Rural – outside MSA	1,227	89.3	93.4	1,273	91.0	96.9
Rural – inside MSA	884	89.4	95.3	903	93.4	97.3
Unknown	125	8.8	14.9	135	12.0	21.3
School size (total enrollment)						
1 to 299	2,114	85.2	88.0	2,340	94.8	97.4
300 to 499	3,385	84.1	89.6	3,628	91.3	96.0
500 to 749	2,808	83.3	86.5	3,039	91.4	93.7
750 or more	1,687	78.6	83.4	1,844	86.9	91.2
Unknown	21	1.3	2.8	21	1.7	3.9

See note at end of table.

Table 5-11. Number of completed child-level cases and child-level completion rates for the student records abstract and teacher-level questionnaire for children sampled in the base year, by school characteristics: School year 2003–04—Continued

School characteristics[1]	Student records abstract			Teacher-level questionnaire		
		Completion rates			Completion rates	
	Completes[2]	Weighted	Unweighted	Completes[2]	Weighted	Unweighted
Percent non-White enrolled						
0 – 10	3,274	89.0	91.7	3,465	93.9	97.1
11 – 49	3,303	81.4	86.7	3,608	89.2	94.7
50 – 89	1,686	80.2	84.0	1,854	90.5	92.3
90 – 100	1,712	80.2	84.0	1,902	91.4	93.4
Unknown	40	2.1	5.2	43	2.8	7.6
Region						
Northeast	1,635	71.8	76.1	1,970	86.9	91.7
Midwest	2,788	90.4	93.1	2,913	94.7	97.3
South	3,359	83.4	90.7	3,531	90.1	95.3
West	2,233	81.3	84.1	2,458	90.0	92.5

[1] School characteristics are for schools attended by children in the ECLS-K third-grade sample and are based on ECLS-K survey data, not data from the sampling frame.
[2] A completed questionnaire was defined as one that was not completely left blank.
SOURCE: U.S. Department of Education, National Center for Education Statistics, Early Childhood Longitudinal Study, Kindergarten Class of 1998–99 (ECLS-K), spring 2004.

Table 5-12. Number of completed child-level cases and child-level completion rates for the child-level teacher questionnaires for children sampled in the base year, by school characteristics: School year 2003–04

School characteristics[1]	Child-level reading teacher questionnaire			Child-level mathematics teacher questionnaire			Child-level science teacher questionnaire		
		Completion rates			Completion rates			Completion rates	
	Completes[2]	Weighted	Unweighted	Completes[2]	Weighted	Unweighted	Completes[2]	Weighted	Unweighted
All schools	10,793	79.3	90.0	5,339	78.1	89.3	5,405	79.5	89.9
School affiliation									
Public	8,780	89.5	93.3	4,357	87.9	92.6	4,376	89.8	93.1
Private	2,013	92.9	96.0	982	93.6	96.0	1,029	92.1	95.8
Catholic	1,306	96.7	97.8	635	96.5	97.2	670	96.7	98.1
Other private	707	88.5	92.9	347	90.2	93.8	359	86.6	91.8
Type of locale									
Large city	1,753	90.2	91.9	853	90.3	91.8	880	88.0	89.9
Mid-size city	1,811	93.6	95.6	893	90.3	94.3	898	94.0	94.7
Urban fringe of large city	3,111	86.5	91.4	1,547	84.4	90.4	1,551	87.9	91.7
Urban fringe of mid-size city	741	88.3	93.6	366	87.6	93.1	374	88.0	93.7
Large town	283	99.7	99.6	140	99.4	99.3	143	100.0	100.0
Small town	808	90.3	96.8	383	85.0	96.5	432	96.8	98.6
Rural – outside MSA	1,258	90.1	95.7	634	92.8	95.5	620	86.8	95.4
Rural – inside MSA	893	91.8	96.2	456	90.5	95.6	441	94.3	97.8
Unknown	135	12.0	21.3	67	13.1	20.9	66	10.5	21.0
School size (total enrollment)									
1 to 299	2,323	93.8	96.7	1,144	94.2	96.6	1,175	93.7	96.4
300 to 499	3,605	91.0	95.4	1,830	90.7	95.8	1,768	90.2	94.6
500 to 749	3,014	90.6	92.9	1,462	88.8	91.7	1,537	91.3	93.2
750 or more	1,830	85.7	90.5	892	81.7	88.0	917	88.0	90.9
Unknown	21	1.7	3.9	11	2.2	4.0	8	0.7	3.0

See notes at end of table.

Table 5-12. Number of completed child-level cases and child-level completion rates for the child-level teacher questionnaires for children sampled in the base year, by school characteristics: School year 2003–04—Continued

School characteristics[1]	Child-level reading teacher questionnaire			Child-level mathematics teacher questionnaire			Child-level science teacher questionnaire		
		Completion rates			Completion rates			Completion rates	
	Completes[2]	Weighted	Unweighted	Completes[2]	Weighted	Unweighted	Completes[2]	Weighted	Unweighted
Percent non-White enrolled									
0 – 10	3,444	93.7	96.5	1,695	93.2	96.3	1,745	94.1	96.5
11 – 49	3,584	88.1	94.1	1,791	87.2	93.8	1,793	89.0	94.4
50 – 89	1,823	89.1	90.8	919	85.8	89.7	893	90.8	90.8
90 – 100	1,899	91.3	93.2	911	90.4	91.7	956	88.8	91.6
Unknown	43	2.8	7.6	23	3.3	7.9	18	1.9	6.5
Region									
Northeast	1,956	86.7	91.1	967	84.6	90.6	969	87.0	89.6
Midwest	2,909	94.4	97.2	1,423	94.5	96.9	1,482	94.5	97.1
South	3,498	89.0	94.4	1,747	87.5	93.6	1,743	89.2	94.8
West	2,430	89.0	91.5	1,202	87.4	90.5	1,211	89.1	91.2

[1] School characteristics are for schools attended by children in the ECLS-K third-grade sample and are based on ECLS-K survey data, not data from the sampling frame.
[2] A completed questionnaire was defined as one that was not completely left blank.
SOURCE: U.S. Department of Education, National Center for Education Statistics, Early Childhood Longitudinal Study, Kindergarten Class of 1998–99 (ECLS-K), spring 2004.

level teacher questionnaires in table 5-12: 85.7 to 99.7 percent for reading, 81.7 to 99.4 percent for mathematics, and 86.8 to 100 percent for science. These rates are higher than in any previous years of the ECLS-K, in all likelihood due to the higher incentives employed in fifth grade.

As noted above, the rate at which the survey instruments were completed varies markedly by mover status and, within movers, by whether the child was located and followed. As shown in table 5-13, the completion rate for the child assessment was 98.2 percent for children still enrolled in their base year school. For movers it dropped by about 6 points to 91.9 percent for those who were located and followed, and for those not located or followed due to a move to a non-ECLS-K PSU, it was zero. The parent interview completion rates varied from 91.6 percent for nonmovers to 87.1 percent for movers who were located and followed for the purposes of the child assessment, to 85.7 percent for movers who could either not be located or were not followed for the purposes of the child assessment. Even though children who had moved to a non-ECLS-K PSU were not administered the child assessment, a parent interview was conducted by telephone wherever possible, leading to the 86 percent response rate for this category.

The school administrator questionnaire completion rate is 15 points lower for movers, even when the children were located and followed; for the facilities checklist, it is 14 points lower (table 5-14). There are several reasons for this difference: located movers were not always assessed in schools; new schools in which movers enrolled had a lower level of commitment to the ECLS-K and often refused to complete the school administrator questionnaire; and some of these schools were contacted too late in the school year for them to consider completing it. The completion rate for nonmovers was 97.1 percent for the school administrator questionnaire and 98.8 percent for the facilities checklist. For located and followed movers it was 82.4 and 84.8 percent for the school administrator questionnaire and for the facilities checklist, respectively. The rates for the student records abstract are 90.1 percent for nonmovers and 72.3 percent for movers who were located and followed (table 5-15).

The teacher-level questionnaire completion rate, as shown in table 5-15, is about 14 points lower for movers who could be located and followed (82.2 percent) than for nonmovers (96.5 percent). Movers who could not be located were all nonrespondents for this instrument, pulling the overall completion rate for movers downward to 63.9 percent.

Table 5-13. Number of completed child-level cases and child-level completion rates for the child assessment and parent interview for children sampled in the base year, by child's mover status: School year 2003–04

Mover status[1]	Child assessment			Parent interview		
		Completion rates			Completion rates	
	Completes[2]	Weighted	Unweighted	Completes[3]	Weighted	Unweighted
All children	11,260	84.7	93.6	10,913	89.1	90.7
Mover status						
Mover	1,814	71.8	74.7	2,094	86.8	86.2
Located, followed	1,814	91.9	92.1	1,704	87.1	86.5
Other[4]	0	0.0	0.0	390	85.7	85.0
Nonmover	9,446	98.2	98.4	8,819	91.6	91.9

[1] This is the mover status used in weighting, which does not consider children who moved into identified destination schools as movers. A destination school is a school that received at least four students from the school where they had just completed the highest grade.
[2] Reading, mathematics, or science assessment was scorable, or child was disabled and could not be assessed.
[3] Family structure portion of parent interview was completed.
[4] This category includes movers who could not be located, and movers who moved into nonsampled PSUs.
SOURCE: U.S. Department of Education, National Center for Education Statistics, Early Childhood Longitudinal Study, Kindergarten Class of 1998–99 (ECLS-K), spring 2004.

Table 5-14. Number of completed child-level cases and child-level completion rates for the school administrator questionnaire and facilities checklist for children sampled in the base year, by child's mover status: School year 2003–04

Mover status[1]	School administrator questionnaire			Facilities checklist		
		Completion rates			Completion rates	
	Completes[2]	Weighted	Unweighted	Completes[2]	Weighted	Unweighted
All children	10,937	77.1	89.6	11,154	78.8	91.4
Mover status						
Mover	1,589	59.2	61.1	1,647	61.0	63.3
Located, followed	1,589	82.4	82.1	1,647	84.8	85.1
Other[3]	0	0.0	0.0	0	0.0	0.0
Nonmover	9,348	97.1	97.4	9,507	98.8	99.0

[1] This is the mover status used in weighting, which does not consider children who moved into identified destination schools as movers. A destination school is a school that received at least four students from the school where they had just completed the highest grade.
[2] A completed questionnaire was defined as one that was not completely left blank.
[3] This category includes movers who could not be located, and movers who moved into nonsampled PSUs.
SOURCE: U.S. Department of Education, National Center for Education Statistics, Early Childhood Longitudinal Study, Kindergarten Class of 1998–99 (ECLS-K), spring 2004.

Table 5-15. Number of completed child-level cases and child-level completion rates for the student records abstract and teacher-level questionnaire for children sampled in the base year, by child's mover status: School year 2003–04

Mover status[1]	Student records abstract			Teacher-level questionnaire		
		Completion rates			Completion rates	
	Completes[2]	Weighted	Unweighted	Completes[2]	Weighted	Unweighted
All children	10,015	69.9	82.1	10,872	79.9	90.6
Mover status						
Mover	1,393	52.0	53.6	1,587	63.9	66.3
Located, followed	1,393	72.3	72.0	1,587	82.2	82.0
Other[3]	0	0.0	0.0	0	0.0	0.0
Nonmover	8,622	90.1	89.8	9,285	96.5	96.7

[1] This is the mover status used in weighting, which does not consider children who moved into identified destination schools as movers. A destination school is a school that received at least four students from the school where they had just completed the highest grade.
[2] A completed questionnaire was defined as one that was not completely left blank.
[3] This category includes movers who could not be located, and movers who moved into nonsampled PSUs.
SOURCE: U.S. Department of Education, National Center for Education Statistics, Early Childhood Longitudinal Study, Kindergarten Class of 1998–99 (ECLS-K), spring 2004.

Table 5-16 shows the completion rates for all three child-level teacher questionnaires. These rates are between 95 and 96 percent for nonmovers, and between 80 and 82 percent for movers who were located and followed. Children who could not be located were all nonrespondents for the child-level teacher instruments. The reasons for lower completion rates from teachers if the child moved are similar to the reasons that affected the school administrator questionnaire and facilities checklist completion rates for movers.

Tables 5-17 to 5-20 present child-level weighted and unweighted completion rates for the spring-fifth grade data collection for children who were sampled as part of the kindergarten cohort in the base year, this time broken out by child characteristics. When the “unknown” categories are not included, the differences in completion rates by sex and by year of birth are inconsequential, but for race and ethnicity they are more substantial. Table 5-17 shows that for the child assessment the completion rate was highest for Asians (87.6 percent) and lowest for American Indians or Alaska Natives (78.3 percent). For the parent interview it is the opposite; the rate was highest for American Indians or Alaska Natives (95.2 percent) and lowest for Asian children (82.8 percent).

Table 5-18 shows that, excluding the “unknown” categories, the highest completion rates for the school administrator questionnaire and for the facilities checklist are for Pacific Islanders (85.7 percent and 86.7 percent, respectively), and the lowest completion rates are for American Indians or Alaska Natives (65.8 percent and 71.1 percent, respectively). Table 5-19 shows that the completion rate for the student records abstract is highest for children with “other” race (72.8 percent) and the lowest is for Black (63.9), excluding the “unknown” categories.

For the teacher-level questionnaires (table 5-19), the highest rate is for Pacific Islanders (84 percent) and the lowest rate is for American Indians or Alaska Natives (77 percent), excluding the “unknown” categories. For the child-level reading teacher questionnaire (table 5-20), the highest rate is for Asians (82 percent) and the lowest rate is for American Indians or Alaska Natives (76 percent). For the child-level mathematics teacher questionnaire, the highest rate is for Hispanic and for children whose race/ethnicity is not among the listed (80 percent) and the lowest rate is for Black and American Indians or Alaska Natives (76 percent). For the child-level science teacher questionnaire, the highest rate is for Asians (86 percent) and the lowest rate is for Blacks and for children whose race/ethnicity is not among the listed (76 percent).

In addition to the child assessment, parent interview, school administrator questionnaire, facilities checklist, student records abstract, and teacher questionnaires, whose completion rates have been summarized in the preceding tables, data were also collected in fifth grade from the special education teachers for children who followed individualized special education programs. Table 5-21 presents counts of completes and weighted and unweighted completion rates at the overall student level for the special education questionnaires A and B. The number of special education teacher questionnaires is small but their completion rates are high, 92.2 percent for part A, which captures teacher information, and 93.7 percent for part B, which relates to children who receive individualized special education services. These rates are not broken down by school and child characteristics because of the small sample sizes.

Table 5-16. Number of completed child-level cases and child-level completion rates for the child-level teacher questionnaires for children sampled in the base year, by child's mover status: School year 2003–04

Mover status[1]	Child-level reading teacher questionnaire			Child-level mathematics teacher questionnaire			Child-level science teacher questionnaire		
		Completion rates			Completion rates			Completion rates	
	Completes[2]	Weighted	Unweighted	Completes[2]	Weighted	Unweighted	Completes[2]	Weighted	Unweighted
All children	10,793	79.3	90.0	5,339	78.1	89.3	5,405	79.5	89.9
Mover status									
Mover	1,568	63.3	65.5	818	62.2	65.2	734	63.2	64.4
Located, followed	1,568	81.4	81.0	818	79.7	80.4	734	81.7	80.0
Other[3]	0	0.0	0.0	0	0.0	0.0	0	0.0	0.0
Nonmover	9,225	95.8	96.1	4,521	95.3	95.7	4,671	95.6	95.8

[1] This is the mover status used in weighting, which does not consider children who moved into identified destination schools as movers. A destination school is a school that received at least four students from the school where they had just completed the highest grade.
[2] A completed questionnaire was defined as one that was not completely left blank.
[3] This category includes movers who could not be located, and movers into nonsampled PSUs.
SOURCE: U.S. Department of Education, National Center for Education Statistics, Early Childhood Longitudinal Study, Kindergarten Class of 1998–99 (ECLS-K), spring 2004.

Table 5-17. Number of completed child-level cases and child-level completion rates for the child assessment and parent interview for children sampled in the base year, by child characteristics: School year 2003–04

Child characteristics[1]	Child assessment			Parent interview		
		Completion rates			Completion rates	
	Completes[2]	Weighted	Unweighted	Completes[3]	Weighted	Unweighted
All children	11,260	84.7	93.6	10,913	89.1	90.7
Sex[4]						
Male	5,675	84.1	93.3	5,525	89.9	90.8
Female	5,585	85.4	94.1	5,388	88.5	90.8
Race/ethnicity						
White, non-Hispanic	6,466	84.8	94.4	6,394	91.3	93.4
Black, non-Hispanic	1,273	83.3	93.3	1,151	83.4	84.3
Hispanic	2,093	85.9	92.4	2,036	89.5	89.9
Asian	788	87.6	92.5	707	82.8	83.0
Pacific Islander	144	85.4	92.3	136	87.7	87.2
American Indian or Alaska Native	210	78.3	92.1	222	95.2	97.4
Other	272	86.3	93.8	253	84.5	87.2
Unknown	14	55.3	50.0	14	58.0	50.0
Year of birth						
1992	3,307	83.9	93.6	3,211	88.7	90.8
1993	7,896	85.1	93.8	7,646	89.4	90.8
Other/unknown	57	71.0	76.0	56	76.0	74.7

[1] Based on ECLS-K survey data and not on data from the sampling frame.
[2] Reading, mathematics, or science assessment was scorable, or child was disabled and could not be assessed.
[3] Family structure portion of parent interview was completed.
[4] There is a small number of children whose gender is unknown and who did not have completed child assessment and parent interview. The completion rates for these children, being zero, are not included in the table.
SOURCE: U.S. Department of Education, National Center for Education Statistics, Early Childhood Longitudinal Study, Kindergarten Class of 1998–99 (ECLS-K), spring 2004.

Table 5-18. Number of completed child-level cases and child-level completion rates for the school administrator questionnaire and facilities checklist for children sampled in the base year, by child characteristics: School year 2003–04

Child characteristics[1]	School administrator questionnaire			Facilities checklist		
		Completion rates			Completion rates	
	Completes[2]	Weighted	Unweighted	Completes[2]	Weighted	Unweighted
All children	10,937	77.1	89.6	11,154	78.8	91.4
Sex[3]						
Male	5,517	77.1	89.5	5,621	78.6	91.2
Female	5,420	77.2	90.0	5,533	79.2	91.9
Race/ethnicity						
White, non-Hispanic	6,356	78.8	92.6	6,413	80.1	93.4
Black, non-Hispanic	1,256	74.5	89.9	1,266	75.4	90.6
Hispanic	1,957	75.4	83.5	2,061	79.0	88.0
Asian	752	77.9	85.1	776	80.9	87.8
Pacific Islander	145	85.7	92.4	146	86.7	93.0
American Indian or Alaska Native	191	65.8	82.3	208	71.1	89.7
Other	266	77.7	90.8	270	79.1	92.2
Unknown	14	50.1	46.7	14	50.1	46.7
Year of birth						
1992	3,223	77.0	90.0	3,275	78.7	91.5
1993	7,660	77.2	89.7	7,825	79.1	91.6
Other/unknown	54	64.0	70.1	54	59.5	70.1

[1] Based on ECLS-K survey data and not on data from the sampling frame.
[2] A completed questionnaire was defined as one that was not completely left blank.
[3] The completion of the school-level instruments does not depend on whether the child has completed assessment or parent interview data. Hence, while all children with completed assessment or parent interview data have known value of gender; there are children with completed school-level data whose gender is unknown.
SOURCE: U.S. Department of Education, National Center for Education Statistics, Early Childhood Longitudinal Study, Kindergarten Class of 1998–99 (ECLS-K), spring 2004.

Table 5-19. Number of completed child-level cases and child-level completion rates for the student records abstract and teacher-level questionnaire for children sampled in the base year, by child characteristics: School year 2003–04

Child characteristics[1]	Student records abstract			Teacher-level questionnaire		
		Completion rates			Completion rates	
	Completes[2]	Weighted	Unweighted	Completes[2]	Weighted	Unweighted
All children	10,015	69.9	82.1	10,872	79.9	90.6
Sex[3]						
Male	5,034	69.6	81.6	5,482	79.3	90.4
Female	4,981	70.4	82.7	5,390	80.7	91.0
Race/ethnicity						
White, non-Hispanic	5,878	72.6	85.6	6,300	80.6	92.3
Black, non-Hispanic	1,101	63.9	78.8	1,238	77.1	90.8
Hispanic	1,769	68.4	75.5	1,962	80.2	86.9
Asian	692	70.7	78.3	751	82.6	88.2
Pacific Islander	125	69.6	79.6	138	84.0	88.5
American Indian or Alaska Native	197	66.0	84.9	206	77.0	91.6
Other	242	72.8	82.6	263	81.9	91.3
Unknown	11	30.5	36.7	14	55.3	50.0
Year of birth						
1992	2,996	71.5	83.7	3,213	80.1	91.2
1993	6,976	69.4	81.7	7,610	80.1	90.6
Other/unknown	43	57.4	55.8	49	54.4	65.3

[1] Based on ECLS-K survey data and not on data from the sampling frame.
[2] A completed questionnaire was defined as one that was not completely left blank.
[3] The completion of the school-level instruments does not depend on whether the child has completed assessment or parent interview data. Hence, while all children with completed assessment and parent interview data have known value of gender; there is one child with completed facilities checklist data whose gender is unknown.
SOURCE: U.S. Department of Education, National Center for Education Statistics, Early Childhood Longitudinal Study, Kindergarten Class of 1998–99 (ECLS-K), spring 2004.

Table 5-20. Number of completed child-level cases and child-level completion rates for the child-level teacher questionnaires for children sampled in the base year, by child characteristics: School year 2003-04

Child characteristics[1]	Child-level reading teacher questionnaire			Child-level mathematics teacher questionnaire			Child-level science teacher questionnaire		
		Completion rates			Completion rates			Completion rates	
	Completes[2]	Weighted	Unweighted	Completes[2]	Weighted	Unweighted	Completes[2]	Weighted	Unweighted
All children	10,793	79.3	90.0	5,339	78.1	89.3	5,405	79.5	89.9
Sex[3]									
Male	5,436	78.7	89.7	2,670	78.0	88.9	2,739	78.2	89.5
Female	5,357	80.0	90.5	2,669	78.4	89.9	2,666	81.1	90.4
Race/ethnicity									
White, non-Hispanic	6,255	80.0	91.7	3,111	78.6	91.4	3,147	81.4	92.0
Black, non-Hispanic	1,231	76.7	90.2	591	75.7	88.5	629	76.2	90.4
Hispanic	1,952	79.5	86.4	959	79.6	85.2	963	77.2	84.9
Asian	743	81.8	87.3	362	77.8	85.8	378	85.1	88.1
Pacific Islander	132	79.1	84.6	75	77.5	85.2	59	81.1	86.8
American Indian or Alaska Native	204	76.4	90.7	109	75.6	91.6	93	77.0	87.7
Other	262	81.4	91.0	127	80.2	89.4	129	76.2	88.4
Unknown	14	55.3	50.0	5	47.5	35.7	7	31.2	50.0
Year of birth									
1992	3,186	79.1	90.4	1,565	77.9	90.1	1,618	79.9	90.5
1993	7,559	79.6	90.0	3,751	78.4	89.1	3,762	79.7	89.8
Other/unknown	48	53.5	64.0	23	55.0	63.9	25	52.5	64.1

[1] Based on ECLS-K survey data and not on data from the sampling frame.
[2] A completed questionnaire was defined as one that was not completely left blank.
[3] There is a small number of children whose gender is unknown and who did not have completed teacher questionnaire data. The completion rates for these children, being zero, are not included in the table.
SOURCE: U.S. Department of Education, National Center for Education Statistics, Early Childhood Longitudinal Study, Kindergarten Class of 1998–99 (ECLS-K), spring 2004.

Table 5-21. Number of completed instruments and child-level completion rates for the special education teacher questionnaires for children sampled in the base year: School year 2003–04

Category	Completes	Completion rates	
		Weighted	Unweighted
Special education part A[1]	960	92.2	93.8
Special education part B[1]	967	93.7	94.4

[1] A completed instrument was defined as one that was not completely left blank.
SOURCE: U.S. Department of Education, National Center for Education Statistics, Early Childhood Longitudinal Study, Kindergarten Class of 1998–99 (ECLS-K), spring 2004.

5.7.2 Children Sampled in First Grade

In spring-first grade the student sample was freshened to include first-graders who had no chance of selection in the base year because they had not attended kindergarten in the United States or had been in first grade in the fall of 1998. (For a detailed description of the freshening procedure see section 4.3.2.) This same group of children was followed into spring-fifth grade, unless they belonged in the excluded groups. Nonresponse in the freshened student sample could occur at two stages: during the procedure for sampling schools for freshening and identifying children to be used as freshening links in spring-first grade (first component) and then during data collection from the freshened children in spring-fifth grade (second component). The first component alone can further be decomposed into two sources: attrition due to the refusal of entire schools to implement the freshening procedure (*the school term*), and attrition because ECLS-K sampled children had moved to other schools (*the child term*). To contain costs, children who transferred from schools targeted for freshening were not used as links to identify freshened children, even when they were otherwise followed for data collection. These movers were considered freshening nonrespondents in the *child term*.

Table 5-22 presents weighted and unweighted completion rates for freshened children who were fielded in fifth grade. The two components of the completion rates are presented separately in table 5-22. The overall completion rates are the products of the two components. The first component is separated into *a school term* and *a child term* as described earlier. For this component, the completion rate is defined as the freshening completion rates, as opposed to the survey instrument completion rates found in the second component. The weighted freshening completion rate for children in schools targeted for freshening (*the school term*) is 67.2 percent. The reasons for non-participation in the freshening process included refusal or inability to provide the requested information in order to complete the procedures.

Within the schools that agreed to freshen, the freshening completion rate is 98.2 percent, the slight loss due to children who transferred to other schools (*the child term*). Multiplying these two terms together gives a first component completion rate of 66.0 percent. Note that the first component rate for spring-fifth grade is not identical to the first component rate for spring-first grade and spring-third grade because of the exclusion of children in special groups as explained in section 4.5.

The second component varies by survey instrument. The rates for the paper-and-pencil instruments range from 67.0 percent for the student records abstract to 100.0 percent for the special education questionnaire part A. The rate for the child assessment at 78.6 percent is about 6 points lower than for the kindergarten sample and the parent interview, at 81.9 percent, is about 7 points lower. The rates for the school instruments and the student records abstract are also lower than for the kindergarten sample, but by a smaller amount. The rates for some of the teacher instruments are higher than for the kindergarten sample. The final completion rate for each instrument is the product of the two components. Because of the low rates at the first stage, these range from a high of 53.6 percent for the teacher-level questionnaire to a low of 44.2 percent for the student records abstract.

5.7.3 Spring-Fifth Grade Completion Rates—All Children

Table 5-23 presents final fifth-grade completion rates for children sampled in kindergarten, children sampled in first grade, and all children combined. Because children sampled in first grade represent such a small fraction of the total population of children, their inclusion in the computation of the completion rate brings down the combined rate by less than one percent relative to the rates of children sampled in kindergarten, even though the completion rates for children sampled in first-grade rates are much lower than the kindergarten rates. The spring-fifth grade overall completion rates for the child assessment and the parent interview are 83.9 percent and 88.3 percent, respectively. These rates are higher than in third grade by about 4 percentage points for the child assessment and by about 11 percentage points for the parent interview. The unweighted completion rates are almost always higher than the weighted completion rates, by as much as 23 percent at the overall level. Where there is a large difference, it is due to movers who have larger weights and higher nonresponse rates than nonmovers. The weights of the movers were increased to account for the subsampling of movers. They also responded at a much lower rate than nonmovers, as shown earlier in table 5-6. This difference is larger than in previous years because movers in fifth grade have much larger weights than in previous years (many more movers were not included in fifth grade, necessitating larger adjustment factors). Note that the unweighted completion rates follow the traditional ECLS-K pattern, that is, rates for the child assessment are higher than rates for the parent interview (93.4 percent for the child assessment and 90.5 percent for the parent interview). This is again due to movers with large weights and to the fact that there are more parent-responding movers than child-responding movers. Thus, the weighted completion rates are higher for the parent interview than for the child assessment.

Table 5-24 shows the completion rates for the parent interviews and the school and teacher instruments for children who have nonzero child weights (C6CW0>0). These are children whose fifth-grade reading, mathematics or science assessment were scorable, or children who could not be assessed because of disabilities. For these children, the completion rate for the child assessment should be 100 percent. The rate slightly less than 100 percent, shown when children sampled in kindergarten are combined with children sampled in first grade, is due to the school freshening nonresponse for children sampled in first grade.

Table 5-22. Number of completed child-level cases and child-level completion rates for children sampled in first grade: School year 2003-04

Category	Completes	Completion rates[1]	
		Weighted	Unweighted
First component (first-grade sample freshening)	7,032	66.0	79.1
School term[2]	7,089	67.2	79.7
Child term[3]	7,135	98.2	99.2
Second component (fifth-grade data collection)			
Child assessment[4]	86	78.6	86.0
Parent interview[5]	83	81.9	83.0
School administrator questionnaire[6]	86	74.4	83.5
Facilities check list[6]	89	77.8	86.4
Student records abstract[6]	75	67.0	72.8
Teacher-level questionnaire[6]	87	81.2	87.9
Reading teacher questionnaire (child level)[6]	84	79.8	84.8
Mathematics teacher questionnaire (child level)[6]	41	81.0	83.7
Science teacher questionnaire (child level)[6]	40	75.2	80.0
Special education part A[6]	15	100.0	100.0
Special education part B[6]	14	93.4	93.3
Overall completion rates			
Child assessment[4]	86	51.9	68.0
Parent interview[5]	83	54.0	65.6
School administrator questionnaire[6]	86	49.1	66.0
Facilities check list[6]	89	51.3	68.3
Student records abstract[6]	75	44.2	57.6
Teacher-level questionnaire[6]	87	53.6	69.5
Reading teacher questionnaire (child level)[6]	84	52.7	67.0
Mathematics teacher questionnaire (child level)[6]	41	53.5	66.2
Science teacher questionnaire (child level)[6]	40	49.6	63.2
Special education part A[6]	15	66.0	79.1
Special education part B[6]	14	61.6	73.8

[1] In the first component, this is the completion rate for freshening. In the second component, this is the completion rate for the survey instruments. The product of the two components is the overall completion rate for the survey instruments.
[2] The freshening completes and completion rates for children in schools targeted for freshening.
[3] The freshening completes and completion rates for children in schools that agreed to the freshening procedure.
[4] Reading, mathematics, or science assessment was scorable, or child was disabled and could not be assessed.
[5] Family structure portion of parent interview was completed.
[6] A completed questionnaire was defined as one that was not completely left blank.
SOURCE: U.S. Department of Education, National Center for Education Statistics, Early Childhood Longitudinal Study, Kindergarten Class of 1998–99 (ECLS-K), spring 2004.

Table 5-23. Number of completed child-level cases and child-level completion rates, for children sampled in kindergarten and first grade, by survey instruments: School year 2003-04

Survey instrument	Children sampled in kindergarten			Children sampled in first grade			All children		
		Completion rates			Completion rates			Completion rates	
	Completes	Weighted	Unweighted	Completes	Weighted	Unweighted	Completes	Weighted	Unweighted
Child assessment[1]	11,260	84.7	93.6	86	51.9	68.0	11,346	83.9	93.4
Parent interview[2]	10,913	89.1	90.7	83	54.0	65.6	10,996	88.3	90.5
School administrator questionnaire[3]	10,937	77.1	89.6	86	49.1	66.0	11,023	76.4	89.4
Facilities check list[3]	11,154	78.8	91.4	89	51.3	68.3	11,243	78.1	91.2
Student records abstract[3]	10,015	69.9	82.1	75	44.2	57.6	10,090	69.3	81.9
Teacher-level questionnaire[3]	10,872	79.9	90.6	87	53.6	69.5	10,959	79.3	90.4
Reading teacher questionnaire (child level)[3]	10,793	79.3	90.0	84	52.7	67.0	10,877	78.7	89.8
Mathematics teacher questionnaire (child level)[3]	5,339	78.1	89.3	41	53.5	66.2	5,380	77.5	89.1
Science teacher questionnaire (child level)[3]	5,405	79.5	89.9	40	49.6	63.2	5,445	78.8	89.7
Special education part A[3]	960	92.2	93.8	15	66.0	79.1	975	91.6	93.7
Special education part B[3]	967	93.7	94.4	14	61.6	73.8	981	92.9	94.2

[1] Reading, mathematics, or science assessment was scorable, or child was disabled and could not be assessed.
[2] Family structure portion of parent interview was completed.
[3] A completed questionnaire was defined as one that was not completely left blank.
SOURCE: U.S. Department of Education, National Center for Education Statistics, Early Childhood Longitudinal Study, Kindergarten Class of 1998–99 (ECLS-K), spring 2004.

Table 5-24. Number of completed child-level cases and child-level completion rates, for children with scorable reading, mathematics or science assessment or children not assessed due to disabilities, by survey instruments: School year 2003-04

Survey instrument	Children sampled in kindergarten			Children sampled in first grade			All children		
		Completion rates			Completion rates			Completion rates	
	Completes	Weighted	Unweighted	Completes	Weighted	Unweighted	Completes	Weighted	Unweighted
Child assessment[1]	11,260	100.0	100.0	86	78.3	86.8	11,346	99.5	99.9
Parent interview[2]	10,445	92.4	92.8	77	71.5	77.6	10,522	91.9	92.7
School administrator questionnaire[3]	10,794	93.4	96.1	83	74.7	83.7	10,877	93.0	96.0
Facilities check list[3]	11,015	95.7	98.0	86	78.3	86.8	11,101	95.3	97.9
Student records abstract[3]	9,986	85.6	88.9	74	68.3	74.6	10,060	85.2	88.8
Teacher-level questionnaire[3]	10,799	93.4	96.1	84	77.3	84.8	10,883	93.0	96.0
Reading teacher questionnaire (child level)[3]	10,774	93.3	95.9	84	77.3	84.8	10,858	92.9	95.8
Mathematics teacher questionnaire (child level)[3]	5,331	92.7	95.5	41	76.2	82.7	5,372	92.3	95.4
Science teacher questionnaire (child level)[3]	5,394	92.8	95.4	40	74.9	80.7	5,434	92.4	95.3
Special education part A[3]	947	93.0	94.1	14	78.3	86.8	961	92.6	94.0
Special education part B[3]	959	95.1	95.3	14	78.3	86.8	973	94.7	95.2

[1] Reading, mathematics, or science assessment was scorable, or child was disabled and could not be assessed.
[2] Family structure portion of parent interview was completed.
[3] A completed questionnaire was defined as one that was not completely left blank.
SOURCE: U.S. Department of Education, National Center for Education Statistics, Early Childhood Longitudinal Study, Kindergarten Class of 1998–99 (ECLS-K), spring 2004.

When the completion rates are conditioned on the presence of the child weight, they are at least 13 points higher than the unconditional completion rates for all instruments but the parent interview and the special education questionnaires. For these last two instruments, the difference between the number of completes for the conditional and unconditional rates is very small; hence the conditional rates are not affected as much as for the other instruments. For the parent interview, the unconditional rate is fairly high for the reason explained earlier that is, movers in fifth grade have much larger weights than in previous years and there are more parent-responding movers than child-responding movers (the weighted completion rates are higher for the parent interview than for the child assessment). This results in the smaller difference between conditional and unconditional of about 4 percent. For all the other instruments, the conditional completion rates are higher by 13.6 points for child-level science teacher questionnaire and as high as 17.2 points for the facilities checklist.

As explained in section 4.5, four groups of children were excluded from the fifth-grade data collection. These are (1) children who became ineligible in an earlier round (because they had died or had moved out of the country), (2) children who were subsampled out in previous rounds because they had moved out of the original schools and were not subsampled to be followed, (3) children whose parents emphatically refused to cooperate (hard refusals) in any of the data collection rounds since spring-kindergarten, and (4) children eligible for the third-grade data collection for whom there are neither first-grade nor third-grade data. Table 5-25 shows the completion rates for all instruments had children in the last two exclusion groups been counted as nonrespondents. These are children who would have been eligible for the fifth-grade collection but past experience showed that they would be most likely nonrespondents. When compared to table 5-23, the completion rates for all instruments in table 5-25 are lower as expected, but only by about 2 percent, with the smallest difference for the student records abstract and the largest difference for the parent interview. Note that the rates for mathematics and science teacher appear to be unchanged. Recall that only about half of the children had mathematics teacher questionnaires and the other half had science teacher questionnaires. Since the mathematics/ science sampling flags were not assigned to children not included in the sample, we were not able to compute a correct completion rate for these two instruments for these children. But the pattern of completion rates would be the same as for the other instruments and we would expect a drop of about 2 percent for the mathematics and science teacher questionnaire.

Table 5-25. Number of completed child-level cases and child-level completion rates, if excluded children were fielded, by survey instruments: School year 2003-04

Survey instrument	Children sampled in kindergarten			Children sampled in first grade			All children		
		Completion rates			Completion rates			Completion rates	
	Completes	Weighted	Unweighted	Completes	Weighted	Unweighted	Completes	Weighted	Unweighted
Child assessment[1]	11,260	82.7	86.7	86	47.6	57.7	11,346	81.9	86.5
Parent interview[2]	10,913	87.0	84.0	83	49.7	55.8	10,996	86.1	83.8
School administrator questionnaire[3]	10,937	75.3	83.3	86	45.2	56.3	11,023	74.6	83.1
Facilities check list[3]	11,154	77.0	84.9	89	47.3	58.2	11,243	76.3	84.7
Student records abstract[3]	10,015	68.3	76.3	75	40.7	49.1	10,090	67.6	76.1
Teacher-level questionnaire[3]	10,872	78.0	84.1	87	49.2	58.9	10,959	77.3	83.9
Reading teacher questionnaire (child level)[3]	10,793	77.4	83.5	84	48.3	56.8	10,877	76.7	83.3
Mathematics teacher questionnaire (child level)[3]	5,339	78.1	89.3	41	52.1	65.2	5,380	77.5	89.1
Science teacher questionnaire (child level)[3]	5,405	79.5	89.9	40	48.3	62.3	5,445	78.8	89.7
Special education part A[3]	960	92.2	93.8	15	64.3	77.9	975	91.5	93.7
Special education part B[3]	967	93.7	94.4	14	60.0	72.7	981	92.9	94.2

[1] Reading, mathematics, or science assessment was scorable, or child was disabled and could not be assessed.
[2] Family structure portion of parent interview was completed.
[3] A completed questionnaire was defined as one that was not completely left blank.
SOURCE: U.S. Department of Education, National Center for Education Statistics, Early Childhood Longitudinal Study, Kindergarten Class of 1998–99 (ECLS-K), spring 2004.

This page is intentionally left blank.

6. DATA PREPARATION

As described in chapter 5, two types of data collection instruments were used for the Early Childhood Longitudinal Study, Kindergarten Class of 1998–99 (ECLS-K) data collection in the spring-fifth-grade: computer-assisted interviews (CAI) and self-administered paper forms (hard copy). The data preparation approach differed with the mode of data collection. The direct child assessments and parent interview were conducted using CAI techniques. Editing specifications were built into the computer programs used by assessors or interviewers to collect these data. The teacher and school administrator forms were self-administered. When the field supervisors returned these forms, coders recorded the receipt of these forms into a project-specific forms tracking system. Coders reviewed the questionnaires to ensure readability of data for transfer into an electronic format. Upcoding was conducted after the data were keyed. Once they finished this review, the coders sent the instruments to data entry to be manually transferred to an electronic format and reviewed for range and logic consistency. The following sections describe the data preparation activities for both modes of data collection in more detail.

6.1 Coding and Editing Specifications for Computer-Assisted Interviews (CAI)

The very nature of designing a computer-assisted interview forces decisions about edit specifications to be made up front. Both acceptable ranges and logic consistency checks were preprogrammed into the electronic questionnaire. The next few sections describe the coding and editing of the data collected using CAI. Though the child assessments and the parent interviews were both collected using CAI, the child assessments did not contain some of the additional range and edit checks contained in the parent interview. The following sections describe the coding and editing that were conducted on the CAI parent interview.

6.1.1 Range Specifications

Within the CAI parent interview instruments, respondent answers were subjected to both “hard” and “soft” range edits during the interviewing process. (The child assessment did not have such hard and soft ranges.) 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 attended 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 assessor or 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, "0-5 times" for the number of times the child, in the previous 5 days, ate a breakfast that was not school provided. 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 assessor or 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 and among 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, in the household roster, one could not be recorded as both a sister and male. When a logical error such as this occurred during a session, a message appeared requesting verification of the last response and a resolution of the discrepancy. In some instances, if the verified response still resulted in a logical error, the assessor or interviewer recorded the problem either in a comment or on a problem report. Consistency checks were not applicable to the child assessments.

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, occupation, and race/ethnicity. Assessors or 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. In this section, we describe the coding activities for the CAI instruments.

Review of "Other, specify" items. The "Other, specify" open-ended parent interview responses were reviewed to determine if they should be coded into one of the existing response categories. During data collection, when a respondent selected an "other" response in the parent interview, the assessor or 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. There were no "Other, specify" items in the child assessments.

Parent occupation coding. As in the kindergarten, first-grade, and third-grade data collections, occupations were coded using the Manual for Coding Industries and Occupations (NCES 2000-077). This coding manual was created for the Adult Education Survey of the National Household Education Surveys Program (AE-NHES)(1999) and used an aggregated version of industry and occupation codes. The industry and occupation codes used by NHES were originally developed for the 1989–90 National Postsecondary Student Aid Study (NPSAS)(1990) and contained one to four digits. Analysis of the NPSAS categories revealed that some categories had very small numbers of cases and some categories that are similar had similar participation rates, suggesting that the separate codes could be collapsed without significant loss of information. The NHES industry and occupation code categories use a two-digit code, the highest level of aggregation, to have sufficient numbers of cases to support analysis without collapsing categories. There are 13 industry codes and 22 occupation codes in the NHES coding scheme. If an industry or occupation could not be coded using this manual, the *Index of Industries and Occupations*—1980 (U.S. Department of Commerce 1982) and *Standard Occupational Classification Manual—1980* (U.S. Department of Commerce 1980) were used. Both of these manuals use an expanded coding system and, at the same time, are directly related to the much more condensed NHES coding scheme. These manuals were used as references in cases where the NHES coding scheme did not adequately cover a particular situation. (See chapter 7, section 7.5.7 for an expanded description of the industry and occupation codes.)

Occupation coding began with an autocoding procedure using a computer string match program developed for the NHES. The program searched the responses for strings of text for each record/case and assigned an appropriate code. A little over a third of the cases were autocoded (37.8 percent).

Cases that could not be coded using the autocoding system were coded manually using a customized coding 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, and name of the employer, was available for the coders. The coders used this information to ensure that the occupation code assigned to each case was appropriate. Over half the cases (62.2 percent) were manually coded.

The cases were then verified. Verification of coding is an important tool for ensuring quality control and extending coder training. As a verification step, two coders independently assigned codes (i.e., a double-blind coding process) to industry and occupation cases. A coding supervisor arbitrated disagreements between the initial code and the verification code. The arbitration by the supervisor served to further train coders by presenting concrete examples of appropriate coding. Initially 100 percent of each coder's work was reviewed. Once the coder's error rate had dropped to 1 percent or less, 10 percent of the coder's work was reviewed. Of the cases that were autocoded, 8.9 percent required adjudication because the verifier disagreed with the autocoding. Of the cases that were manually coded, 21.2 percent required adjudication because the manual coder and the verifier disagreed.

Race/ethnicity coding. The same coding rules used in the kindergarten year were used to code all race/ethnicity variables for children, resident parents, and nonresident parents. (See chapter 7, section 7.5.1 for details on how the race variables were coded and how the race/ethnicity composite was created.)

Partially complete parent interviews. A "completed" parent instrument was defined by whether the section on family structure (FSQ) was completed by the respondent. Only completed **interviews** were retained in the final data file. A small number of interviews in fifth grade (83, less than 1 percent) terminated the parent interview after the FSQ section but before the end of the instrument. These interviews were considered as "partially complete" cases and were included in the data file. All instrument items after the interview termination point were set to -9 for "not ascertained."

Household roster in the parent interview. Several tests were run on the household roster to identify missing or inaccurate information. These tests are the same tests run on the first-grade and third-grade files. One flag was used to identify cases that were edited for any of the reasons described below. The flag is P6EDIT; the flag was set to 1 if the case was edited in the given wave. There were 446 cases requiring edits in fifth grade.

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. Problems found were corrected on the basis of data from prior data collections wherever possible. 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 whenever clear errors and a clear resolution existed. Lastly, the relationship of an individual to both the focal child and the reference person was examined, as there were cases in which the relationship of an individual to the focal child conflicted with his or her status as the spouse/partner of the reference person. For example, in a household containing a child's grandparents but not his or her parents, the grandmother may be designated the "mother" figure, and the grandfather thus becomes the "father" (for the purposes of some questions in the interview) 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. In the fifth-grade data, there are 76 cases with these types of errors after the roster tests were run; the cases can be identified by the flag "P6ERRFLG."

Teacher responses to key child items. Teachers of sampled children were asked to respond to child-level questionnaires for the reading, mathematics, and science domains. In many cases, teachers had more than one sampled child in a class. The items in the child-level questionnaire that collected information about classroom characteristics were redundant under these circumstances. The key child approach was designed to minimize the burden on the teachers by designating one questionnaire in which the classroom characteristics items were to be completed. See section 5.3.2 for a description of the key child design and procedures.

Once the child-level questionnaires were keyed and loaded into the editing system, a review was conducted to identify cases in which teachers reported classroom characteristics on a different questionnaire than the one designated as the key child instrument for the given class. This process involved three steps: the review of missing data for classroom characteristics items within each domain (reading, mathematics, and science) for key child records, a detailed review of all data records in classes with multiple children and missing values for selected classroom characteristics items, and the updating of appropriate records.

In the first step, data records for key children in all classrooms with more than one sampled child were selected. Frequency distributions of the classroom items were examined for the level of missing data within each domain. All classroom characteristics items were included in this review. The results of this initial review indicated that missingness was largely confined to the items concerning the race and sex composition of the classroom.

In the second step, all returned instruments were selected for classrooms with multiple children that had missing data for the race and sex composition items. These cases were reviewed to ascertain whether the teacher had mistakenly reported the classroom characteristics items on a questionnaire other than that designated for the key child.

In the third step, update specifications were prepared, directing data preparation staff to apply the classroom characteristics data to the key child record for the classroom. Updates were made to 10 reading records, 5 mathematics records, and 3 science records as a result of this review.

A review was also conducted to identify classrooms with multiple sampled children for which no key child instrument was returned. There were 5 such cases for reading, 7 such cases for mathematics, and 3 such cases for science. Another child for whom an instrument was returned was designated as the key child in these classrooms.

6.2 Coding and Editing Specifications for Hard-Copy Questionnaires

6.2.1 Receipt Control

In order to monitor the more than 40,000 documents that were to be received in the fifth-grade year, the project-specific receipt and document control system developed in the kindergarten year was used, with some modifications. The receipt and document control system was initially loaded with the identifying information, such as identification numbers for schools, teachers, and children; the links between teachers and children; and the questionnaires that were expected from each school and teacher for each cooperating school in the sample. As data were collected in the field, field supervisors completed transmittal forms for each school to indicate which questionnaires were being mailed to the home office.

Once data collection started, receipt control clerks reviewed the questionnaires returned from the field for accuracy and completeness. The identification number on each form was matched against the identification numbers in the tracking system to verify that the appropriate number of forms for each school was returned. When the clerks verified that the correct questionnaires were returned, they were logged into the receipt and document control system. Once forms were logged in, if they had any data (some forms had no data due to refusal by the respondent to complete them), they were then coded. The data were then keyed into electronic format and edited.

The following sections describe the coding, data entry, and editing processes for hard-copy questionnaires.

6.2.2 Coding

The hard-copy questionnaires required coding of race/ethnicity for teachers, review of "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, and consistent across variables and that the numbers were converted to the appropriate unit of measurement prior to converting data to an electronic format. The coding staff were trained on the coding procedures and had coding manuals to support the coding process. This staff also edited the data after data entry was complete. Senior coders verified coding. The verification rate was set at 100 percent for each coder until an error rate of less than 1 percent was established. After that point, work was reviewed at a rate of 10 percent.

Review of "Other, specify" items. The "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 text responses that remained after upcoding did not fit into any preexisting category.

Coding teacher/race ethnicity. "Other, specify" text responses for race/ethnicity in the teacher questionnaire part B were coded using the kindergarten, first-grade and third-grade procedures. Many of these "others" included more than one response (e.g., African American/Asian or American Indian/White). The open responses were coded into one or more of the following seven categories: one Hispanic category; White, non-Hispanic; Black or African American, non-Hispanic; American Indian or

Alaska Native; Asian; Native Hawaiian or other Pacific Islander; and one unspecified multirace-ethnicity category.

Coding teacher language. "Other, specify" text responses for language in the teacher questionnaire part A were coded using the kindergarten, first-grade, and third-grade procedures. Languages beyond the options provided were recorded in "Other, specify." Groups of languages were created based on geographic boundaries. Additional languages included African language; Eastern European language; Native American language; sign language; Middle Eastern language; Western European language; Indian subcontinent language; Southeast Asian language; Pacific Islander language; and other language.

6.2.3 Data Entry

Westat data entry staff keyed the forms in each batch. The data were rekeyed by more senior data entry operators at a rate of 100 percent to verify the data entry. The results of the two data entry passes were compared and differences identified. The hard-copy form was pulled and examined to determine what corrections had to be made to the keyed data. These corrections were rekeyed, resulting in an accuracy rate exceeding 99 percent. The verified batches were then transmitted electronically to Westat's computer system for data editing.

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.

Range specifications. Hard-copy range specifications set the parameters for high and low acceptable values for a question. Where values were printed on the forms, these were used as the range parameters. For open-ended questions, such as, "Counting this school year, how many years have you taught in your *current school* including part-time teaching?," high and low ranges were established as acceptable values. Data frequencies were run on the range of values to identify any errors. Values outside the range were identified as errors and were printed on hard copy for a data editor to review. Cases identified with range errors were identified, and the original response was updated. In some cases, range violations were retained in the data because the value was checked and found to be the value reported by the teacher or school. These were marked as “keep as is” cases. Data frequencies were then rerun and reviewed. This iterative process was repeated until no further range errors were found.

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 school administrator questionnaire, the number of children eligible for free breakfast could not exceed the total number of children enrolled in the school. 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 batches of 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 corrected (or checked and marked “keep as is”) 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.

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 teacher questionnaire part A, an item asking about languages other than English spoken in the classroom included a response option of “No language other than English.” If a respondent circled that response, but also answered (in subsequent items) that other languages besides English were spoken in the classroom, then the response was left as recorded by the respondent because the discrepancy could not be resolved.

This page is intentionally left blank.

7. DATA FILE CONTENT AND COMPOSITE VARIABLES

This chapter describes the content of the Early Childhood Longitudinal Study, Kindergarten Class of 1998-99 (ECLS-K) Fifth-Grade Data File and focuses largely on the composite variables that have been created. The fifth-grade data file can be used for longitudinal analysis in combination with the files from the base year (kindergarten year), first grade, and third grade; see chapter 9 for details about longitudinal analyses. For reference, the Base-Year, First-, and Third-Grade User's Manuals are included in appendix C of the fifth-grade electronic codebook (ECB).

There is one child-level fifth-grade data file or catalog, as noted in chapter 1. Each child record contains data from the various respondents associated with the child (the child herself/himself, a parent, one or more teachers, and a school administrator), as well as from the facilities checklist, school records, and the Field Management System (FMS).

The fifth-grade child catalog contains one record for each of the 11,820 participating students in spring-fifth grade. Included in the file are cases with either a child assessment, a parent interview, or both. Fifth-grade school- and teacher-level data, including composites, are also stored in the child catalog. The file, named child5p.dat for the public use data file and child5r.dat for the restricted-use data file, 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 ECB software available on the CD-ROM rather than access the ASCII file directly. 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 7.4 focus on the conventions used in the study and describe identification variables, the structure of the teacher variables, missing values, and variable names. Section 7.5 provides details about the creation of composite variables on the fifth-grade data file. Section 7.6 focuses on the methodological variables. Section 7.7 discusses variables used to identify children who changed schools. Section 7.8 contains a table of the composite variables. Finally, section 7.9 describes masked variables.

7.1 Identification Variables

The fifth-grade data file contains a child identification (ID) variable (CHILDID) that uniquely identifies each record. Teachers on the child records are identified with the ID variables J61T_ID (reading teacher ID) and J62T_ID (mathematics or science teacher ID). The structure of the teacher data is different in spring-fifth grade than in previous rounds of the study because, rather than one main teacher, reading and mathematics or science teachers were asked to provide data. Information about how to use these data and how they are stored is provided in section 7.2. In addition to teacher identification numbers, there are also identification numbers that indicate whether a child was assigned to a particular class (reading and math/science). For reading, the ID variable name is J61CLASS. For math/science, it is J62CLASS.

Schools are identified by the ID variable S6_ID (spring-fifth grade). The ID variable S6_ID indicates the school the child attended at the time of the spring-fifth grade data collection. Schools that joined the ECLS-K in the fifth grade have an "A" as the first digit. Another identification variable indicates whether the child moved within spring-fifth grade. Section 7.7 provides further details on identifying children who changed schools.

Each type of respondent (child, parent, reading teacher, mathematics or science teacher, special education teacher, and school) has a unique ID number. The original school ID number (S_ID) is the base for all the subsequent ID numbers as children, parents, and teachers were sampled from schools during the base year. The school ID number is a four-digit number assigned sequentially to sampled schools. The number has a series of ranges: 0001–1299 for originally sampled schools; 2000 series for new schools added to the sample during the first grade sample freshening process; 3000 series for substitute schools that replaced nonresponding original sample schools; and 4000 through 6000 series for transfer schools, which were assigned during processing at the home office. (See chapter 4 for a complete description of the ECLS-K sample.) There is also a 9000 series of S_ID numbers that refers to children who do not attend regular school because they are schooled at home (S_ID numbers 9101 through 9499). There are also several specific 9000 series codes for children who were not located or not followed at the end of a round. The school ID numbers start with 999 for these cases. These are described in section 7.6.

The child ID number (CHILDID) is a concatenation of the school ID where the child was sampled, a three-digit student number and the letter "C." For example, 0001010C is the ID number of the tenth child sampled in school 0001. The teacher ID numbers (J61T_ID and J62T_ID) are a concatenation of the school ID where the teacher was sampled, the letter "T," and a two-digit teacher number. In previous rounds of the study, the numbering for the two-digit teacher number started with 01, such that 0001T01 was the ID number for the first teacher sampled in school 0001. In spring-fifth grade, the two-digit teacher numbers started numbering with T60 so that the teachers from this round of the study could be identified easily. Thus, in spring-fifth grade 0002T60 is the ID number for the first teacher sampled in school 0002. The parent ID number (PARENTID) is linked to the child ID number and is a concatenation of the four digit school ID, the three digit student number, and the letter "P." It is the same number as the child ID with a letter "P" instead of a letter "C" at the end. For example, 0001010P is the ID number of the parent of the tenth child sampled in school 0001. If twins are sampled, the ID of the first child sampled is used to generate the parent ID. For twins, there are two child-level records with the same parent ID. Children with the same teacher can be identified by finding all children on the child file with the same teacher ID.

It should be noted that there is a difference in the variable names between the base year and the first-, third-, and fifth-grade special education teacher IDs. In the base year of the study, information from special education teachers was included in a separate file and was not part of the child or teacher catalogs. The ID number for special education teachers in the base year special education file was T_ID. In the fifth-grade data file (and the first- and third-grade data files), the special education teacher information is included with the rest of the data, necessitating ID numbers to distinguish special education teachers from regular education teachers. In the fifth grade file, J61T_ID and J62T_ID are used to identify regular education teachers and D6T_ID is used to identify special education teachers.

If there is no special education teacher, D6T_ID will be missing. If there is a special education teacher, D6T_ID will be filled whether or not the special education teacher responded. In either case, it should be noted that there could be missing data for special education data in the part B questionnaire. It is left to users to determine how they would like to set "Not Applicable" versus "Not Ascertained" codes for such combinations. Users interested in links to special education services, regardless of whether the source of the information was the starting or ending school, can use the composite variable F6SPECS that is based on information from the FMS system rather than the receipt of particular special education questionnaires.

7.2 Using Teacher Variables

In the fifth grade, more students were expected to have different teachers for reading, math, and science than in previous rounds of the study and the teacher questionnaires were changed to be specific to each subject to reflect this. For the spring-fifth grade data collection, all students were assigned to have a reading teacher complete questionnaires. Half of students were assigned to have a mathematics teacher complete questionnaires, and the other half of students were assigned to have a science teacher complete questionnaires. Thus, each student was linked to a maximum of two teachers: one for reading, and one for either mathematics or science. However, a teacher could be linked to any number of students. In addition, although each student was only linked for two subjects, a teacher could be linked for three subjects (e.g., linked to student 1 for reading/math, and linked to student 2 for reading/science).

There are two types of data collected from teachers, taken from four questionnaires. The first type is data about the teacher's background and topics such as instructional practices, classroom resources, views on teaching, and the school, collected in the teacher questionnaire (one per each teacher linked to a responding ECLS-K student). The second type is data about the child, as reported by the reading, math, and science teacher.

As discussed in section 7.1, teachers on the child records are identified with the ID variables J61T_ID (reading teacher ID) and J62T_ID (mathematics or science teacher ID). These ID variables indicate the teacher ID that links to the child regardless of whether there were data received from that teacher. To determine whether data were receipted from a teacher, flag variables must be used. These flags are described below.

7.2.1 Teacher Flags (J61TQUEX, J62TQUEX, F6MTHSCI, T6SAMTCH)

There are three teacher flags on the file (J61TQUEX, J62TQUEX, F6MTHSCI) that identify the presence or absence of teacher data and indicate if the data are from the reading, math, or science teacher. There is also a flag (T6SAMTCH) that indicates if the teacher linked to the child for reading and math/science was the same. In previous rounds of the study, there was only one teacher (other than a special education teacher, if applicable) assigned to answer questions about the child, and there were flags corresponding to each of the three teacher questionnaires (parts A, B, and C) given to this teacher. In spring-fifth grade, the flags also correspond to different teacher questionnaires but the data were collected from reading, math, and science teachers.

The flag J61TQUEX indicates whether there were reading teacher data collected (0 = False; 1 = True) and the flag J62TQUEX indicates whether there were mathematics or science teacher data collected (0 = False; 1 = True). To determine whether the child was linked to a mathematics or science teacher, the flag F6MTHSCI should be used (1=Math, 2=Science).

Using the flags J6TQUEX and F6MTHSCI together will indicate the presence or absence of data and whether the data were for mathematics or science. For example, if a user sought to examine science teacher data, he or she would first determine whether mathematics or science teacher data had been collected (J62TQUEX = 1) and, if so, examine data for children who were linked to a science teacher (F6MTHSCI = 2) rather than a mathematics teacher (F6MTHSCI = 1). If the child had science teacher data, the user would look at science questionnaire variables (all of which begin with the prefix N6). Mathematics teacher data (variables beginning with the prefix M6) would be missing for that child. Further information on variable prefixes is in section 7.4 below.

There is also a flag (T6SAMTCH) that indicates if the same teacher was linked to the child for both reading and math/science. If the value of the flag is 1 (True), then the teacher linked to the child for reading and math/science was the same person. If the value of the flag is 0 (False), then the teachers linked to the child for reading and math/science were different.

7.3 Missing Values

All variables in the ECLS-K 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. Missing values codes, School years 1998–99, 1999–2000, 2001–02, and 2003–04

Value	Description
-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, Kindergarten Class of 1998–99 (ECLS-K), fall 1998, spring 1999, fall 1999, spring 2000, spring 2002, and spring 2004.

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 interview, 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, a question about a sibling's age is not asked when the respondent has indicated that the child has no siblings. A "Not Applicable" code is also used in the direct child assessment if a child did not participate in any section due to a disability. For the teacher and school 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 school administrator questionnaire Question 21. Question 21 asks whether the school received Federal Title I funds for this school year. If the answer is "yes," the questionnaire skips to question 22 about whether the school is operating a Title I targeted assistance or schoolwide program. If the answer is "no," the questionnaire skips to question 24. If question 21 was left blank by the respondent, question 22 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 (-8) and the "Not Ascertained" code (-9), indicates item nonresponse. The "Refused" code rarely appears in the school and teacher data 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 did 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, a “-8” will not be coded for those that choose 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 school and teacher 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.

“System Missing” appears as a blank when viewing codebook frequencies and in the ASCII data file. System missing codes (blanks) in the fifth- grade data file indicate that an entire instrument or assessment is missing due to unit nonresponse. (Note that in the first grade, system missing also indicated that some questions were not asked in the school administrator questionnaire for returning schools but were asked in another form of a questionnaire for new schools. This issue does not apply to the third- or fifth-grade files because only one form of the school administrator questionnaire was used.) An example of system missing is non-participation in the parent interview by a child’s parent. In this case, all questions from the parent interview 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 (e.g., -1, -7, -8, -9, and system missing) may need to be recoded. It is advised that users cross-tabulate all lead questions (e.g., whether the child received child care from a relative) and follow-up questions (e.g., hours of child care from a relative) before proceeding with any recodes or use of the data.

Missing values for composite variables were coded using the same general coding rules as those used for other variables. If a particular composite was inappropriate for a given household—as the variable P6MOMID 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, except in the calculations of the height, weight, and body mass index (BMI) composites for spring-fifth grade.[1]

The ECLS-K Fifth-Grade Restricted-Use Data File is provided on a CD-ROM and is accessible through an ECB that 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 file for analysis. The child data file on the ECB is referred to as a “catalog.” Instructions for using the CD-ROM and ECB are provided in chapter 8.

7.4 Variable Naming Conventions

Variables were named according to the data source (e.g., parent interview, teacher questionnaire) and the data collection point. (A number is used to indicate in which round of data collection the variable was obtained, as follows: 6 for spring-fifth grade, 5 for spring-third grade, 4 for spring-first grade, 3 for fall-first grade, 2 for spring-kindergarten, and 1 for fall-kindergarten. This numbering system is used for all variables except those beginning with “W.” For those variables, 5 indicates fifth grade, 3 third grade, 1 first grade, and K kindergarten.) These variable names are used consistently throughout the catalog. The prefixes listed here are in two categories: (1) fifth-grade variables, and (2) cross-sectional and cross-round longitudinal weights (exhibit 7-2). In general, variable names start with the prefixes listed in exhibit 7-2. For a discussion of the weights, see section 4.7 for cross-sectional weights and section 9.3 for longitudinal weights.

[1] Children’s height and weight measurements were each taken twice to prevent error and provide an accurate reading. Children’s BMI was calculated based on height and weight. The rules for using “Don’t Know” and “Not Ascertained” codes for these values was as follows. If both the first and second measurement of height in the child assessment were coded as -8 (Don’t Know), then the height composite was coded as -8 (Don’t Know). If both the first and second measurements of weight were coded as -8 (Don’t Know), the weight composite was coded as -8 (Don’t Know). If either the height or weight composites were coded as not ascertained (-9), the BMI composite was coded as not ascertained (-9). If neither the height nor weight composites were coded as not ascertained, and either the height or weight composite was coded as -8 (Don’t Know), then the BMI composite was coded as -8 (Don’t Know).

Exhibit 7-2. Prefixes for fifth-grade variables and cross-sectional and cross-round longitudinal weights: School year 2003–04

Category	Description
Fifth-grade variables	
C6	Data/scores collected/derived from spring-fifth grade direct child assessment and spring-fifth grade weight variables
D6	Data collected from spring-fifth grade special education teacher questionnaire A
E6	Data collected from spring-fifth grade special education teacher questionnaire B
F6	Data from spring-fifth grade Field Management System (FMS)
G6	Data collected/derived from spring-fifth grade reading teacher child-level questionnaire
IF	Imputation flags
J6	Data collected/derived from spring-fifth grade teacher questionnaire
K6	Data collected/derived from spring-fifth grade school facilities checklist
M6	Data collected/derived from spring-fifth grade mathematics teacher child-level questionnaire
N6	Data collected/derived from spring-fifth grade science teacher child-level questionnaire
P6	Data/scores collected/derived from spring-fifth grade parent interview
R6	Derived child demographic or child status variables for spring-fifth grade
S6	Data collected/derived from spring-fifth grade school administrator questionnaire
U6	Data collected/derived from spring-fifth grade student record abstract
W5	Fifth-grade (cross-round) parent composite variables
Cross-Sectional and Cross-Round Longitudinal Weights	
C6C	Child-level panel weight variable from spring-fifth grade
C6P	Child-level panel weight for parent data from spring-fifth grade
C6CPTR	Child-level panel weight for combined parent, child, and teacher data from spring-fifth grade
C6CPTM	Child-level panel weight for combined parent, child, and teacher data from spring-fifth grade, if using data from mathematics teacher
C6CPTS	Child-level panel weight for combined parent, child, and teacher data from spring-fifth grade, if using data from science teacher
C56C	Child-level panel weight variable from spring-third grade and spring-fifth grade

See note at end of exhibit.

Exhibit 7-2. Prefixes for fifth-grade variables and cross-sectional and cross-round longitudinal weights: School year 2003–04—Continued

Category	Description
Cross-Sectional and Cross-Round Longitudinal Weights —Continued	
C56P	Child-level panel weights for parent data from spring-third grade and spring-fifth grade
C456C	Child-level panel weight variable from spring-first grade, spring-third grade, and spring-fifth grade
C456P	Child-level panel weights for parent data from spring-first grade, spring-third grade, and spring-fifth grade
C2_6FC	Child-level panel weight variable from spring-kindergarten, spring-first grade, spring-third grade, and spring-fifth grade
C2_6FP	Child-level panel weights for parent data from spring-kindergarten, spring-first grade, spring-third grade, and spring-fifth grade
C1_6FC	Child-level panel weight variable from fall-kindergarten, spring-kindergarten, spring-first grade, spring-third grade, and spring-fifth grade
C1_6FP	Child-level panel weights for parent data from fall-kindergarten, spring-kindergarten, spring-first grade, spring-third grade, and spring-fifth grade
C1_6SC	Child-level panel weight variable from fall-kindergarten, spring-kindergarten, fall-first grade, spring-first grade, spring-third grade, and spring-fifth grade
C1_6SP	Child-level panel weights for parent data from fall-kindergarten, spring-kindergarten, fall-first grade, spring-first grade, spring-third grade, and spring-fifth grade

SOURCE: U.S. Department of Education, National Center for Education Statistics, Early Childhood Longitudinal Study, Kindergarten Class of 1998–99 (ECLS-K), fall 1998, spring 1999, fall 1999, spring 2000, spring 2002, and spring 2004.

A few exceptions that do not follow the prefix convention below are as follows:[2]

- The identifiers CHILDID, PARENTID, and S6_ID.
- The composite T6GLVL. This variable indicates the grade level of the child.
- The composite variable R6R5SCHG. This variable indicates change in school between spring-third grade and spring-fifth grade. Source variables and other details for this and all other composite variables can be found in table 7-15.

[2] It should be noted that in past rounds derived child demographic variables for gender, race/ethnicity, and date of birth (GENDER, RACE, DOBMM, DOBDD, and DOBYY) in the kindergarten and first grade files did not follow the prefix conventions above because they combined information across data collection points and/or several sources. In spring-third and spring-fifth grades, these same demographic variables begin with the prefix R5 (e.g., R5RACE) for spring-third grade and R6 (e.g., R6RACE) for spring-fifth grade. This was done because reports of these variables from parent data were prioritized over other sources starting in spring-third grade and a prefix change was used to indicate the difference to users.

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. Variables based on the child assessment include height, weight, and BMI. Variables based on the teacher data include reading, math, and science class sizes, percentage of limited-English-proficient children in the class, and student grade level. Variables constructed from the school data include the percentage of minority students, school type, and school instructional level. Variables constructed from the parent interview data include parent identifiers, parent demographics, household composition, household income, and poverty, child care, and child demographics. Certain composites were created using data from the Field Management System (FMS).

Table 7-15 lists all the composite variables for the fifth grade. All basic child demographic variables are presented first. Child care variables follow the demographics and then household composition. Imputed variables are listed next, followed by demographics for parents (resident father and mother characteristics are followed by characteristics of nonresident biological parents and nonresident adoptive parents). Teacher, classroom, and school variables are listed last. Once the user identifies the composites of interest, he or she can refer to exhibit 8-8 for instructions on accessing the variables from the ECB.

7.5.1 Child Composite Variables

There are many child-level composite variables on the child catalog. Table 7-15 describes all of the composites. Some of these variables are described in further detail here.

7.5.1.1 Child's Age at Assessment (R6AGE)

The child's age was calculated by determining the number of days between the date when the child completed the ECLS-K direct child assessment and the child's date of birth (R6DOBMO, R6DOBDA, R6DOBYR). The total number of days was then divided by 30 to calculate the age in months. The child assessment date was tested for the appropriate range (March to July 2004). If the assessment date fell outside these ranges, the modal assessment date for the child's school was used.

It should be noted that the date of assessment used for R6AGE may be different from the set of assessment dates and times incorporated into methodological variables that are described further in section 7.6. These variables were not edited like those for R6AGE and are text variables that note both date and time.

7.5.1.2 Gender (R6GENDER)

The fifth-grade gender composite was taken from the third-grade gender composite, if it was not missing. The third-grade gender composite was derived using the gender indicated in the parent interview (INQ.016), child report (AIQ.050), and the FMS. Because of the discrepancies found in the third grade of reports of a child's sex from different sources, the most frequently reported gender was used for the child. If there were an equal number of reports for male and female from these sources, the following hierarchy of rules was used: if the data were from the parent interview in previous rounds, then the third-grade gender composite, R5GENDER, was equal to gender from that parent data. Otherwise, gender was updated from the third-grade parent interview question. If the parent interview data were missing, gender was updated from child report. Otherwise, the third-grade gender composite was equal to the composite GENDER from a previous round (because GENDER in previous rounds incorporated the FMS, this last step meant that the FMS was used as the final source of data).

If the third-grade gender composite was missing, R6GENDER was decided based on the most frequently reported gender from all sources of data, across all rounds of data collection. (The composite variable for R6GENDER is on the file but not the source variables). For most of the cases the data were collected in the base year. Gender was not asked in the fifth-grade parent interview.

7.5.1.3 Child's Date of Birth (R6DOBYY, R6DOBMM, and R6DOBDD)

In the fifth grade, the child's date of birth was derived from the third-grade date of birth composites, if they were not missing. The third-grade date of birth composites were derived from one of three sources: the parent report (CHILDDOB), the child report (AIQ.040), or the FMS. If the child's date of birth had been reported in a parent interview from a previous round, that value was used. Otherwise, the value from the third-grade parent interview was used. If those data were not available or outside the criteria for inclusion (June 1, 1990 to March 31, 1995), the date of birth from the child interview was used. Finally, if the child report was not available or outside the criteria for inclusion, the FMS value was used. If the date of birth given was before June 1, 1990, or after March 31, 1995, the data were excluded from the third-grade composite.

It should be noted that in the kindergarten and first grade files, the child date of birth composites (DOBYY, DOBMM, and DOBDD) were created using two rather than three sources of data. The two sources were parent interview data and, in cases in which the parent interview data did not exist or were outside reasonable boundaries, FMS data. In spring-third grade, a third source—the child—was added and used in the creation of the third-grade composite.

If the third-grade composite was missing, the fifth-grade composite for date of birth was taken from a previous parent interview. Otherwise, date of birth was taken from the FMS.

7.5.1.4 Race/Ethnicity (W5AMERIN, W5ASIAN, W5PACISL, W5BLACK, W5WHITE, W5HISP, W5MT1RAC, W5RACETH, and R6RACE)

In spring-fifth grade, the race of the focal child was no longer collected in the parent interview; thus, race information is based on information collected in previous parent interviews and the FMS. The composites for the child's race/ethnicity are presented in the ECLS-K files in three ways: (1) as dichotomous variables for each race/ethnicity category (W5AMERIN, W5ASIAN, W5PACISL W5BLACK, W5WHITE, W5HISP, W5MT1RAC) from the parent interview data; (2) as a single race/ethnicity composite taken from the parent interview data (W3RACETH); and (3) as a race/ethnicity composite taken from either the parent data or the FMS, with FMS data used only if parent data were missing (R6RACE).

Respondents were allowed to indicate that their child belonged to more than one of the five race categories (White, Black or African American, American Indian or Alaska Native, Asian, Native Hawaiian or 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 the five specified race groups. In addition, one more dichotomous variable was created for those who had simply indicated that their child was biracial or multiracial without specifying a race. The retention of the dichotomous variables on the file allows users to create different composites as needed.

Data were collected on ethnicity as well. Specifically, respondents were asked whether or not their child was Hispanic. Using the six race dichotomous variables and the Hispanic ethnicity variable (e.g., from spring-third grade P5HSP_1 to P5HSP_25, depending on household size), the race/ethnicity composite variables for the child (W5RACETH and R6RACE) were created. The categories were: White, non-Hispanic; Black or African American, non-Hispanic; Hispanic, race specified; Hispanic, no race specified; Asian; Native Hawaiian or other Pacific Islander; American Indian or Alaska Native, and more than one race specified, non-Hispanic. The child composites W5RACETH (race/ethnicity) and R6RACE (race/ethnicity) both share these categories; however, FMS data were used to fill in missing parent report data for the variable R6RACE and only parent report data were used for the variable W3RACETH. A child was classified as Hispanic if a respondent indicated the child's ethnicity was Hispanic regardless of whether a race was identified and what that race was.

For W5RACETH, if the child's race/ethnicity information was available from the parent interview composite in a prior data collection (e.g., W3RACETH, W1RACETH, WKRACETH), the value from the most recent year composite was used and copied forward.[3] If the data were missing for a child from one of these composites, W5RACETH was -9 (Not Ascertained).

For R6RACE, responses from the parent interview composite from third grade (R5RACE) were copied forward. If the third-grade composite, R5RACE, was missing, then the race variable based on parent interview data in the first grade was used (W1RACETH). If the first-grade composite was missing, the race variable based on parent interview data in kindergarten was used (WKRACETH). If the parent interview data were missing, then FMS data from a previous round were used. If previous round FMS data were missing, then FMS data on race from the fifth grade were used.

It should be noted that for both fifth- and third-grade variables R6RACE and R5RACE, previous parent interviews were prioritized over the FMS. This is different from the method used to derive the variable RACE in the first grade. In the first grade, the composite RACE was copied forward from previous rounds and FMS data were used if parent reports were not available. Because parent

[3] A number of respondents, both in this and in prior rounds, gave some variant of "biracial" as the other-specify response to child race. In previous rounds, these responses had been considered to be uncodeable, and the relevant children were given a value of –9 (not ascertained) for WKRACETH and W1RACETH. In spring-third and spring-fifth grades, these responses were treated as multiracial, and the relevant children were given a value of 8 (multiracial) for W3RACETH.

reports were expected to be more accurate than school records, if new information about race was obtained in the third-grade parent interview it was used rather than previous information obtained from the FMS. Therefore, the fifth- and third-grade variables R6RACE and R5RACE are different from RACE in previous rounds for a minority of cases.

7.5.1.5 Child's Height (C6HEIGHT)

To obtain good measurements, each child's height was measured twice. For the height composite C6HEIGHT, if the two height values from the instrument (i.e., C6HGT1 and C6HGT2 for spring-fifth grade) were less than two inches apart, the average of the two height values was computed and used as the composite value. Otherwise, the value that was closest to 57 inches, the median height for 11-year-olds as developed by the National Center for Health Statistics (NCHS) in collaboration with the National Center for Chronic Disease Prevention and Health Prevention (NCCDPHP), was used as the composite value.

7.5.1.6 Child's Weight (C6WEIGHT)

Each child's weight was also measured twice. For the weight composite (C6WEIGHT), if the two weight values from the instrument (i.e., C6WGT1 and C6WGT2 for spring-fifth grade) were less than 5 pounds apart, the average of the two values was computed and used as the composite value. Otherwise, the value that was closest to 82.0 pounds, the median weight for 11-year-olds as developed by NCHS in collaboration with the NCCDPHP, was used as the composite value.

7.5.1.7 Child's Body Mass Index (C6BMI)

Composite Body Mass Index (BMI; variable name C6BMI) was calculated by multiplying the composite weight in pounds by 703.0696261393 and dividing by the square of the child's composite height in inches.

7.5.1.8 Child's Disability Status (P6DISABL)

A composite variable was created to indicate whether a child had a disability diagnosed by a professional. Questions in the parent interview about disabilities in spring-fifth grade asked about the child's ability to pay attention and learn, overall activity level, overall behavior and relations to adults, overall emotional behavior such as anxiety or depression, ability to communicate, difficulty in hearing and understanding speech, and eyesight. For each disability or behavior, a question was asked about whether a diagnosis of a problem was obtained by a professional (CHQ.050, CHQ.110, CHQ.170, CHQ.210, CHQ.300, CHQ.335, CHQ.360). A question was also asked about receipt of therapy services or participation in a program for children with disabilities (CHQ.520).

The composite variable P6DISABL was coded 1 (Yes) if any of the source variables (CHQ.050, CHQ.110, CHQ.170, CHQ.210, CHQ.335, CHQ.360, CHQ.520) about diagnosis or therapy services were coded 1 (Yes). This was done even if data for some of the source variables were missing. In spring-fifth grade, unlike previous rounds, another source variable used to code P6DISABL was CHQ.300 for vision-related problems. If the source variable for the vision diagnosis (CHQ.300) was coded 1 (Yes) and the follow-up question (CHQ.316) was coded NOT "correctable by glasses" (i.e., either only "improvable with glasses" or "not correctable with glasses"), the composite P6DISABL was coded 1 (Yes). Also, in spring-fifth grade, the composite P6DISABL was coded 1 (Yes) if the child had vision problems such that the child's best eyesight (CHQ.320) allowed him or her to see large print in books, form and/or color of objects but not detail, shadows, lights, or saw no light or had no light perception. If data for all the source variables were missing, the composite was coded -9 (Not Ascertained). Otherwise, P6DISABL was coded 2 to indicate no reported disability.

It should be noted that both the spring-third and spring-fifth grade composites are somewhat different from the composites in previous rounds of the study because questions were added about overall behavior and relations to adults and about emotional behavior such as anxiety or depression. Only diagnosed emotional or behavioral problems were included in the composite. In addition, unlike the disability composite in fall-kindergarten which included a question about children's coordination in using their arms or legs, the disability composites since spring-first grade have not included that question. In addition, the disability composite in spring-fifth grade is different from other years of the study because it excludes children who have a diagnosis, but the diagnosis was that the child had "no problem." It also excludes children with correctable vision. In addition, any answers that indicate, for children who do not have correctable vision, what a child's <u>best</u> eyesight allows him or her to see are also counted as having a disability. The question about what a child's best eyesight allows him or her to see asks if the child can see large print in books; form and/or color of objects, but not detail; whether the child can see shadows and lights; or if the child sees no light or has no light perception. Also, in spring-fifth grade, questions asked if the child ever had a disability rather than whether they had a disability since the last round of data collection as had been done in earlier rounds of the study. Thus, disabilities that were diagnosed before spring-fifth grade are included.

7.5.1.9 Nonparental Care (P6CARNOW)

There are several composite variables on the file that can be used to describe child care arrangements based on information from the parent interview. One of these (P6CARNOW) describes whether the child had any type of nonparental care at the time of the interview. The creation of P6CARNOW was as follows. If the child was receiving care from a relative (CCQ.010), a nonrelative (CCQ.150), or a day care center or before or after school program at a school or in a center (CCQ.260), P6CARNOW was equal to 1 (Yes). Otherwise, if any of the three variables was unknown, P6CARNOW was coded as -9 (Not Ascertained). If the respondent indicated that the child was not currently receiving any of the three types of care (CCQ.010, CCQ.150, and CCQ.260 all equaled 2 [No]), P6CARNOW was coded as 2 (No).

It should be noted that the nonparental care as defined by P6CARNOW does not have to be received on a regular basis. However, for the composite P6HRSNOW (hours per week in child care) described below, if the nonparental care is not regular, the number of hours in care is coded as zero. This is because the child must have a regular arrangement in order for hours per week in care to be reported. Users should be aware of the differences in definitions when comparing P6CARNOW with P6HRSNOW.

7.5.1.10 Hours Per Week in Child Care (P6HRSNOW)

Another set of child care composites indicates the number of hours per week the child spent in child care. P6HRSNOW indicates the total number of hours per week the focal child spent in care at the time of the spring-fifth grade interview. The variable combines hours in child care arrangements in which the child spent the most time with hours from additional regular child care arrangements. It was coded as follows. If the relevant child care receipt variables for relative, nonrelative, and center-based care (CCQ.010, CCQ.150, or CCQ.260) were equal to 2 (No Receipt), or if the indicator for regular receipt of that type of care (CCQ.080, CCQ.180, and CCQ.340) was equal to 2 (No Regular Receipt), the number of hours for that type of care was coded to 0. If the receipt variables or 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.

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 (questions CCQ.060, CCQ.165, and CCQ.325 indicated number of regular providers of each type) was equal to 1, then additional hours were coded to 0. Otherwise, if the number of providers or the number of additional hours (questions CCQ.140, CCQ.250, and CCQ.403 indicated number of hours spent with additional providers) was refused or unknown, then the number of additional hours was coded as -9 (Not Ascertained). Otherwise the number of additional hours was coded to equal the appropriate number of additional hours variables in the instrument (CCQ.140, CCQ.250, or CCQ.403).

This process was followed three times, once each for relative care, nonrelative care, and center-based care. If any of the three primary caregiver hour variables or the three additional hours 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.

It should be noted that in rounds prior to spring-third grade, if the primary care arrangement hours were not missing and the additional hours were missing, the primary caregiver hours were used for the composite. In both spring-third and spring-fifth grades, if any of the primary or additional hours variables were missing, the composite was missing. This change makes the variable represent all types of regular care rather than prioritizing primary arrangements.

Although P6HRSNOW was created almost identically to the same composite variable in kindergarten (P1HRSNOW), with the exception noted above, there was one other difference. In kindergarten, questions were asked about whether the child was ever in a particular type of care. If not, P1HRSNOW was set to 0. Because questions about the child having ever been in a particular type of care were not included after the kindergarten year, they were not part of the composite variable definition for the fifth-, third-, or first-grade variables.

7.5.1.11 Number of Child Care Arrangements (P6NUMNOW)

Another composite variable (P6NUMNOW) was used to indicate the total number of all types of care arrangements the focal child had at the time of the spring-fifth grade parent interview. The variable was created as follows. If any of the child care receipt variables for relative, nonrelative, or center-based care (CCQ.010, CCQ.150, or CCQ.260) was refused, unknown, or missing, then P6NUMNOW was coded as -9 (Not Ascertained). If any of the care receipt variables was equal to 1 (Yes), but its corresponding number of arrangements variable (CCQ.060, CCQ.165, and CCQ.325) was refused, unknown, or missing, then P6NUMNOW was again coded as -9 (Not Ascertained). Otherwise, the number of arrangements indicated in CCQ.060, CCQ.165, and CCQ.325 were summed to obtain the total number of current child care arrangements.

The differences in how missing data are handled for each of the child care composites are important to note when combining variables. For example, because P6NUMNOW requires that the number of child care arrangements be known, it is possible for a child to have P6CARNOW =1 (child was in nonparental care) and have P6NUMNOW be -9 (Not Ascertained).

7.5.1.12 Primary Nonparental Child Care Arrangement (P6PRIMNW)

A composite variable (P6PRIMNW) was created to indicate the primary, nonparental child care arrangement in which the child spent the most hours per week at the time of the spring-fifth grade interview. This variable is for children in a regular care arrangement. 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=Nonrelative care in child's home
- 4=Nonrelative care in another home
- 5=Center-based program
- 6=Two or more programs
- 7=Location of care varies

To obtain the composite, hours were compared for relative care in the child's home (CCQ.090) or in other home (CCQ.070); nonrelative care in child's home (CCQ.190) or in other home (CCQ.170); and center/program care (CCQ.355). First, the composite P6HRSNOW, described earlier, was used to code individuals missing current hours of care (P6HRSNOW=-9) or with no hours of nonparental care (P6HRSNOW=0). Those with missing hours of care were coded as -9 (Not Ascertained); those with no hours of care or no regularly schedule care were coded as 0.

For the remaining cases, if the number of hours of either relative or nonrelative care (given in CCQ.090 and CCQ.190) were higher than all other hours of care, the variable indicating location of care for that type was examined using instrument items CCQ.070 and CCQ.170. If location of care was missing, then P6PRIMNW was coded as -9 (Not Ascertained); if P6PRIMNW was not missing, then P6PRIMNW was coded 1, 2, 3, or 4, depending on the type (relative/nonrelative) and location (child's home/other home) of care. Otherwise, if the number of hours of care in center-based programs (CCQ.355) was higher than for relative or nonrelative care, then P6PRIMNW was coded as 5. If the number of hours of care was equal for two or more types of care, P6PRIMNW was coded as 6. P6PRIMNW was coded as 7 if the location of care varied between two homes.

It should be noted that it is possible to have missing data for the primary child care arrangement (P6PRIMNW), but have information on the number of hours of child care a child has (P6HRSNOW), because there must be information about the location of care in order to have a valid value for P6PRIMNW.

7.5.2 Family and Household Composite Variables

Many composites were created to capture information about the sampled children's family and household characteristics. Several of these are described below. All of the family and household composites are listed and described in table 7-15.

7.5.2.1 Number of Siblings (P6NUMSIB)

The composite P6NUMSIB indicates the total number of siblings (full, step, adoptive, or foster) with whom the child lived in the household (FSQ.160 and FSQ.170). 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 Parent and Household Members' Age (P6LESS18, P6OVER18, P6HDAGE, and P6HMAGE)

There are several composite variables on the file that refer to the ages of adults and children in the household. These are P6LESS18 (total number of people in the household under age 18, including focal child, siblings, and other children), P6OVER18 (total number of people in the household age 18 or older, siblings, and other children), P6HDAGE (age of resident father), and P6HMAGE (age of resident mother). The ages of these persons in the household were collected during the fall of kindergarten in the household matrix. However, in subsequent years of the study, questions about age were not asked for household members who were previously in the household. In the fifth grade, ages were collected for new household members. Otherwise, ages were incremented in spring-third grade based on the round in which the person joined the household, and then updated again in spring-fifth grade by adding two years to the age calculated in spring-third grade. Age changes were made to increase the ages of all household members other than the focal child and twin (the ages of the focal child, and twin, if applicable, were updated based on birth date).

The ages of all household members who were not new to the study in spring-fifth grade (other than the focal child and twin) were increased by spring-fifth grade by the numbers shown in table 7-1. Ages were increased incrementally each round of the study. The numbers in table 7-1 reflect the total number of years added to the first reported age for a household member when the household joined the study. The guidelines for creating these were as follows: (1) half years could not be included, and (2) the same number of years was added for those who entered the study during the same school year. The numbers were made to err on the side of making persons older rather than younger because this would cause fewer problems with range checks and displays in the parent interview if there was a discrepancy between actual age and imputed age.

Table 7-1. Incremented ages of previous household members based on round entered study: School year 2003–04

Round in which household member joined study	Number of years added by spring-fifth grade to first age reported when household joined study
Fall 1998	+6
Spring 1999	+6
Fall 1999	+5
Spring 2000	+5
Spring 2002	+2

SOURCE: U.S. Department of Education, National Center for Education Statistics, Early Childhood Longitudinal Study, Kindergarten Class of 1998–99 (ECLS-K), spring 2004.

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 (P6WORRFD through P6NOMONY) in the spring-fifth grade parent interview. The questions measured a wide range of food insecurity and reduced food intake issues. They were combined into a scale using statistical methods based on the Rasch measurement model. The items and the food security scale based on them have been validated using both ethnographic and statistical methods. For spring-fifth grade, composites were created for Household Food Security scale variables, Children's Food Security scale variables, and Adult Food Security scale variables (for spring-kindergarten and spring-first grade, composites were created only for Household Food Security scale variables; for spring-third grade, composites were created only for Household Food Security and Children's Food Security scale variables). Calculations of the Household Food Security scale variables were carried out in accordance with the standard methods described in *Guide to Measuring Household Food Security, Revised 2000* (U.S. Department of Agriculture 2000). Calculations of the Children's Food Security scale variables 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 2002). Analysis of the ECLS-K data using Rasch methods indicated that use of the standard benchmark household scores was appropriate.

7.5.2.4 Food Security Status: Continuous Measures (P6FSSCAL, P6FSCHSC, and P6FSADSC)

P6FSSCAL is the scale score presentation of the Household Food Security items. 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 (P6FSRAW). Valid values range from 1 to 13, 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. P6FSSCAL 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.

P6FSCHSC is similar to P6FSSCAL but 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 to 13, with higher values indicating more severe food deprivation. The scale score is undefined for households that affirmed no child-referenced items (see discussion of P6FSSCAL above).

P6FSADSC 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. It is a continuous, interval-level measure based on the Rasch measurement model and is appropriate for linear models, such as correlation, regression, or analysis of variance. It is on the standard (logistic-unit) metric described in *Measuring Children's Food Security in U.S. Households, 1995-99* (for households without children). Valid values range from 1 to 12, with higher values indicating more severe food deprivation. The scale score is undefined for households that affirmed no adult-referenced items (see discussion of P6FSSCAL above).

7.5.2.5 Food Security Status: Categorical Measures (P6FSSTAT, P6FSCHST, and P6FSADST)

P6FSSTAT is a categorical measure of Household Food Security status formed by dividing P6FSSCAL into three ordered categories: food secure, food insecure without hunger, food insecure with hunger. In previous rounds, the third category of “food insecure with hunger” was broken into two categories: “food insecure with hunger (moderate)” and “food insecure with hunger (severe).” In spring-fifth grade, these categories have been collapsed into one. P6FSSTAT is appropriate for comparing prevalence rates of food insecurity and hunger across subpopulations and can be used as a categorical variable in associative models. There are few cases in the most severe category, so, for most prevalence reporting purposes, the two categories of food insecure with hunger (moderate and severe) should be collapsed and reported as a single category. 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.

P6FSCHST is 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. This variable is appropriate for comparing prevalence rates of hunger among children across subpopulations. There were few households (n=55, 0.5 percent) that reported hunger among children, so the analytic utility of this variable is limited. 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 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-K study if there are other children in the household.

P6FSADST is a categorical measure of Adults’ Food Security status that identifies households as food secure, food insecure without hunger, or food insecure with hunger among adults. This variable is appropriate for comparing prevalence rates of food insecurity and hunger among adults across subpopulations.

7.5.2.6 Food Security Status: Raw Scores (P6FSRAW, P6FSCHRA, and P6FSADRA)

The Household Food Security raw score, P6FSRAW, is a count of affirmative responses to the 18 items. This is an ordinal-level measure of food insecurity. It is not recommended for direct use in analysis, but can be used to identify categories of food insecurity additional to the categorical measures provided in the NCES data file. The Children's Food Security raw score, P6FSCHRA, is a count of affirmative responses to child-referenced items. Responses to items skipped because of screening are assumed to be negative. Families with no valid responses are coded as missing (-9). P6FSADRA is the adult food security raw score, a simple count of the number of household- and adult-referenced food security items affirmed by the parent. It ranges from 0 to 10.

7.5.2.7 Socioeconomic Status (SES) and Poverty (W5DADSCR, W5MOMSCR, W5SESL, W5SESQ5, W5INCCAT, W5POVRTY)

Socioeconomic status (SES) was computed at the household level using data for the set of parents who completed the parent interview in spring-fifth grade. The SES variable reflects the socioeconomic status of the household at the time of data collection for spring-fifth grade (spring 2004). The components used to create the 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.

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 1980 Census occupational categories covered by the ECLS-K occupation. Table 7-15 provides details on the prestige score values (W5DADSCR, W5MOMSCR).

The variables were collected as follows:

- **Income.** The information about income was collected in spring-fifth grade. 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 7-2.

  Households that were determined to meet the size and income criteria related to poverty shown in table 7-3 were asked to report income to the nearest $1,000. (We call this exact income for simplicity.) 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.

- **Parent's education.** The information about parent's education was collected or updated in spring-fifth grade.

- **Parent's occupation.** The information about parent's occupation was collected or updated in spring-fifth grade.

Table 7-2. Levels of the detailed income range, spring-fifth grade: School year 2003–04

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, Kindergarten Class of 1998–99 (ECLS-K), spring 2004.

Table 7-3. Households asked to report income to the nearest $1,000, spring-fifth grade: School year 2003–04

Household size	Total household income
One	$10,000 or less
Two or three	$15,000 or less
Four	$20,000 or less
Five or six	$25,000 or less
Seven	$30,000 or less
Eight	$35,000 or less
Nine or more	$40,000 or less

SOURCE: U.S. Department of Education, National Center for Education Statistics, Early Childhood Longitudinal Study, Kindergarten Class of 1998–99 (ECLS-K), spring 2004.

Not all parents completed the parent interview; among those who did, not all responded to every question. 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 7-4).

Table 7-4. Missing data for SES source variables, spring-fifth grade: School year 2003–04

Variable	Number missing	Percent
Mother's education	159	1.49
Father's education	160	1.83
Mother's occupation	100	0.90
Father's occupation	166	1.90
Detailed income range	883	8.10

SOURCE: U.S. Department of Education, National Center for Education Statistics, Early Childhood Longitudinal Study, Kindergarten Class of 1998–99 (ECLS-K), spring 2004.

A two-stage procedure was used to impute missing values for each component of the SES composite variable. First, if a parent had completed an interview in the kindergarten-, first-, or third-grade year, missing values for the spring-fifth grade education, occupation, and detailed income range were filled in with values from the previous years. The rationale for this approach was that the best source of data for an individual or a household was the data from a previous year.

This first imputation stage was implemented as follows:

- Education level was brought forward from the most recent previous round. This was done only if the same person was the parent figure both in spring-fifth grade and in the earlier round.

- Occupation was brought forward only if the individual was in the labor force (i.e., was working at a paid job, on vacation from a paid job, or looking for a job). It was also required that the same person be the parent figure both in spring-fifth grade and in the earlier round. NOTE: Prestige scores were not assigned to individuals unless they were in the labor force, regardless of whether they reported an occupation.

- Detailed income category was brought forward from the most recent previous round.

Second, data still missing after this initial step 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 were the best predictors of the variables to be imputed. These relationships had been determined previously by Chi-squared Automatic Interaction Detector (CHAID) analyses of the base year data. Missing values for the education, occupation, and detailed income range variables were imputed by the hot deck method for all households. 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 education, occupation, and income category.

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 interview 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 both parents were present, the order of hot deck imputation was as follows:

- Mother's education;
- Father's education;
- Mother's labor force status;
- Mother's occupation;
- Father's labor force status;
- Father's occupation;
- Detailed income range, where the broad income range was known; and
- Detailed income range, where the broad income range was unknown.

At this point, all of the missing values had been imputed. 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 all 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.

All missing values of the SES components were imputed by the process described above. Tables 7-5 through 7-8 summarize the results.

Table 7-5. Selected statistics on imputed parental education variables, spring-fifth grade: School year 2003–04

SES component	Total missing	Number of values filled from previous rounds	Number of values imputed by hot deck	Number of cases resolved
Mother's education	159	108	51	159
Father's education	160	97	63	160

SOURCE: U.S. Department of Education, National Center for Education Statistics, Early Childhood Longitudinal Study, Kindergarten Class of 1998–99 (ECLS-K), spring 2004.

Table 7-6. Selected statistics on imputed labor force status, spring-fifth grade: School year 2003–04

Labor Force Status	Number of values filled from previous rounds	Number of values imputed by hot deck	Number of cases resolved
Mother			
Total missing			106
In labor force	68	7	75
Not in labor force	26	5	31
Father			
Total missing			88
In labor force	66	13	79
Not in labor force	7	2	9

SOURCE: U.S. Department of Education, National Center for Education Statistics, Early Childhood Longitudinal Study, Kindergarten Class of 1998–99 (ECLS-K), spring 2004.

Table 7-7. Selected statistics on imputed occupation variables, spring fifth grade: School year 2003–04

Occupation	Number of values filled from previous rounds	Number of values imputed by hot deck	Number of cases resolved
Mother			
Total missing			131
Occupation	18	82	100
Not in labor force[1]	26	5	31
Father			
Total missing			175
Occupation	16	150	166
Not in labor force[1]	7	2	9

[1] No occupation was imputed if "not in labor force" was filled from previous rounds or imputed by hot deck.
SOURCE: U.S. Department of Education, National Center for Education Statistics, Early Childhood Longitudinal Study, Kindergarten Class of 1998–99 (ECLS-K), spring 2004.

Table 7-8. Selected statistics on imputed detailed income range, spring-fifth grade: School year 2003–04

SES component	Total missing	Number of values filled from previous rounds		Number of values imputed by hot deck		Number of cases resolved
		Broad income range known	Broad income range unknown	Broad income range known	Broad income range unknown	
Detailed income range	883	786	97	84	13	883

SOURCE: U.S. Department of Education, National Center for Education Statistics, Early Childhood Longitudinal Study, Kindergarten Class of 1998–99 (ECLS-K), spring 2004.

Once the components of the SES variable were imputed, their corresponding *z*-scores or normalized values were computed. 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 the number of nonmissing SES components for the *i*-th household. W5SESL is the continuous variable for the SES composite that ranges from –2.48 to 2.54. 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 multivariate regressions, W5SESL is the variable to use. A categorical SES variable (W5SESQ5) 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 spring-fifth grade parent weights. For categorical analyses, use W5SESQ5 and the parent weight.

Note that for households with only one parent present, not all the components were defined. In these cases, SES was computed averaging the available components.

The imputed detailed income range variable (W5INCCAT) was also used to create a household-level poverty variable (W5POVRTY). Income was compared to Census poverty thresholds for 2003, which vary by household size. Table 7-9 shows the detailed income categories used in the ECLS-K parent interview 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 Census.[4] Households whose income fell below the appropriate threshold were classified as poor (see table 7-9). For example, if a household contained two members, and the household income was lower than $12,015, then the household was considered to be below the poverty threshold.

Table 7-9. ECLS-K and Census poverty thresholds for 2003: School year 2003–04

Household size	ECLS-K income categories	Census weighted average thresholds for 2003[1]
2	Less than or equal to $15,000	$12,015
3	Less than or equal to $15,000	$14,680
4	Less than or equal to $20,000	$18,810
5	Less than or equal to $25,000	$22,245
6	Less than or equal to $25,000	$25,122
7	Less than or equal to $30,000	$28,544
8	Less than or equal to $35,000	$31,589
9+	Less than or equal to $40,000	$37,656

[1] U.S. Census Bureau, Current Population Survey. http://www.census.gov/hhes/poverty/threshld/thresh03.html.
SOURCE: U.S. Department of Education, National Center for Education Statistics, Early Childhood Longitudinal Study, Kindergarten Class of 1998–99 (ECLS-K), spring 2004.

7.5.2.8 Parent Education (W5PARED, W5DADED, and W5MOMED)

There are three parent education composites on the file. These are W5PARED (the highest level of education for the child's parents or nonparent guardians who reside in the household), W5DADED (father's highest level of education), and W5MOMED (mother's highest level of education). The variables include both parent (birth, adoptive, step, and foster) and nonparent guardians. For example, if the child had no parents but had a guardian, the education of the guardian and his or her spouse were used in the creation of the composites if the guardian was specified as such in the relationship variable or if the guardian was the respondent/respondent's spouse and there were no other parent figures in the household.

[4] The ECLS-K provides an approximate, but not exact measure of poverty. Income category thresholds used in the parent questionnaire are similar, but not identical to those from weighted Census averages.

In spring-fifth grade, parent education level was updated from the spring-third grade composite variable value for education if it was a household that had been part of the spring-third grade round of the study. Respondents were asked if they or their corresponding parent figures, if applicable, had completed any additional grades of school or had received any diplomas or degrees (PEQ.010). If so, PEQ.020 asked what grade the parent had completed or what degree had been received. Another question, PEQ.021, verified whether the parent had a high school diploma or its equivalent, such as a GED. If there was no education information to update from spring-third grade, respondents were asked for their highest education level in PEQ.020. If this education level was less than the education level reported in a previous round, the higher education level was kept for the spring-fifth grade composite.

If both parents/guardians resided in the household, W5PARED was the highest value for education level from either the mother/guardian in W5MOMED or the father/guardian in W5DADED. If the household only had one parent or guardian, then W5PARED was equal to either W5MOMED or W5DADED depending on which parent or guardian resided with the child. If the education data for either of the parents were missing[5] it was imputed, and the composite W5PARED was created based on both the reported and imputed data.

7.5.2.9 Parent Race/Ethnicity (P6HDRACE and P6HMRACE)

The composites for race/ethnicity for the parents were calculated in the same way as those for the child, except that there is not a variable that supplements parent-reported race/ethnicity with FMS data similar to the variable R6RACE for children. All data on parent race/ethnicity are derived from the parent interview. Race/ethnicity for parents is presented in the spring-fifth grade data file as a categorical race/ethnicity composite (for the father/male guardian it is P6HDRACE, and for the mother/female guardian it is P6HMRACE).

Respondents were allowed to indicate that they belonged to more than one of the five race categories (White, Black or African American, American Indian or Alaska Native, Asian, Native Hawaiian or other Pacific Islander). From these responses, a series of five dichotomous race variables were coded that indicated separately whether the respondent belonged to each of the five specified race groups. In addition, one more variable was coded for those who had simply indicated that they were

[5] Missing data were due to "refused" or "don't know" answers from respondents.

biracial or multiracial without specifying the race.[6] The dichotomous codes for each of the race variables are not provided on the spring-fifth grade file, but the composite derived from the responses is provided.

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 her boyfriend, if he 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.

7.5.3 Teacher Composite Variables

Details about how two of the teacher composites, child grade level and class type, were created are provided here. All of the teacher composites are listed and described in table 7-15.

7.5.3.1 Grade-Level Composite (T6GLVL)

To create the grade-level composite (T6GLVL), five possible sources of information were used: (1) reading teacher questionnaire (Q1 G6GRENRL for grade level); (2) special education teacher part B (Q2 E6ENRGR for grade level); (3) child assessment introductory section (AIQ.030 C6INGRAD); (4) child assessment closing section (ACQ.005 C6FIFTH and ACQ.010 C6GRADE, completed by interviewer), and (5) FMS information about grade level.

If conflicts existed among these five sources, the grade level indicated by the majority of the nonmissing sources was used for T6GLVL. When there were five, four, or three sources of information and three were in agreement, the grade level indicated by the three sources was taken. When there were four sources of information and only two were in agreement, the grade level indicated by the two sources in agreement was taken. When there were three sources of information and two were in agreement, the grade level indicated by the two sources in agreement was taken. When there were four or five sources of

[6] In a previous round of the study, respondents who reported they were "biracial" in the "other" category were classified as "uncodeable." These responses were reclassified as "multiracial" in spring-fifth grade.

data, and two sources indicated one option and the other two indicated another option, the grade indicated in a particular source was selected, according to the hierarchy presented below.

1. Classroom reading teacher, G6GRENRL
2. Special education teacher, E6ENRGR
3. Assessment introduction, C6INGRAD
4. Assessment closing, C6FIFTH and C6GRADE
5. FMS

In establishing this hierarchy, it was assumed that teachers have the best knowledge and that school records (on which the FMS are based) were more apt to be in error. It was also assumed that children are reliable reporters of their own grade level, so their reports are prioritized over the FMS. When equal numbers of sources are in conflict (1 vs. 1) or (2 vs. 2) or (1 vs. 1 vs. 1), the decision was made by using the information from the source highest on the list above.

One exception to this hierarchy was made. Because the FMS and AIQ grade-level information did not allow for ungraded classrooms, the FMS and AIQ information were not considered in any case in which at least one source indicated an "ungraded" classroom.

It should be noted that in spring-first grade, there was information about grade level from the student record abstract; however, there were no grade level questions in the child assessment at that time. In both spring-third and spring-fifth grades, grade level was not asked in the student record abstract, but was included as part of the child assessment instead.

It should also be noted that in spring-fifth grade, data from the reading teacher is prioritized. Although there are data about individual children in the reading, math, and science teacher questionnaires, the reading questionnaire is the only one that asks about the grade level of the child. In previous rounds of the study, there was one teacher sampled for all subjects and grade level was provided by that teacher.

7.5.3.2 Class Size (G6CLSZ, M6CLSZ, N6CLSZ)

In spring-fifth grade, there are composites for class sizes of reading (G6CLSZ), mathematics (M6CLSZ), and science (N6CLSZ) classes, all of them similarly created. For example, for mathematics the totals for race/ethnicity (Q6, M6TOTRA) and sex (Q7, M6TOTGEN) were compared. If the two totals matched, the total shown by the two matching sources was used. If not, the total for the composite was set using, in order of priority, the sex and race/ethnicity total. Otherwise, the class size variable was coded as -9 (Not Ascertained). Some class sizes appear high and are likely due to teachers erroneously reporting the class sizes of all of their classes, rather than only the subject child's class.

It should be noted that the class size composites in both third grade (A5CLSZ) and fifth grade (G6CLSZ, M6CLSZ, and S6CLSZ) were used as the denominator for the composite variables that describe percent minority, percent Hispanic, percent Black, and percent Limited English Proficiency (LEP) in the classroom. In fall-kindergarten and spring-first grade, the total class size used in the calculation of these variables was based on the total number of children in the question about numbers of children by race/ethnicity. It should also be noted that in spring-third grade, a total of the class by age, in addition to the totals by sex and race/ethnicity, was used in the calculation of class size. In spring-fifth grade, teachers were no longer asked to provide age information about students in the class.

7.5.4 School and Class Composite Variables

Variables on school and class characteristics were constructed from the teacher and school data and the sample frame. Details on how some of the variables were created follow.

7.5.4.1 School Type (S6SCTYP)

In spring-fifth grade, S6SCTYP was created as follows. Questions 5 (S6PUBLIC) (whether school is public) and 7 (S6CATHOL, S6OTHREL) (type of private school) from the School Administrator Questionnaire, along with school sample frame data, were used to create the school type composite variable. If the response to question 5 (Is this a public school?) was "Yes," then S65SCTYP was coded "public." If the response to question 7.a. (S6CATHOL) (Is your school a Catholic school?) was "Yes," then the school was coded as "Catholic." Otherwise, if the response to question 7.b.

(S6OTHREL) (Is your school private with another religious affiliation?) was “Yes,” then S6SCTYP was coded as “private, other religious.” Otherwise, because the skip pattern to question 7 was used only if the school was private, if the response to question 7.c. (S6NAISKL, private school accredited by NAIS), question 7.d. (S6OTHPRI, other private), question 7.e. (S6PVTSPD, special education school-primarily serves children with disabilities), or question 7.f. (S6PVTEAR, an early childhood center-school or center includes preschool and/or early elementary grades) was “Yes,” then S6SCTYP was coded as “other private.”

If S6SCTYP could not be coded from the School Administrator Questionnaire, reports of school type from the same school in previous rounds were used (in spring-third grade, school type was taken from a questionnaire called the school fact sheet; and the variable name was S6SCTYP; in previous rounds, school type had been asked in the school administrator questionnaire and the variable names were S4SCTYP, S3SCTYP, S2KSCTYP, and CS_TYPE2). If those sources were unavailable, a variable from the school master file (taken from the 2001–2002 Private School Survey/2003-2004 Common Core of Data frame) was used to code S6SCTYP. If S6SCTYP could not be coded, S6SCTYP was coded as -9 (Not Ascertained). If the child was schooled at home, the composite was coded as -1 (Not Applicable).

7.5.4.2 Public or Private School (S6PUPRI)

S6PUPRI is a less detailed version of school type (with only two categories—public and private) and is derived from the school type composite S6SCTYP described above. In spring-fifth grade, it was created as follows. If S6SCTYP was 4 (public), then S6PUPRI was coded as public (1). If S6SCTYP was 1–3 (Catholic, other religious, other private) then S6PUPRI was coded as private (2). If S6SCTYP was coded as Not Ascertained (-9), then S6PUPRI was -9 (Not Ascertained). If S6SCTYP was coded “Not Applicable,” then S6PUPRI was coded “Not Applicable.”

7.5.4.3 School and Grade-Level Enrollment (S6ENRLS, S6ENRL5)

There are two composite enrollment variables on the fifth-grade file: total school enrollment (S6ENRLS) and fifth-grade enrollment (S6ENRL5). Total school enrollment was created using the school enrollment variable from the school administrator questionnaire (S6ANUMCH). If this variable was missing, data for private schools were taken from the 2001–2002 Private School Survey (PSS) and data for public schools were taken from the 2003-2004 CCD public school universe. If these were also missing, the variable was coded -9 (Not Ascertained). If the child was schooled at home, the composites were coded -1 (Not Applicable).

Fifth-grade enrollment was not obtained during data collection. The fifth-grade enrollment data for private schools came from the 2001–2002 PSS data. The enrollment data for public schools came from the 2003-2004 CCD public school universe data.

7.5.4.4 Percent Minority Students in the School (S6MINOR)

The composite variable S6MINOR indicates the percentage of minority students in a school in spring-fifth grade. The composite is based on a question in the school administrator questionnaire (Q8) that was used to ask about the number or percentage of students in the following categories: Hispanic, regardless of race; Black, not of Hispanic origin; White, not of Hispanic origin; Asian or Pacific Islander; American Indian or Native Alaskan; and other. The composite was based on the sum of percentages for all categories except White, not of Hispanic origin. In some cases, the composite could not be obtained from the data because of missing data or errors. If the composite could not be derived from the data, percent minority was obtained from the CCD (for public schools) or the PSS (for private schools). If these data were missing, the composite was coded -9 (Not Ascertained). If the child was schooled at home, the composite was coded as -1 (Not Applicable).

In spring-fifth grade, as in spring-third and spring-first grades, school administrators were allowed to report their answers to the student racial composition questions as either numbers or percents, whereas in spring-kindergarten they were asked to report those answers as percents. All answers recorded as numbers in spring-fifth grade were converted to percentages for the composite variable. The sum of the answers across all categories was allowed to add within +/- 5 percent of the reported total. In a few cases, this produced answers slightly over 100 percent. These were topcoded to 100 percent.

A flag for each individual race/ethnicity variable indicates whether the answer was reported as a number or a percent.[7] Because the composite is calculated as a percent, these flags will not be needed by users unless the analyst is interested in examining how answers were reported. If the flags (S6ASNFL, S6HSPFL, S6BLKFL, S6WHTFL, S6INDFL, and S6OTHFL) were equal to 1 for each of the race variables S6ASNPCT, S6HISPPCT, S6BLKPCT, S6WHTPCT, S6INDPCT, S6OTHPCT, these 6 race/ethnicity variables were reported by the respondent as percentages.

It should be noted that the spring-fifth grade composite was created in the same way as the composites for spring-third and first grades. However, the composites from first grade forward are slightly different from the one used in spring-kindergarten (S2MINOR) because the school administrator questionnaire item that asked about the percent of minority students in the school had different response options. In spring-kindergarten, the percent of minority students was derived from answers to the school administrator questionnaire by determining the percentage of children who were either of Hispanic or Latino origin (question 14) and the percentage of children who were American Indian or Alaska Native, Asian, black or African American, or Native Hawaiian or Other Pacific Islander (question 15) to create the percent minority composite. In spring-first, third, and fifth grades, ethnicity and race were included in the same question.

7.5.4.5 School Instructional Level (S6SCLVL)

The purpose of this composite is to classify schools based on the highest grade taught in the school. This composite is taken in spring-fifth grade from the School Administrator Questionnaire (Q4, S6PRKNDR, S6KINDER, S6GRADE1, S6SECOND, S6THIRD, S6FOURTH, S6FIFTH, S6SIXTH, S67TH, S68TH, S6NINTH, S6TENTH, S611TH, S612TH). The highest grade level circled on the form was determined, and the grade level was classified accordingly. If data were missing, data were used from the school master file (based on the 2001–2002 PSS and the 2003-2004 CCD) to fill in instructional level. If school master file data were unavailable for a particular school, data from previous school administrator questionnaires from spring-third grade, spring-first grade, or spring-kindergarten schools (S5SCVL,

[7] There were also other questions in the school administrator questionnaire that allowed for answers to be recorded as either a number or percent. The flags for these variables are S6ADAFLG (average daily attendance reported as number/percent), S6ASNFLG (question about Asian or Pacific Islander teachers reported as number or percent), S6HSPFLG (question about Hispanic teachers reported as number or percent), S6BLKFLG (question about black teachers reported as number or percent), S6WHTFLG (question about white teachers reported as number or percent), S6INDFLG (question about American Indian or Native Alaskan teachers reported as number or percent), and S6OTHFLG (question about teachers of other races reported as number or percent). In all cases, the final variables related to these flags are reported as percentages, but the flags indicate how the answers were originally recorded by respondents.

S4SCLVL, and S2KSCLVL) were used to determine instructional level. If those sources were also not available, S6SCLVL was coded as -9 (Not Ascertained). If the child was schooled at home, the composite was coded as -1 (Not Applicable).

In spring-third grade, this composite was taken from the school fact sheet. In other rounds of the study, it was taken from the school administrator questionnaire. Also, in spring-kindergarten and spring-first grade, if the question about grade levels in the school was left blank, another question from the school administrator questionnaire about grade levels that participated in special programs was used. If the respondent did not answer either of these questions, then school sample frame data were used to determine the value for the composite in previous rounds.

7.5.4.6 School Lunch Composites (S6FLCH_I, S6RLCH_I)

The school lunch composites were computed at the school level for the set of public schools that have at least one child or parent respondent (i.e., the child has nonzero child weight, C6CW0, or nonzero child-level parent weight, C6PW0) in spring-fifth grade. There are two school lunch composites as follows:

- Percent of children eligible for free school lunch; and
- Percent of children eligible for reduced-price lunch.

The data that are used to create the school lunch composites were collected in the school administrator questionnaire (SAQ). Specifically, school principals were asked to report on the total enrollment in the school (S6ANUMCH), the number of children in the school who were eligible for free school lunch (S6ELILNC), and the number of children who were eligible for reduced-price school lunch (S6ELIRED). The percent of children eligible for free school lunch is computed as the ratio of S6ELILNC over S6ANUMCH. Likewise, the percent of children eligible for reduced-price school lunch is the ratio of S6ELIRED over S6ANUMCH.

Not all schools completed the SAQ, and among those who did, not all responded to all three questions needed to compute the school lunch composites. Therefore, there were missing values for some of the components of the school lunch composite variables. Prior to fifth grade, if the source variables have missing value, then the composites were filled in with values computed using the most recent CCD if they are not missing from the CCD, or left missing if they are missing from the CCD. In grade five, the composites were computed as they had been in the past, but if they had missing values, they were imputed. The source variables, however, were not imputed. Table 7-10 shows the level of missing data for the school lunch composite variables among the 2,008 public schools that have child or parent respondents in the fifth grade of the ECLS-K.

Table 7-10. Public schools with missing values of the school lunch composites, spring-fifth grade: School year 2003–04

School lunch composite	Number missing	Percent missing
Free lunch	691	34.4
Reduced-price lunch	712	35.5

SOURCE: U.S. Department of Education, National Center for Education Statistics, Early Childhood Longitudinal Study, Kindergarten Class of 1998–99 (ECLS-K), spring 2004.

A two-stage procedure was used to impute missing values for each school lunch composite variable. First, if a school had nonmissing value of the school lunch composite in the kindergarten, first-grade, or third-grade year, missing values for the spring-fifth grade school lunch composites were filled in with values from the previous years. The rationale for this approach was that the best source of data for a school was the data from a previous year.

Second, data still missing after this initial step were imputed using a hot deck methodology. Imputation cells were created using the Title I status of the school and the school latitude and longitude. The Title I status is a derived variable using the data on whether the school received Title I funds (S6TT1) and whether Title I funds were targeted or used school-wide (S6TT1TA), both collected in the SAQ. If these two variables have missing values for fifth grade, then data the most recent available data (from third grade or first grade or kindergarten) were used. If these data were missing from the SAQ for all rounds, then the information from the most recent CCD (2002-03) was used.

The resolution of cases having missing data is shown for each school lunch composite in table 7-11 (for schools) and table 7-12 (for children). Schools that were imputed by hot deck are generally transfer schools with few sample children in those schools. This is reflected in tables 7-11 and 7-12 where the percent of children with hot deck values of the school composites is much smaller than the percent of schools with hot deck values of the school composites.

Table 7-11. Imputation of school lunch composites at the school level, spring-fifth grade: School year 2003–04

School lunch composite	Number missing	Values from previous round		Imputed by Hot deck	
		n	Percent	n	Percent
Free lunch	691	256	37.0	435	63.0
Reduced-price lunch	712	265	37.2	447	62.8

SOURCE: U.S. Department of Education, National Center for Education Statistics, Early Childhood Longitudinal Study, Kindergarten Class of 1998–99 (ECLS-K), spring 2004.

Table 7-12. Results of imputation of school lunch composites at the child level, spring-fifth grade: School year 2003–04

School lunch composite	Number missing	Values from previous round		Imputed by Hot deck	
		n	Percent	n	Percent
Free lunch	2,545	1,777	69.8	768	30.2
Reduced-price lunch	2,601	1,819	69.9	782	30.1

SOURCE: U.S. Department of Education, National Center for Education Statistics, Early Childhood Longitudinal Study, Kindergarten Class of 1998–99 (ECLS-K), spring 2004.

Since children were designated as eligible for either free lunch or reduced-price lunch but not for both services, the two school lunch composites should sum to no more than 100 percent. A very small number of schools (less than 2 percent) had imputed values of the two school lunch composites summing to more than 100 percent. These values came from two sources: (1) from values reported by the school in another year, or (2) from the hot-deck imputation. The reporting error has been present in all rounds of the ECLS-K, and the decision was to keep the reported values in the data file. If the erroneous values came from the hot-deck imputation, then they were corrected so that the two school lunch composites do not add to more than 100 percent.

7.5.4.7 School Year Start and End Dates (S6SCHBDD, S6SCHBMM, S6SCHBYY, S6SCHEDD, S6SCHEMM, S6SCHEYY)

The composite for school year start and end dates was taken from the school administrator questionnaire (Q10, S6SYRSMM, S6SYRSDD, S6SYRSYY, S6SYREMM, S6SYREDD, S6SYREYY). If those data were missing, the values were taken from the FMS. In spring-fifth grade, the answers for the starting date, year (S6SCHBYY) and the ending date, year (S6SCHEYY) had already been filled in for the school administrator when he or she received them. For this reason, the starting date, year was always 2003 and the ending date year was always 2004. This was done to prevent errors.

S6SCHBDD	S6 School Year Starting Date, Day
S6SCHBMM	S6 School Year Starting Date, Month
S6SCHBYY	S6 School Year Starting Date, Year
S6SCHEDD	S6 School Year Ending Date, Day
S6SCHEMM	S6 School Year Ending Date, Month
S6SCHEYY	S6 School Year Ending Date, Year

It should be noted that in spring-third grade, the question about school year starting and ending dates was in the school fact sheet. Also, in spring-first grade and spring-kindergarten the composites for school year start and end dates were created differently because they were based on different questions. The question was in the student record abstract rather than the school fact sheet and was based on responses to multiple questions about start and end dates for school terms (e.g., semesters, trimesters). Composite variable names in past rounds started with a "L" prefix in spring-third grade (this was the prefix for the school fact sheet), and a "U" prefix in spring-first grade and spring-kindergarten (this was the prefix for the student record abstract). If the start and end dates varied for children in the same school, the composite was created by using the school start and end dates reported for the majority of children in a school. Because school start and end dates were only collected once in both the spring-third grade and spring-fifth grade, discrepancies in questionnaire reports for children in the same school were not an issue.

7.5.5 Student Record Abstract and FMS Composite Variables

The composite variables created from FMS data and the student record abstract follow.

7.5.5.1 Year-Round Schools (F6YRRND)

This composite was created using data from the FMS. The FMS flag was “1” if the child was in a year-round school. The values for the year-round school composite variable are 1 (Year round school) and 2 (Not year round school). If the child was schooled at home, the composite was coded as -1 (Not Applicable).

7.5.5.2 Indicator of Whether Child Received Special Education Services (F6SPECS)

The composite variable F6SPECS indicates whether or not the child received special education services in the spring of fifth grade, based on the presence or absence of a link to a special education teacher in the FMS. The values are 1 if the child received special education services, 2 if the child did not receive special education services, and -9 if the link was missing between the child and his or her teacher in the FMS.

7.5.5.3 Indicator of Whether Child Has an Individualized Education Plan (IEP) on Record at School (U6RIEP)

The variable U6RIEP indicates whether or not the child had an IEP or Individualized Family Service Plan (IFSP) on record at his/her school or another school in the spring of fifth grade. This information was recorded on the student record abstract. The values for the variable are 1 (child has an IEP/IFSP on record at his or her school or at another school) and 2 (child does not have an IEP/IFSP on record at his or her school). If the information was missing, U6RIEP was coded as -9 (Not Ascertained).

7.5.6 Parent Identifiers and Household Composition (P6DADID, P6MOMID, P6HPARNT, P6HDAD, P6HMOM, P6HFAMIL, P6MOMTYP, P6DADTYP)

The construction of parent identifiers and the household composition variables from the parent interview data was a two-step process. First, individuals identifying themselves as the child’s mother/father were located within the household roster, and the type of their relationship to the child (biological, adoptive, foster, step, partner of parent, or unknown) was established. For households containing more than one father or mother, a hierarchy was used to designate the "current," or residential, parent of each gender. 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, or "unknown-type" parent. If there were more than one father or mother of the same type, the one with the lower person number on the household roster was selected. Person number refers to the number each household member has on the roster list. Household members are listed in the order they are reported by the respondent. 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 P6HPARNT, P6HDAD, P6HMOM, and P6HFAMIL and parent-type variables P6MOMTYP, and P6DADTYP.

After the residential parents were identified and the composite variables were constructed, in any household without a parent, the household respondent (and his or her spouse/partner, if applicable) was assigned as a "parent figure." Parent demographic variables (including age, race/ethnicity, and education) were then constructed for all parents/parent figures. It should be noted, however, that these parent figures were not defined as parents (meaning biological, step-, adoptive, or foster) in the construction of the household composition composite variables described earlier. For example, for P6HFAMIL, 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 were placed in the "other" category for this composite. Likewise, for the composite P6HPARNT, parent figures were placed in categories 8 or 9 for related and unrelated guardians, respectively. Similarly, parent figures were included in the category "no resident mother" for P6HMOM and "no resident father" for P6HDAD. Thus, although persons reported as children's parent/guardians and the spouses/partners of the parent/guardians are included in the definitions of all the household composites, individuals later identified as parent figures in households in which no parents are present are not considered to be parents in the coding of the household composites.

Some parent-specific variables do include persons who were later identified as parent figures. These are as follows (variables for fathers are listed below but those for mothers are created in the same way):

- P6DADID (Household roster number of resident father, male guardian, or father figure);
- P6HDAGE (Age of resident father, male guardian, or father figure);
- P6HDRACE (Race and ethnicity of the father, male guardian, or father figure in the household);
- P6HDEMP (The employment status of the father, male guardian, or father figure in the household);
- P6DADOCC (Father, male guardian, or father figure's occupation);
- W5DADED (The father, male guardian, or father figure's highest level of education); and
- W5DADSCR (Father, male guardian, or father figure's occupation prestige score).

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 (P6REL_1, 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 (identified by P6PER_1, etc.) is the child's parent and the respondent's spouse/partner (identified from P6SPOUSE) is the same sex as the respondent.

There are two sections in the parent interview that asked questions specific to the parent-figure:

- PEQ, Parent education
- EMQ, Employment

Each of these sections was completed during the parent interview for up to two parents or parent-figures. To indicate which household member or members were the subject of each section, "pointer" variables that hold the original number of the household member on the household roster were used. To illustrate how the pointer variables work, suppose there is a household with both a mother and a father who were listed third and fourth in the household roster. If household member #3, the mother, was the first person to receive the PEQ education section, then the pointer variable P6EDUP1 will equal "3." The answers to the education questions for the mother will be contained in interview items in this section that end with the suffix "_1" (e.g., P6NDEG_1, P6DEGT_1, P6ENR_1, etc.). The suffix "_1" indicates that the data are for the first subject of the questions. Similarly, if household member #4, the father, was the second person to receive the PEQ education section, then the pointer variable P6EDUP2 will equal "4." The answers to the education questions for the father will be contained in interview items in this section that end with the suffix "_2" (e.g., P6NDEG_2, P6DEGT_2, P6ENR_2, etc.). The suffix "_2" indicates that the data are for the second subject of the questions. Table 7-13 identifies the pointer variables.

7.5.7 Industry and Occupation Codes Used in ECLS-K

This section describes the aggregated categories that were used for coding occupation in the ECLS-K.

1. **Executive, Administrative, and Managerial 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 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 consultants, library directors, custom house builders, and location managers. Legislators are also included in this category.

2. **Engineers, Surveyors, and Architects**

   The 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.

Table 7-13. Pointers to parent figure questions, spring-fifth grade: School year 2003–04

Person pointer		Interview item	
P6EDUP1	P6 PEQ010–060 HH PERSON POINTER 1	P6NDEG_1	P6 PEQ010 PERS 1 COMPLETED NEW DEGREE
		P6DEGT_1	P6 PEQ020 PERS 1 DEGREE TYPE COMPLETED
		P6ENR_1	P6 PEQ030 IF PERS 1 ENROLLED IN COURSES
		P6FPT_1	P6 PEQ040 PERS 1 COURSE FULL/PART TIME
		P6TRN_1	P6 PEQ050 IF PERS 1 GETS JOB TRAINING
		P6HRTR_1	P6 PEQ060 PERS 1 HR/WK SPEND ON TRAINING
P6EDUP2	P6 PEQ010–060 HH PERSON POINTER 2	P6NDEG_2	P6 PEQ010 PERS 2 COMPLETED NEW DEGREE
		P6DEGT_2	P6 PEQ020 PERS 2 DEGREE TYPE COMPLETED
		P6ENR_2	P6 PEQ030 PERS 2 ENROLLED IN COURSES
		P6FPT_2	P6 PEQ040 PERS 2 COURSE FULL/PART TIME
		P6TRN_2	PR PEQ050 IF PERS 2 GETS JOB TRAINING
		P6HTR_2	PR PEQ060 PERS 2 HR/WK SPEND ON TRAINING
P6EMPP1	P6 EMQ010–150 HH PERSON POINTER 1	P6CHJB_1	P6 EMQ010 PERS 1 CHNGD JOB SNC SPR 2002
		P6PAY_1	P6 EMQ020 PERS 1 HAD PAID JOB LAST WEEK
		P6VAC_1	P6 EMQ030 IF PERS 1 ON LEAVE PAST WEEK
		P6JOB_1	P6 EMQ040 PERSON 1 NUMBER OF ALL JOBS
		P6HRS_1	P6 EMQ050 PERSON 1 HOURS/WK AT ALL JOBS
		P6LOK_1	P6 EMQ060 PERS 1 SOUGHT JOB LAST 4 WEEKS
		P6DO1_1	P6 EMQ070 PERS 1 CHKD W/PUB EMPL AGENCY
		P6DO2_1	P6 EMQ070 PERS 1 CHKD W/PRIV EMP AGENCY
		P6DO3_1	P6 EMQ070 PERS 1 CHKD W/EMPLOYER DIRECTLY
		P6DO4_1	P6 EMQ070 PERS 1 CHKD W/FRIENDS & REL
		P6DO5_1	P6 EMQ070 PERS 1 PLACED OR ANSWERED ADS
		P6DO6_1	P6 EMQ070 PERS 1 READ WANT ADS
		P6DO7_1	P6 EMQ070 PERS 1 DID SOMETHING ELSE
		P6DOW_1	P6 EMQ080 WHAT PERSON 1 DOING LAST WEEK
		P6TAK_1	P6 EMQ100 PERS 1 JOB AVAILABLE LAST WEEK
		P6OCC_1	P6 EMQ130–50 1ST PERSON OCCUPATION CODE
P6EMPP2	P6EMQ010–150 HH PERSON POINTER 2	P6CHJB_2	P6 EMQ010 PERS 2 CHNGD JOB SNC SPRING 2002
		P6PAY_2	P6 EMQ020 PERS 2 HAD PAID JOB LAST WEEK
		P6VAC_2	P6 EMQ030 IF PERS 2 ON LEAVE PAST WEEK
		P6JOB_2	P6 EMQ040 PERSON 2 NUMBER OF ALL JOBS
		P6HRS_2	P6 EMQ050 PERSON 2 HOURS/WK AT ALL JOBS
		P6LOK_2	P6 EMQ060 PERS 2 SOUGHT JOB LAST 4 WEEKS
		P6DO1_2	P6 EMQ070 PERS 2 CHKD W/PUB EMPL AGENCY
		P6DO2_2	P6 EMQ070 PERS 2 CHKD W/PRIV EMP AGENCY
		P6DO3_2	P6 EMQ070 PERS 2 CHKD W/EMPLOYER DIRECTLY
		P6DO4_2	P6 EMQ070 PERS 2 CHKD W/FRIENDS & REL
		P6DO5_2	P6 EMQ070 PERS 2 PLACED OR ANSWERED ADS
		P6DO6_2	P6 EMQ070 PERS 2 READ WANT ADS
		P6DO7_2	P6 EMQ070 PERS 2 DID SOMETHING ELSE
		P6DOW_2	P6 EMQ080 WHAT PERSON 2 DOING LAST WEEK
		P6TAK_2	P6 EMQ100 PERS 2 JOB AVAILABLE LAST WEEK
		P6OCC_2	P6 EMQ130–50 2ND PERSON OCCUPATION CODE

SOURCE: U.S. Department of Education, National Center for Education Statistics, Early Childhood Longitudinal Study, Kindergarten Class of 1998–99 (ECLS-K), spring 2004.

3. **Natural Scientists and Mathematicians**

This category includes those engaged primarily in the application of scientific principles to research and development. Natural scientists are those in the physical sciences (e.g., chemistry, physics) and the life sciences (e.g., biology, agriculture, medicine). In addition, this category includes those in computer science, mathematics (including statistics), and operations research.

4. **Social Scientists, Social Workers, Religious Workers, and Lawyers**

This category includes occupations concerned with the social needs of people and in basic and applied research in the social sciences.

5. **Teachers: College, University, and Other Postsecondary Institution; Counselors, Librarians, and Archivists**

This category includes those who teach at higher education institutions and at other postsecondary (after high school) institutions, such as vocational institutes. In addition, vocational and educational counselors, librarians, and archivists are included here.

6. **Teachers, except Postsecondary Institution**

This category includes prekindergarten and kindergarten teachers, elementary and secondary teachers, special education teachers, instructional coordinators, and adult education teachers (outside postsecondary).

7. **Physicians, Dentists, and Veterinarians**

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 optometrists, podiatrists, chiropractors, hypnotherapists, and acupuncturists.

8. **Registered Nurses, Pharmacists, Dieticians, Therapists, and Physician's Assistants**

This category 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.

9. **Writers, Artists, Entertainers, and Athletes**

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, singing in entertainment; and participating in sports and athletics as a competitor or player and administering and directing athletic programs.

10. Health Technologists and Technicians

This category includes occupations concerned with providing technical assistance in the provision of health care. For example, clinical laboratory technologists and technicians, dental hygienists, radiologic technicians, licensed practical nurses (LPNs), and other health technologists are included here.

11. Technologists and Technicians, except Health

This category includes those providing technical assistance in engineering and scientific research, development, testing, and related activities, as well as operating and programming technical equipment and systems.

12. Marketing and Sales Occupations

This category includes occupations involving selling goods or services, purchasing commodities and property for resale, and conducting wholesale or retail business.

13. Administrative Support Occupations, including Clerks

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 teller, data entry keyer, etc.

14. Service Occupations

The category includes occupations providing personal and protective services to individuals, and current maintenance and cleaning for building and residences. Some examples include food service, health service (e.g., aides or assistants), cleaning services other than household, and personal services.

15. Agricultural, Forestry, and Fishing 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).

16. **Mechanics and Repairers**

Mechanics and repairers are persons who do adjustment, maintenance, part replacement, and repair of tools, equipment, and machines. Installation may be included if installation is usually done in conjunction with other duties of the repairers.

17. **Construction and Extractive 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, etc. Extractive occupations include oil well drillers, mining machine operators, and so on.

18. **Precision Production Occupations**

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 occupations including butchers and bakers. Plant and system operators include water and sewage, gas, power, chemical, petroleum, and other plant or system operators.

19. **Production Working Occupations**

This category 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.

20. **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.

21. **Handlers, Equipment Cleaners, Helpers, and Laborers**

This category includes occupations that involve helping other workers and performing routine nonmachine tasks. A wide variety of helpers, handlers, etc., are included in this category. Examples include construction laborers, freight, stock, and material movers, garage and service station related occupations, parking lot attendants, and vehicles washers and equipment cleaners.

22. 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 (e.g., occupations that are strictly military, such as “tank crew member” and “infantryman”).

7.6 Methodological Variables

To facilitate methodological research, eleven variables are included on the fifth-grade data file. The identifiers for parent interview work area (F6PWKARE), parent interviewer (F6PINTVR), child assessment work area (F6CWKARE), and child assessor (F6CASSOR) were extracted from the field management system.

Start and end times for both the child assessment (C6ASMSTM, C6ASMETM) and the parent interview (P6INTSTM, P6INTETM) were created from the keystroke-by-keystroke record of each parent interview and child assessment in the Blaise data. All four are text variables in the form *MM/DD/YY hour:minute:second AM/PM*. It should be noted that there may be more than one attempt to complete an interview or assessment, and those attempts could span several days. For example, an interviewer could begin a parent interview on one evening, and complete the remainder of the interview several days later. For this reason, variables to indicate the number of attempts necessary to complete the parent interview (P6ATTMPT) and the number of attempts necessary to complete the child assessment (C6APPMPT) have also been included on the file.

Finally, an indicator variable (F6PREFCV, Parent Interview Refusal Conversion) was created to flag cases that had, at any time, refused to respond to the parent interview but later agreed to participate. The values for F6PREFCV are 1=YES (refused but were converted to be a participant) and 2=NO (did not refuse).

7.7 Children Who Changed Schools

There are several variables in the file that can be used to determine if a child moved to a different school between rounds or moved to a different school during the fifth-grade data collection period.

7.7.1 Children Who Changed Schools During Fifth-Grade Data Collection

The variable S6_ID is a school identification number that indicates which school the child was in at the end of the fifth-grade data collection. There is another school ID variable, S6_ST_ID that indicates where the child was at the beginning of the round. By comparing school ID variables, users can determine whether the child physically moved from one school to another during round 6. For the vast majority of the children these two variables will be identical but, for those that moved during the data collection period, they will be different. If it was not known where the child was at the beginning or the end of the round, the scheme shown in table 7-14 for assigning ID numbers was used.[8]

Table 7-14. Case status and school ID numbers for children not followed or located, spring-fifth grade: School year 2003–04

Case status	S6_ID/S6_ST_ID
Not in the United States. The student now lives outside the U.S.	9993
Deceased. Information about the student indicates that he/she is deceased.	9994
Unlocatable. Field staff were unable to locate a transfer student in his/her new school.	9995
End of field period. Information on the transfer student's new school was identified too late in the field period for the case to be re-fielded for the assessment.	9996
Moved to nonsampled PSU. The transfer student enrolled in a school that was outside of ECLS-Ks sampled PSUs; field staff did not attempt to collect the assessment but did attempt to collect the parent interview.	9997
Do not follow. The transfer student was flagged by the statisticians as "do not follow" because of subsampling of transfer students due to cost constraints. If the child moved from his/her original school, field staff did not "follow" him or her to the new school and did not collect a child assessment or parent interview.	9998

SOURCE: U.S. Department of Education, National Center for Education Statistics, Early Childhood Longitudinal Study, Kindergarten Class of 1998–99 (ECLS-K), spring 2004.

[8] It should be noted that there were some children who could not be located during the field period for the beginning of the round but were located during the field period for the end of the round. Children who could not be located at the beginning of the round but who were located and enrolled in sampled schools at the end of the round will have S6_ST_ID values that begin with 999 and S6_ID values that are ID numbers for schools. Others who could not be located at the beginning or end round have 999 codes for both S6_ST_ID and S6_ID.

7.7.2 Children Who Changed Schools Between Rounds (R6DEST, R6R5SCHG)

Children moved between schools for a variety of reasons, but one factor was that a school terminated before the fifth grade and most of the students went to fifth grade at another school. This is known as a “destination school” and the move is known as a “destination move.” Destination schools were schools for which it was determined before data collection that at least four ECLS-K children would move into them from a school that ended before the fifth grade or a school that had closed. The variable on the file that indicates destination moves is R6DEST (moved to a spring-fifth grade destination school).

It should be noted that the destination school may also have been an originally sampled school; in this case, the school was a destination school only for the new students, not for the originally sampled students. The variable R6DEST was set to 1 (True) if a child made a school change and destination move to a spring-fifth grade destination school. If a child did not move to a spring-fifth grade destination school or did not move between schools at all, the composite is coded 0 (False). If the data are missing about whether the move was a destination move, the composite was coded -9 (Not Ascertained). If the child was schooled at home, the composite was coded as -1 (Not Applicable).

Another variable on the file that will be of interest to users examining school change is R6R5SCHG (school type change between spring-third grade and spring-fifth grade). It is used in the creation of R6DEST (R6R5SCHG must indicate a school change for R6DEST to be set to “1”) and indicates whether the child changed schools and, if so, the school type of the previous and the new school (e.g., whether the change was from public to private school, private to private school, etc.). R6R5SCHG is created by comparing the school IDs from spring-third grade and spring-fifth grade for children who were in the spring-third grade data collection. A difference in IDs indicated a change. If there was no difference in IDs, R6R5SCHG was coded 1 (child did not change schools). For children who changed schools, the spring-third grade school type variable S5SCTYP was compared to the spring-fifth grade school type variable S6SCTYP. Categories were assigned as appropriate (2 = child transferred from public to public; 3 = child transferred from private to private; 4 = child transferred from public to private; 5 = child transferred from private to public; and 6 = child transferred, other). Category six was used for those children who transferred schools, but school type was unknown. Children who were not in the spring-third grade data collection were coded -9, “Not Ascertained” on R6R5SCHG. Children who were homeschooled in spring-third grade or spring-fifth grade were coded -1, “Not Applicable” for R6R5SCHG.

► *Please note that the last two columns of table 7-15 in section 7.8 contain information that is file specific. Information for the restricted-use file is contained in the second to last column while information for the public-use and the K–5 longitudinal files is contained in the last column of table 7-15.*

7.8 Composite Table

Table 7-15 describes the composite and derived variables that are on the ECLS-K child catalog. Note that a few of the variables specified in the "derived from" column are intermediary variables that were not included in the final data set. An example of an intermediary variable is the child gender variable from parent questionnaires prior to spring-fifth grade, CHILDGEN. If this variable was missing, or had conflicting information across rounds of the study, information about gender was used from the field management system (FMS) or child report. The variable CHILDGEN is not included in the final data set, but the composite R6GENDER is included. Other intermediary variables are taken from either the FMS or the school master file and are not included on the data file.

The "derived from" column also contains the item numbers from the questionnaire, which help in identifying the items used in the creation of these composites. This information allows a user to decide whether to use the composite based on how it was defined.

Some variables in table 7-15 have been recoded or suppressed. Reasons for these data changes are discussed in section 7.9. All values for variables in the public use file are shown in the last column of table 7-15, including those that were recoded.

► *Please note that the following section (7.9) applies to the fifth-grade public-use file and to the K–5 longitudinal file. It does not apply to the fifth-grade restricted-use file.*

7.9 Masked Variables

All the variables from the ECLS-K restricted-use data file are included in the same order on the ECLS-K public-use data 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.

There are three types of modifications on the public-use data file.

- Outliers are top- or bottom- coded to prevent identification of unique schools, teachers, parents, and children without affecting overall data quality.
- Variables with too few cases and a sparse distribution are suppressed in the public-use data file. The values for these variables were set to -2 and labeled "suppressed" in the ECB.
- Certain continuous variables are modified into categorical variables, and certain categorical variables have their categories collapsed in the public-use data file. While this protects the cases from a disclosure risk, these variables can still be used in all different kinds of analysis such as logistic regression analysis.

In addition to these modifications, other procedures were used in all data files (restricted-use, public-use, and kindergarten–fifth grade longitudinal) to modify data based on the disclosure analysis NCES conducted in order to protect the identity of the respondents and children. Certain schools identified as at risk for disclosure had a 5 to 10 percent noise introduced in those variables that posed a risk for disclosure. Also, 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.

There is a comment field in the variable frequency distribution view screen of the electronic codebook that displays a comment for each masked variable indicating whether the variable from the restricted-use file has been recoded or suppressed in the public-use file. 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 7-16 presents the list of the masked variables. The table displays the variable name, variable label, and the comment displayed in the electronic codebook 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.)

All variables from the special education teacher questionnaire part A (i.e., all variables with the prefix D6), from the special education teacher questionnaire part B (i.e., all variables with the prefix E6), and from the student record abstract (i.e., all variables with the prefix U6) have been suppressed in the fifth-grade public-use file. Included in this group of suppressed variables are all teacher and school identifiers, which have last two characters “ID” and prefix D6, E6, or U6.

Table 7-15. Spring-fifth grade composite variables: School year 2003–04

ID	Variable name	Category	Description	Derived from	Restricted-use file values	Public-use file values
1	R6AGE	Child	Child's age in months at the time the direct child assessment occurred.	R6DOBMM, R6DOBDD, R6DOBYY (composites) or previous round date of birth variables if child was not in round 5, assessment date (taken from assessment audit trails)	Continuous	Recoded to the following: 1=Less than 126, 2=126 to less than 132, 3=132 to less than 138, 4=138 to less than 144, 5=144 or more
2	R6GENDER	Child	Child's gender	R6GENDER, CHILDGEN (INQ016 from previous parent interview, not delivered), FMS (variable not delivered), GENDER (composite from previous rounds) *Note: In spring-fifth grade, gender was no longer collected in the parent interview or child assessment.*	1=Male; 2=Female	1=Male; 2=Female
3	R6DOBMM	Child	Child date of birth month	R6DOBMM, DOBMM, CHILDDOB (not delivered) from first data collection in which reported in parent interview, and FMS date of birth variable *Note: In spring-fifth grade, date of birth was no longer collected in the parent interview or child assessment.*	1–12	1–12
4	R6DOBDD	Child	Child's date of birth day	R6DOBDD, DOBDD, CHILDDOB (not delivered) from first data collection in which reported in parent interview, and FMS date of birth variable *Note: In spring-fifth grade, date of birth was no longer collected in the parent interview or child assessment.*	1–31	1–31

See note at end of table.

Table 7-15. Spring-fifth grade composite variables: School year 2003–04—Continued

ID	Variable name	Category	Description	Derived from	Restricted-use file values	Public-use file values
5	R6DOBYY	Child	Child's date of birth year	W3RACETH, W1RACETH, WKRACETH (composites) *Note: In spring-fifth grade, race or ethnicity of the focal child was no longer collected in the parent interview.*	1990–1995	Recoded to a minimum value of 1992 and a maximum value of 1993
6	W5RACETH	Child	Race and ethnicity of the focal child	W3RACETH, W1RACETH, WKRACETH (composites) *Note: In spring-fifth grade, race or ethnicity of the focal child was no longer collected in the parent interview.*	1=White, 2=Black or African American, 3=Hispanic, race specified, 4=Hispanic, no race specified, 5=Asian, 6=Native Hawaiian or other Pacific Islander, 7=American Indian or Alaska Native, 8=More than 1 race, non-Hispanic	1=White, 2=Black or African American, 3=Hispanic, race specified, 4=Hispanic, no race specified, 5=Asian, 6=Native Hawaiian or other Pacific Islander, 7=American Indian or Alaska Native, 8=More than 1 race, non-Hispanic
7	R6RACE	Child	Child race and ethnicity	W5RACETH, W3RACETH, W1RACETH, WKRACETH, RACE from previous round (composites), C_RACE (FMS, not delivered), HI_PSU (FMS, not delivered)) *Note: In spring-fifth grade, race or ethnicity of the focal child was no longer collected in the parent interview.*	1=White, 2=Black or African American, 3=Hispanic, race specified, 4=Hispanic, no race specified, 5=Asian, 6=Native Hawaiian or other Pacific Islander, 7=American Indian or Alaska Native, 8=More than 1 race, non-Hispanic	1=White, 2=Black or African American, 3=Hispanic, race specified, 4=Hispanic, no race specified, 5=Asian, 6=Native Hawaiian or other Pacific Islander, 7=American Indian or Alaska Native, 8=More than 1 race, non-Hispanic
8	W5AMERIN	Child	Child is American Indian or Alaska Native	W3AMERIN, W1AMERIN, WKAMERIN (composites)	1=Yes, 2=No	1=Yes, 2=No

See note at end of table.

Table 7-15. Spring-fifth grade composite variables: School year 2003–04—Continued

ID	Variable name	Category	Description	Derived from	Restricted-use file values	Public-use file values
9	W5ASIAN	Child	Child is Asian	W3ASIAN, W1ASIAN, WKASIAN (composites)	1=Yes, 2=No	1=Yes, 2=No
10	W5BLACK	Child	Child is African American	W3BLACK, W1BLACK, WKBLACK (composites) .	1=Yes, 2=No	1=Yes, 2=No
11	W5PACISL	Child	Child is Native Hawaiian or other Pacific Islander	W3PACISL, W1PACISL, WKPACISL (composites)	1=Yes, 2=No	1=Yes, 2=No
12	W5WHITE	Child	Child is white	W3WHITE, W1WHITE, WKWHITE (composites)	1=Yes, 2=No	1=Yes, 2=No
13	W5MT1RAC	Child	Child is more than one race	W3MT1RAC, W1MT1RAC, WKMT1RAC (composites)	1=Yes, 2=No	1=Yes, 2=No
14	W5HISP	Child	Child is Hispanic	W3HISP, W1HISP, WKHISP (composites)	1=Yes, 2=No	1=Yes, 2=No
15	C6BMI	Child	Child's spring-fifth grade body mass index	C6HEIGHT, C6WEIGHT (composites)	Continuous	Continuous
16	C6HEIGHT	Child	Child's spring-fifth grade composite height	C6HGT1, C6HGT2	Continuous	Continuous
17	C6WEIGHT	Child	Child's spring-fifth grade composite weight	C6WGT1, C6WGT2	Continuous	Continuous

See note at end of table.

Table 7-15. Spring-fifth grade composite variables: School year 2003–04—Continued

ID	Variable name	Category	Description	Derived from	Restricted-use file values	Public-use file values
18	P6DISABL	Child	Child has a disability	P6DIAGNO (CHQ050), P6PROFFD (CHQ110), P6COMMU2 (CHQ170), P6DIFFH3 (CHQ210), P6VISIO2 (CHQ300), P6CORREC (CHQ316), P6RSVTSY (CHQ520), P6DIABEH (CHQ335), P6DIAEMO (CHQ360) P6DGNATT (CHQ.060) P6DGNACT (CHQ.120) P6DGNBEH (CHQ.337) P6DGNEMO (CHQ.365) P6BESTEY (CHQ.320) *Note: In spring-fifth grade, stem questions asked if the child ever had a disability rather than whether they had a disability since the last round of data collection as had been asked in round 5. Also, the spring-fifth grade composite excludes cases that have a diagnosis, but the diagnosis was "no problem," excludes cases with correctable vision, and includes cases that have vision problems such that the child's best eyesight allows him or her to see large print in books, form and/or color of objects but not detail, shadows, lights, or the child sees no light or has no light perception.*	1=Yes, 2=No	1=Yes, 2=No
19	P6CARNOW	Child	Focal child is currently receiving any nonparental care	P6RELNOW (CCQ.010), P6NRNOW (CCQ.150), P6CTRNOW (CCQ.260)	1=Yes, 2=No	1=Yes, 2=No
20	P6HRSNOW	Child	Total number of hours per week the focal child currently spends in all nonparental child care	P6RHRS (CCQ.090), P6NHRS (CCQ.190), P6CHRS (CCQ.355), P6RELNOW (CCQ.010), P6RELNUM (CCQ.060), P6RHROTH (CCQ.140), P6NRNOW (CCQ.150), P6NRNUM (CCQ.165), P6NHROTH (CCQ.250), P6CTRNOW (CCQ.260), P6CTRNUM (CCQ.325), P6CHROTH (CCQ403), P6RWEEK (CCQ.080), P6NWEEK (CCQ.180), P6CWEEK (CCQ.340)	Continuous	Continuous

See note at end of table.

Table 7-15. Spring-fifth grade composite variables: School year 2003–04—Continued

ID	Variable name	Category	Description	Derived from	Restricted-use file values	Public-use file values
21	P6NUMNOW	Child	Total number of all types of care arrangements the focal child currently has on a regular basis	P6RELNUM (CCQ.060), P6NRNUM (CCQ.165), P6CTRNUM (CCQ.325), P6RELNOW (CCQ.010), P6NRNOW (CCQ.150), P6CTRNOW (CCQ.260)	Continuous	Continuous
22	P6PRIMNW	Child	Primary, regular, nonparental child care arrangement in which the child currently spends the most hours per week	P6HRSNOW (composite), P6RHRS (CCQ.090), P6NHRS (CCQ.190), P6RPLACE (CCQ.070), P6NPLACE (CCQ.170), P6CHRS (CCQ.355)	0=No nonparental care, 1=Relative care in child's home, 2=Relative care in another home, 3=Nonrelative care in child's home, 4=Nonrelative care in another home, 5=Center-based program, 6=2 or more programs, 7 = Location of care varies	0=No nonparental care, 1=Relative care in child's home, 2=Relative care in another home; 3=Nonrelative care in child's home, 4=Nonrelative care in another home, 5=Center-based program, 6=2 or more programs, 7 = Location of care varies
23	F6SPECS	Child	This variable indicates whether or not the child received special education services based on the presence or absence of a link to a special education teacher in the FMS.	T_ID and TYPE (FMS variables not on file)	1 = Child got special education services 2 = Child did not get special education services	1 = Child got special education services 2 = Child did not get special education services

See note at end of table.

Table 7-15. Spring-fifth grade composite variables: School year 2003–04—Continued

ID	Variable name	Category	Description	Derived from	Restricted-use file values	Public-use file values
24	U6RIEP	Child	This variable indicates whether or not the child has an Individualized Education Program (IEP) or Individualized Family Service Plan (IFSP) on record at his/her school or at another school according to information from the student record abstract.	U6IEP (Student Record Abstract item 8)	1 = Child has IEP/IFSP on record at his/her school or another school 2 = Child does not have an IEP/IFSP	1 = Child has IEP/IFSP on record at his/her school or another school 2 = Child does not have an IEP/IFSP
25	R6DEST	Child	Moved to spring-fifth grade destination school	DESTSCH (School Master file variable not on file), R6R5SCHG	0 = False, 1= True	0 = False, 1= True
26	R6R5SCHG	Child	School type change between spring-third grade and spring-fifth grade	School ID, S6SCTYP, S5SCTYP, S4SCTYP, S3SCTYP, S2KSCTYP	1 = Child did not change schools 2 = Child transferred from public to public 3 = Child transferred from private to private 4 = Child transferred from public to private 5 = Child transferred from private to public 6 = Child transferred, other	1 = Child did not change schools 2 = Child transferred from public to public 3 = Child transferred from private to private 4 = Child transferred from public to private 5 = Child transferred from private to public 6 = Child transferred, other

See note at end of table.

Table 7-15. Spring-fifth grade composite variables: School year 2003–04—Continued

ID	Variable name	Category	Description	Derived from	Restricted-use file values	Public-use file values
27	R6ELIG	Child	Eligibility status of child	Child raw assessment status, ASSESSME (not on file)	1 = Eligible, 2 = Ineligible, out of scope, 3 = Ineligible, moved out of the country, 4 = Ineligible, deceased, 5 = Ineligible, mover not followed	1 = Eligible, 2 =Ineligible, out of scope, 3 = Ineligible, moved out of the country, 4 = Ineligible, deceased, 5 = Ineligible, mover not followed
28	C6ASMTST	Child	Child assessment status	C6RDGFLG, C6MTHFLG, C6SCIFLG, statistical flag SCORE_FG (not on file)	1 = Completely scoreable assessment data, 2 = Partially completed scoreable assessment data, 3= Category not in use in round 6; 4 = Child with disability, not assessed, 5 = Nonrespondent	1 = Completely scoreable assessment data, 2 = Partially completed scoreable assessment data, 3= Category not in use in round 6; 4 = Child with disability, not assessed, 5 = Nonrespondent
29	P6MOMID	Family/HH	Household roster number of resident mother, female guardian or mother figure	P6REL_1 to P6REL_25 (FSQ.130), P6UNR_1 to P6UNR_25 (FSQ.180), P6SPOUSE (FSQ.120), P6MOM_1 through P6MOM_25 (FSQ.140)	1–25	1–25
30	P6DADID	Family/HH	Household roster number of resident father, male guardian or father figure	P6REL_1 to P6REL_25 (FSQ.130), P6UNR_1 to P6UNR_25 (FSQ.180), P6SPOUSE (FSQ.120), P6DAD_1 through P6DAD_25 (FSQ.150)	1–25	1–25

See note at end of table.

Table 7-15. Spring-fifth grade composite variables: School year 2003–04—Continued

ID	Variable name	Category	Description	Derived from	Restricted-use file values	Public-use file values
31	P6HPARNT	Family/HH	Classification of the focal child's parents who reside in the household	P6REL_1 through P6REL_25 (FSQ.130), P6UNR_1 through P6UNR_25 (FSQ.180), P6HMOM, P6HDAD (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 parent and stepparent, 8=Related guardian(s), 9=Unrelated guardian(s)	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 parent and stepparent, 8=Related guardian(s), 9=Unrelated guardian(s)
32	P6HFAMIL	Family/HH	Family type categories using both parent and sibling information	P6REL_1 through P6REL_25 (FSQ.130), P6UNR_1 through P6UNR_25 (FSQ.180), P6HMOM, P6HDAD, P6NUMSIB (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	1=Two parents and sibling(s), 2=Two parents, no siblings, 3=One parent and sibling(s), 4=One parent, no siblings, 5=Other
33	P6NUMSIB	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	P6REL_1 to P6REL_25 (FSQ.130)	Continuous	Continuous

See note at end of table.

Table 7-15. Spring-fifth grade composite variables: School year 2003–04—Continued

ID	Variable name	Category	Description	Derived from	Restricted-use file values	Public-use file values
34	P6LESS18	Family/HH	Total number of household members younger than 18 years old	HHNUMBER and HH18ANDOVER (parent interview flags not on file)	Continuous	Continuous
35	P6OVER18	Family/HH	Total number of household members age 18 or older	HH18ANDOVER (parent interview flags not on file)	Continuous	Continuous
36	P6HTOTAL	Family/HH	Total number of household members	HHNUMBER (parent interview flag not on file)	Continuous	Continuous
37	P6TWIN	Family/HH	Household has sampled twins	P6PER_1 to P6PER_25 (person type in FSQ roster)	0=No twin in HH, 1=Twin in HH	0=No twin in HH, 1=Twin in HH
38	W5POVRTY	Family/HH	Poverty indicator	P6HILOW (PAQ.100), P6INCCAT(PAQ.110), W5INCCAT, P6HTOTAL (composites), and Census-defined thresholds	1=Below poverty threshold, 2=At or above poverty threshold	1=Below poverty threshold, 2=At or above poverty threshold
39	W5INCCAT	Family/HH	Household income	P6INCCAT(PAQ.110)	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	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

See note at end of table.

Table 7-15. Spring-fifth grade composite variables: School year 2003–04—Continued

ID	Variable name	Category	Description	Derived from	Restricted-use file values	Public-use file values
40	W5SESL	Family/HH	Socioeconomic scale	W5INCCAT, W5MOMED, W5DADED, W5MOMSCR, W5DADSCR (all composites)	Continuous	Continuous
41	W5SESQ5	Family/HH	Quintile indicator for W5SESL	W5SESL (composite)	1=First quintile (lowest), 2=Second quintile, 3=Third quintile, 4=Fourth quintile, 5=Fifth quintile (highest)	1=First quintile (lowest), 2=Second quintile, 3=Third quintile, 4=Fourth quintile, 5=Fifth quintile (highest)
42	W5PARED	Family/HH	Highest level of education for the child's parents or non-parental guardians who reside in the household. If only one parent or guardian resides in the household, W5PARED reflects that parent's education level.	W5MOMED, W5DADED (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	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 7-15. Spring-fifth grade composite variables: School year 2003–04—Continued

ID	Variable name	Category	Description	Derived from	Restricted-use file values	Public-use file values
43	W5MOMSCR	Family/HH	Mother, female guardian, or mother figure's occupation GSS prestige score	1989 GSS prestige scores, EMQ.120 (not on file), EMQ.130 (not on file), and EMQ.140 (not on file).	29.6 Handler, Equip, Cleaner, Helpers, Labor; 33.42 Production Working Occupation; 34.95 Service Occupations; 35.63 Agriculture, Forestry, Fishing Occupations; 35.78 Marketing & Sales Occupation; 35.92 Transportation, Material Moving; 37.67 Precision Production Occupation; 38.18 Administrative Support, Including Clerk; 39.18 Mechanics & Repairs; 39.2 Construction & Extractive Occupations; 48.69 Technologists, Except Health; 52.54 Writers, Artists, Entertainers, Athletes; 53.5 Executive, Admin, Managerial Occupation; 57.83 Health Technologists & Technicians; 59 Social Scientist/Workers, Lawyers; 61.56 Registered Nurses, Pharmacists; 62.87 Natural Scientists & Mathematicians; 63.43.	29.6 Handler, Equip, Cleaner, Helpers, Labor; 33.42 Production Working Occupation; 34.95 Service Occupations; 35.63 Agriculture, Forestry, Fishing Occupations; 35.78 Marketing & Sales Occupation; 35.92 Transportation, Material Moving; 37.67 Precision Production Occupation; 38.18 Administrative Support, Including Clerk; 39.18 Mechanics & Repairs; 39.2 Construction & Extractive Occupations; 48.69 Technologists, Except Health; 52.54 Writers, Artists, Entertainers, Athletes; 53.5 Executive, Admin, Managerial Occupation; 57.83 Health Technologists & Technicians; 59 Social Scientist/Workers, Lawyers; 61.56 Registered Nurses, Pharmacists; 62.87 Natural Scientists & Mathematicians; 63.43.

See note at end of table.

Table 7-15. Spring-fifth grade composite variables: School year 2003–04—Continued

ID	Variable name	Category	Description	Derived from	Restricted-use file values	Public-use file values
43	W5MOMSCR (continued)	Family/HH	Mother, female guardian, or mother figure's occupation GSS prestige score	1989 GSS prestige scores, EMQ.120 (not on file), EMQ.130 (not on file), and EMQ.140 (not on file).	Teacher, Except Postsecondary; 64.89 Engineers, Surveyors, & Architects; 72.1 Teachers; College, Postsecondary Counselors, Librarians; 77.5 Physicians, Dentists, Veterinarians	Teacher, Except Postsecondary; 64.89 Engineers, Surveyors, & Architects; 72.1 Teachers; College, Postsecondary Counselors, Librarians; 77.5 Physicians, Dentists, Veterinarians
44	P6HDAD	Family/HH	Indicates whether the birth, adoptive, step or foster father of the focal child resides in the household with the focal child	P6REL_1 through P6REL_25(FSQ.130), P6DAD_1 through P6DAD_25 (FSQ.150), P6UNR_1 through P6UNR_25 (FSQ.180), P6PARTNR (FSQ.110), P6SPOUSE (FSQ.120)	1=Biological, 2=Adoptive, 3=Step, 4=Foster, 5=Partner, 6=Don't know type, 7= No resident father	1=Biological, 2=Adoptive, 3=Step, 4=Foster, 5=Partner, 6=Don't know type, 7= No resident father
45	P6HDAGE	Family/HH	Age of resident father, male guardian or father figure	P6AGE_1 through P6AGE_25 (FSQ.030), P6DADID	Continuous	Continuous
46	P6HDRACE	Family/HH	Race and ethnicity of the father, male guardian or father figure in the household	RACE1, RACE2, RACE3, RACE4, RACE5, RACE6 (variables coded in parent interview based on P6RC1_1 through P6RC6_1 up to P6RC1_25 through P6RC6_25 (FSQ.195), and P6HSP_1 through P6HSP_25 (FSQ.190)).	1=White, 2=Black or African American, 3=Hispanic, race specified, 4=Hispanic, no race specified, 5=Asian, 6=Native Hawaiian or other Pacific Islander, 7=American Indian or Alaska Native, 8=More than one race, non-Hispanic	1=White, 2=Black or African American, 3=Hispanic, race specified, 4=Hispanic, no race specified, 5=Asian, 6=Native Hawaiian or other Pacific Islander, 7=American Indian or Alaska Native, 8=More than one race, non-Hispanic

See note at end of table.

Table 7-15. Spring-fifth grade composite variables: School year 2003–04—Continued

ID	Variable name	Category	Description	Derived from	Restricted-use file values	Public-use file values
47	W5DADSCR	Family/HH	Father, male guardian, or father figure's occupation GSS prestige score	1989 GSS prestige scores, EMQ.120, EMQ.130, and EMQ.140 (not on file).	29.6 Handler, Equip, Cleaner, Helpers, Labor; 33.42 Production Working Occupation; 34.95 Service Occupations; 35.63 Agriculture, Forestry, Fishing Occupations; 35.78 Marketing & Sales Occupation; 35.92 Transportation, Material Moving; 37.67 Precision Production Occupation; 38.18 Administrative Support, Including Clerk; 39.18 Mechanics & Repairs; 39.2 Construction & Extractive Occupations; 48.69 Technologists, Except Health; 52.54 Writers, Artists, Entertainers, Athletes; 53.5 Executive, Admin, Managerial Occupation; 57.83 Health Technologists & Technicians; 59 Social Scientist/Workers, Lawyers; 61.56 Registered Nurses, Pharmacists; 62.87 Natural Scientists & Mathematicians; 63.43.	29.6 Handler, Equip, Cleaner, Helpers, Labor; 33.42 Production Working Occupation; 34.95 Service Occupations; 35.63 Agriculture, Forestry, Fishing Occupations; 35.78 Marketing & Sales Occupation; 35.92 Transportation, Material Moving; 37.67 Precision Production Occupation; 38.18 Administrative Support, Including Clerk; 39.18 Mechanics & Repairs; 39.2 Construction & Extractive Occupations; 48.69 Technologists, Except Health; 52.54 Writers, Artists, Entertainers, Athletes; 53.5 Executive, Admin, Managerial Occupation; 57.83 Health Technologists & Technicians; 59 Social Scientist/Workers, Lawyers; 61.56 Registered Nurses, Pharmacists; 62.87 Natural Scientists & Mathematicians; 63.43.

See note at end of table.

Table 7-15. Spring-fifth grade composite variables: School year 2003–04—Continued

ID	Variable name	Category	Description	Derived from	Restricted-use file values	Public-use file values
47	W5DADSCR (continued)	Family/HH	Father, male guardian, or father figure's occupation GSS prestige score	1989 GSS prestige scores, EMQ.120, EMQ.130, and EMQ.140 (not on file).	Teacher, Except Postsecondary; 64.89 Engineers, Surveyors, & Architects; 72.1 Teachers; College, Postsecondary Counselors, Librarians; 77.5 Physicians, Dentists, Veterinarians	Teacher, Except Postsecondary; 64.89 Engineers, Surveyors, & Architects; 72.1 Teachers; College, Postsecondary Counselors, Librarians; 77.5 Physicians, Dentists, Veterinarians
48	W5DADED	Family/HH	The father, male guardian, or father figure's highest level of education	P6NDEG_1 through P6NDEG_2 (PEQ.010), P6DEGT _1 through P6DEGT _2 (PEQ.020), P6HSD_1 through P6HSD_2 (PEQ.021) *Note: PEQ.021 is a question in spring-fifth grade that was not in spring-third grade and is used to provide information about whether persons have a high school degree or equivalent. Also, it should be noted that some cases that had lower educations in spring-fifth grade than in an earlier round of the study were asked for their highest level of education again.*	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	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
49	P6HDEMP	Family/HH	The work status of the father, male guardian or father figure in the household.	P6HRS_1, _2 (EMQ.050), P6PAY_1, _2 (EMQ.020), P6VAC_1, _2 (EMQ 030), P6LOK_1, _2 (EMQ.060), P6DO1_1, _2 (EMQ.070), P6DO2_1, _2 (EMQ.070), P6DO3_1, _2 (EMQ.070), P6DO4_1, _2 (EMQ.070), P6DO5_1, _2 (EMQ.070), P6DO6_1, _2 (EMQ.070), P6DO7_1, _2 (EMQ.070), P6CHJB_1, _2 (EMQ.010)	1=35 hours or more per week, 2=Less than 35 hours per week, 3=Looking for work, 4=Not in the labor force	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 7-15. Spring-fifth grade composite variables: School year 2003–04—Continued

ID	Variable name	Category	Description	Derived from	Restricted-use file values	Public-use file values
50	P6DADOCC	Family/HH	Father, male guardian or father figure's occupation	Combination of P6CHJB_1, _2 (EMQ.010), EMQ.120, EMQ.130, and EMQ.140 (not on file)	01 Executive, Admin, Managerial Occupation 02 Engineers, Surveyors, & Architects 03 Natural Scientists & Mathematicians 04 Social Scientist/Workers, Lawyers 05 University Teachers, Postsecondary Counselors, Librarians 06 Teacher, except postsecondary 07 Physicians, Dentists, Veterinarians 08 Registered Nurses, Pharmacists 09 Writers, Artists, Entertainers, Athletes 10 Health Technologists & Technicians 11 Technologists, except Health 12 Marketing & Sales Occupation 13 Administrative Support, incl. Clerk 14 Service Occupations 15 Agriculture, Forestry, Fishing	01 Executive, Admin, Managerial Occupation 02 Engineers, Surveyors, & Architects 03 Natural Scientists & Mathematicians 04 Social Scientist/Workers, Lawyers 05 University Teachers, Postsecondary Counselors, Librarians 06 Teacher, except postsecondary 07 Physicians, Dentists, Veterinarians 08 Registered Nurses, Pharmacists 09 Writers, Artists, Entertainers, Athletes 10 Health Technologists & Technicians 11 Technologists, except Health 12 Marketing & Sales Occupation 13 Administrative Support, incl. Clerk 14 Service Occupations 15 Agriculture, Forestry, Fishing

See note at end of table.

Table 7-15. Spring-fifth grade composite variables: School year 2003–04—Continued

ID	Variable name	Category	Description	Derived from	Restricted-use file values	Public-use file values
50	P6DADOCC (continued)	Family/HH	Father, male guardian or father figure's occupation	Combination of P6CHJB_1, _2 (EMQ.010), EMQ.120, EMQ.130, and EMQ.140 (not on file)	16 Mechanics & Repairs 17 Construction & Extractive Occupations 18 Precision Production Occupation 19 Production Working Occupation 20 Transportation, Material Moving 21 Handler, Equip, Cleaner, Helpers, Labor 22 Unemployed or Retired	16 Mechanics & Repairs 17 Construction & Extractive Occupations 18 Precision Production Occupation 19 Production Working Occupation 20 Transportation, Material Moving 21 Handler, Equip, Cleaner, Helpers, Labor 22 Unemployed or Retired
51	P6HMOM	Family/HH	Indicates whether the birth, adoptive, step, or foster mother of the focal child resides in the household with the focal child	P6REL_1 through P6REL_25(FSQ.130), P6MOM_1 through P6MOM_25 (FSQ.140), P6UNR_1 through P6UNR_25 (FSQ.180), P6PARTNR (FSQ.110), P6SPOUSE (FSQ.120)	1=Biological, 2=Adoptive, 3=Step, 4=Foster, 5=Partner, 6=Don't know type, 7=No resident mother	1=Biological, 2=Adoptive, 3=Step, 4=Foster, 5=Partner, 6=Don't know type, 7=No resident mother
52	P6HMAGE	Family/HH	Age of resident mother, female guardian or mother figure	P6AGE_1 through P6AGE_25 (FSQ.030), P6MOMID	Continuous	Continuous

See note at end of table.

Table 7-15. Spring-fifth grade composite variables: School year 2003–04—Continued

ID	Variable name	Category	Description	Derived from	Restricted-use file values	Public-use file values
53	P6HMRACE	Family/HH	Race and ethnicity of the mother, female guardian, or mother figure in the household	RACE1, RACE2, RACE3, RACE4, RACE5, RACE6 (These variables are coded in parent interview--see W5RACETH specs for details. The original race variables are P6RC1_1 through P6RC6_1 up to P6RC1_25 through P6RC6_25 (FSQ.195), and P6HSP_1 through P6HSP_25 (FSQ.190)).	1=White, 2=Black or African American, 3=Hispanic, race specified, 4=Hispanic, no race specified, 5=Asian, 6=Native Hawaiian or other Pacific Islander, 7=American Indian or Alaska Native, 8=More than one race, non-Hispanic	1=White, 2=Black or African American, 3=Hispanic, race specified, 4=Hispanic, no race specified, 5=Asian, 6=Native Hawaiian or other Pacific Islander, 7=American Indian or Alaska Native, 8=More than one race, non-Hispanic
54	W5MOMED	Family/HH	Mother, female guardian, or mother figure's highest level of education	P6NDEG_1 through P6NDEG_2 (PEQ.010), P6DEGT _1 through P6DEGT _2 (PEQ.020), P6HSD_1 through P6HSD_2 (PEQ.021*)* *Note: PEQ.021 is a question in spring-fifth grade that was not in spring-third grade and is used to provide information about whether persons have a high school degree or equivalent. Also, it should be noted that some cases that had lower educations in spring-fifth grade than in an earlier round of the study were asked for their highest level of education again.*	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	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
55	P6HMEMP	Family/HH	The work status of the mother, female guardian, or mother figure in the household	P6HRS_1, _2 (EMQ.050), P6PAY_1, _2 (EMQ.020), P6VAC_1, _2 (EMQ 030), P6LOK_1, _2 (EMQ.060), P6DO1_1, _2 (EMQ.070), P6DO2_1, _2 (EMQ.070), P6DO3_1, _2 (EMQ.070), P6DO4_1, _2 (EMQ.070), P6DO5_1, _2 (EMQ.070), P6DO6_1, _2 (EMQ.070), P6DO7_1, _2 (EMQ.070), P6CHJB_1, _2 (EMQ.010)	1=35 hours or more per week, 2=Less than 35 hours per week, 3=Looking for work, 4=Not in the labor force	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 7-15. Spring-fifth grade composite variables: School year 2003–04—Continued

ID	Variable name	Category	Description	Derived from	Restricted-use file values	Public-use file values
56	P6MOMOCC	Family/HH	Mother, female guardian, or mother figure's occupation	Combination of P6CHJB_1, _2, EMQ.010, EMQ.120, EMQ.130, and EMQ.140 (not on file)	01 Executive, Admin, Managerial Occupation 02 Engineers, Surveyors, & Architects 03 Natural Scientists & Mathematicians 04 Social Scientist/Workers, Lawyers 05 University Teachers, Postsecondary Counselors, Librarians 06 Teachers, except postsecondary 07 Physicians, Dentists, Veterinarians; 08 Registered Nurses, Pharmacists 09 Writers, Artists, Entertainers, Athletes 10 Health Technologists & Technicians 11 Technologists, except Health 12 Marketing & Sales Occupation 13 Administrative Support, including Clerk 14 Service Occupations 15 Agriculture, Forestry, Fishing Occupations 16 Mechanics & Repairs	01 Executive, Admin, Managerial Occupation 02 Engineers, Surveyors, & Architects 03 Natural Scientists & Mathematicians 04 Social Scientist/Workers, Lawyers 05 University Teachers, Postsecondary Counselors, Librarians 06 Teachers, except postsecondary 07 Physicians, Dentists, Veterinarians; 08 Registered Nurses, Pharmacists 09 Writers, Artists, Entertainers, Athletes 10 Health Technologists & Technicians 11 Technologists, except Health 12 Marketing & Sales Occupation 13 Administrative Support, including Clerk 14 Service Occupations 15 Agriculture, Forestry, Fishing Occupations 16 Mechanics & Repairs

See note at end of table.

Table 7-15. Spring-fifth grade composite variables: School year 2003–04—Continued

ID	Variable name	Category	Description	Derived from	Restricted-use file values	Public-use file values
56	P6MOMOCC (continued)	Family/HH	Mother, female guardian, or mother figure's occupation	Combination of P6CHJB_1, _2, EMQ.010, EMQ.120, EMQ.130, and EMQ.140 (not on file)	17 Construction & Extractive Occupations 18 Precision Production Occupation 19 Production Working Occupation 20 Transportation, Material Moving 21 Handler, Equip, Cleaner, Helpers, Labor 22 Unemployed or Retired	17 Construction & Extractive Occupations 18 Precision Production Occupation 19 Production Working Occupation 20 Transportation, Material Moving 21 Handler, Equip, Cleaner, Helpers, Labor 22 Unemployed or Retired
57	P6ABSDAD	Family/ HH	Type of nonresident father	P6REL_1 through P6REL_25 (FSQ.130), P6CTP_N1, P6CTP_N2, P6CTP_N3, P6CTP_N4 (all from item NRQ.100) *Note: NRQ.100 was not asked in round 6 if the nonresident biological father was recorded as deceased in any previous round. It was also not asked if the nonresident adoptive father was recorded as deceased in one of the two most recent rounds, round 4 or round 5.*	1=Biological only, 2=Both biological and adoptive	1=Biological only, 2=Both biological and adoptive
58	P6ABSMOM	Family/HH	Type of nonresident mother	P6REL_1 through P6REL_25 (FSQ.130), P6CTP_N1, P6CTP_N2, P6CTP_N3, P6CTP_N4 (all from item NRQ.100) *Note: NRQ.100 was not asked in round 6 if the nonresident biological mother was recorded as deceased in any previous round. It was also not asked if the nonresident adoptive mother was recorded as deceased in one of the two most recent rounds, round 4 or round 5.*	1=Biological only, 2=Both biological and adoptive	1=Biological only, 2=Both biological and adoptive

See note at end of table.

Table 7-15. Spring-fifth grade composite variables: School year 2003–04—Continued

ID	Variable name	Category	Description	Derived from	Restricted-use file values	Public-use file values
59	P6FSRAW	Family/HH	Household food security raw score, a simple count of the number of food security items affirmed by the parent.	P6WORRFD (FDQ.130A), P6FDLAST (FDQ.130B), P6BLMEAL (FDQ.130C), P6LOWCST (FDQ.130D), P6NOBAL (FDQ.130E), P6CANTAF (FDQ.130F), P6EVCUT2 (FDQ.140), P6EVCUT (FDQ.150), P6EATLES (FDQ.160), P6HUNGRY (FDQ.170), P6LOSEWT (FDQ.180), P6NOTEAT (FDQ.190), P6NOTEA2 (FDQ.200), P6CUTML (FDQ.210), P6CHSKIP (FDQ.220), P6OFTCUT (FDQ.230), P6CHIEVR (FDQ.240), P6NOMONY (FDQ.250)	Continuous	Continuous
60	P6FSSCAL	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.	P6WORRFD (FDQ.130A), P6FDLAST (FDQ.130B), P6BLMEAL (FDQ.130C), P6LOWCST (FDQ.130D), P6NOBAL (FDQ.130E), P6CANTAF (FDQ.130F), P6EVCUT2 (FDQ.140), P6EVCUT (FDQ.150), P6EATLES (FDQ.160), P6HUNGRY (FDQ.170), P6LOSEWT (FDQ.180), P6NOTEAT (FDQ.190), P6NOTEA2 (FDQ.200), P6CUTML (FDQ.210), P6CHSKIP (FDQ.220), P6OFTCUT (FDQ.230), P6CHIEVR (FDQ.240), P6NOMONY (FDQ.250)	Continuous	Continuous
61	P6FSSTAT	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)	P6WORRFD (FDQ.130A), P6FDLAST (FDQ.130B), P6BLMEAL (FDQ.130C), P6LOWCST (FDQ.130D), P6NOBAL (FDQ.130E), P6CANTAF (FDQ.130F), P6EVCUT2 (FDQ.140), P6EVCUT (FDQ.150), P6EATLES (FDQ.160), P6HUNGRY (FDQ.170), P6LOSEWT (FDQ.180), P6NOTEAT (FDQ.190), P6NOTEA2 (FDQ.200), P6CUTML (FDQ.210), P6CHSKIP (FDQ.220), P6OFTCUT (FDQ.230), P6CHIEVR (FDQ.240), P6NOMONY (FDQ.250)	1 = Food secure; 2 = Food insecure without hunger; 3 = Food insecure with hunger (moderate); 4 = Food insecure with hunger (severe)	1 = Food secure; 2 = Food insecure without hunger; 3 = Food insecure with hunger (moderate); 4 = Food insecure with hunger (severe)

See note at end of table.

Table 7-15. Spring-fifth grade composite variables: School year 2003–04—Continued

ID	Variable name	Category	Description	Derived from	Restricted-use file values	Public-use file values
62	P6FSCHRA	Family/HH	Children's food security raw score, a simple count of the number of child-referenced food security items affirmed by the parent	P6LOWCST (FDQ.130D), P6NOBAL (FDQ.130E), P6CANTAF (FDQ.130F), P6CUTML (FDQ.210), P6CHSKIP (FDQ.220), P6OFTCUT (FDQ.230), P6CHIEVR (FDQ.240), P6NOMONY (FDQ.250)	Continuous	Continuous
63	P6FSCHSC	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.	P6LOWCST (FDQ.130D), P6NOBAL (FDQ.130E), P6CANTAF (FDQ.130F), P6CUTML (FDQ.210), P6CHSKIP (FDQ.220), P6OFTCUT (FDQ.230), P6CHIEVR (FDQ.240), P6NOMONY (FDQ.250)	Continuous	Continuous
64	P6FSCHST	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.	P6LOWCST (FDQ.130D), P6NOBAL (FDQ.130E), P6CANTAF (FDQ.130F), P6CUTML (FDQ.210), P6CHSKIP (FDQ.220), P6OFTCUT (FDQ.230), P6CHIEVR (FDQ.240), P6NOMONY (FDQ.250)	1 = Food secure or food insecure without hunger among children; 2 = Food insecure with hunger among children	1 = Food secure or food insecure without hunger among children; 2 = Food insecure with hunger among children
65	P6FSADRA	Family/HH	Adult food security raw score, a simple count of the number of household- and adult-referenced food security items affirmed by the parent	P6WORRFD (FDQ130A), P6FDLAST (FDQ130B), P6BLMEAL (FDQ130C), P6EVCUT2 (FDQ140), P6EVCUT (FDQ150), P6EATLES (FDQ160), P6HUNGRY (FDQ170), P6LOSEWT (FDQ180), P6NOTEAT (FDQ190), P6NOTEA2 (FDQ200)	Continuous	Continuous

See note at end of table.

Table 7-15. Spring-fifth grade composite variables: School year 2003–04—Continued

ID	Variable name	Category	Description	Derived from	Restricted-use file values	Public-use file values
66	P6FSADSC	Family/HH	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.	P6WORRFD (FDQ130A), P6FDLAST (FDQ130B), P6BLMEAL (FDQ130C), P6EVCUT2 (FDQ140), P6EVCUT (FDQ150), P6EATLES (FDQ160), P6HUNGRY (FDQ170), P6LOSEWT (FDQ180), P6NOTEAT (FDQ190), P6NOTEA2 (FDQ200)	Continuous	Continuous
67	P6FSADST	Family/HH	A categorical measure of adult's food security status that identifies households as food secure, food insecure without hunger, and food insecure with hunger among adults.	P6WORRFD (FDQ130A), P6FDLAST (FDQ130B), P6BLMEAL (FDQ130C), P6EVCUT2 (FDQ140), P6EVCUT (FDQ150), P6EATLES (FDQ160), P6HUNGRY (FDQ170), P6LOSEWT (FDQ180), P6NOTEAT (FDQ190), P6NOTEA2 (FDQ200)	1=Food secure; 2 = Food insecure without hunger; 3 = Food insecure with hunger	1=Food secure; 2 = Food insecure without hunger; 3 = Food insecure with hunger
68	P6RESID	Family/HH	Household roster number of respondent	P6PER_1 to P6PER_25 (parent interview household roster person type)	1–25	1–25
69	P6RESREL	Family/HH	Respondent relationship to focal child	P6REL_1 through P6REL_25(FSQ.130), P6UNR_1 through P6UNR_25 (FSQ.180), P6MOM_1 through P6MOM_25 (FSQ.140), P6DAD_1 through P6DAD_25 (FSQ.150)	1 = Biological mother 2 = Other mother type 3 = Biological father 4 = Other father type 5 = Non-parent relative 6 = Non-relative	1 = Biological mother 2 = Other mother type 3 = Biological father 4 = Other father type 5 = Non-parent relative 6 = Non-relative
70	P6CHLDID	Family/HH	Household roster number of child	P6PER_1 to P6PER_25 (parent interview household roster person type)	1–25	1–25

See note at end of table.

Table 7-15. Spring-fifth grade composite variables: School year 2003–04—Continued

ID	Variable name	Category	Description	Derived from	Restricted-use file values	Public-use file values
71	P6ERRFLG	Family/ HH flag	Household roster has clear errors	P6REL_1 to P6REL_25 (FSQ.130), P6UNR_1 to P6UNR_25 (FSQ.180), P6JOI_1 to P6JOI_25 (round joined study), P6RDP_1 to P6RDP_25 (round departed study), P6REASL1 to P6REAS25 (reason left household) *Note: In spring-fifth grade, the category of "roster error" (category 6) was added back to the parent interview and can be used for setting the error flag.*	0 = False, 1 = True	0 = False, 1 = True
72	P6EDIT	Family/ HH flag	Parent household matrix was edited	HOLDINGS (parent interview editing flag – not on file)	0 = False, 1 = True	0 = False, 1 = True
73	P6SHCHG	Family/ HH flag	Household roster had a change between rounds.	P6JOI_1 to P6JOI_25 (round joined study), P6RDP_1 to P6RDP_25 (round departed study), P6REASL1 to P6REAS25 (reason left household) *Note: In spring-fifth grade, the category of "roster error" (category 6) was added back to the parent interview and can be used for setting the flag.*	0 = False, 1= True	0 = False, 1= True
74	P6PARDAT	Family/ HH flag	Presence of parent data	Presence or absence of parent interview	0 = False, 1 = True	0 = False, 1 = True
75	T6GLVL	Teacher	Grade level of child	G6GRENRL (RDG Q1), E6ENRGR (SPB Q2), C_GRADE (from FMS), C6FIFTH (ACQ.005), C6GRADE (ACQ.010), C6INGRAD (AIQ.030) *Note: Categories have been renumbered since spring-third grade. Category numbers now correspond to grade level numbers.*	0 = Kindergarten; 1 = First grade, 2 = Second grade, 3 = Third Grade, 4 = Fourth Grade, 5 = Fifth Grade, 6 = Sixth Grade, 7 = Seventh Grade, 8 = Eighth Grade, 9= Ungraded classroom	0 = Kindergarten; 1 = First grade, 2 = Second grade, 3 = Third Grade, 4 = Fourth Grade, 5 = Fifth Grade, 6 = Sixth Grade, 7 = Seventh Grade, 8 = Eighth Grade, 9= Ungraded classroom

See note at end of table.

Table 7-15. Spring-fifth grade composite variables: School year 2003–04—Continued

ID	Variable name	Category	Description	Derived from	Restricted-use file values	Public-use file values
76	G6CLSZ	Class	Number of students in reading class	G6TOTRA (RDG Q16), G6TOTGEN (RDG Q17) *Note: There are now separate class size variables for reading, math, and science. The sum of children by age is not a variable in the teacher questionnaires in spring-fifth grade so that is not listed as a source this time for any of the "CLSZ" class size variables. Also, a total of girls and boys is now included in the teacher questionnaires ("TOTGEN") so this total is compared to the total by race.*	Continuous	Recoded to a minimum value of 10 and a maximum value of 35
77	M6CLSZ	Class	Number of students in mathematics class	M6TOTRA (MTH Q6), M6TOTGEN (MTH Q7)	Continuous	Recoded to a minimum value of 10 and a maximum value of 35
78	N6CLSZ	Class	Number of students in science class	N6TOTRA (SCI Q5), N6TOTGEN (SCI Q6)	Continuous	Recoded to a minimum value of 10 and a maximum value of 35
79	G6PLEP	Class	Percentage of limited English proficient children in the reading class	G6NUMLE (RDG Q18), G6CLSZ (composite) *Note: There are now separate "PLEP" variables for reading, math, and science. Also, two variables that were used to create the "PLEP" composite in the past are not included in the questionnaire in spring-fifth grade. The source variables with endings of "OTLA" and "LEP" are no longer used to derive the "PLEP" variables.*	0 - 100	Recoded to the following: 1=Less than 1%, 2=1% to less than 5%, 3=5% to less than 10%, 4=10% to less than 25%, 5=25% or more

See note at end of table.

Table 7-15. Spring-fifth grade composite variables: School year 2003–04—Continued

ID	Variable name	Category	Description	Derived from	Restricted-use file values	Public-use file values
80	M6PLEP	Class	Percentage of limited English proficient children in the mathematics class	M6NUMLE (MTH Q8), M6CLSZ (composite)	0 - 100	Recoded to the following: 1=Less than 1%, 2=1% to less than 5%, 3=5% to less than 10%, 4=10% to less than 25%, 5=25% or more
81	N6PLEP	Class	Percentage of limited English proficient children in the science class	N6NUMLE (SCI Q7), N6CLSZ (composite)	0 - 100	Recoded to the following: 1=Less than 1%, 2=1% to less than 5%, 3=5% to less than 10%, 4=10% to less than 25%, 5=25% or more
82	G6PBLK	Class	Percent of blacks in reading class—child-level data	G6BLACK (RDG Q16), G6CLSZ (composite) *Note: There are now separate "PBLK" variables for reading, math, and science.*	0–100	Recoded to the following: 1=Less than 1%, 2=1% to less than 5%, 3=5% to less than 10%, 4=10% to less than 25%, 5=25% or more
83	M6PBLK	Class	Percent of blacks in mathematics class—child-level data	M6BLACK (MTH Q6), M6CLSZ (composite)	0–100	Recoded to the following: 1=Less than 1%, 2=1% to less than 5%, 3=5% to less than 10%, 4=10% to less than 25%, 5=25% or more

See note at end of table.

Table 7-15. Spring-fifth grade composite variables: School year 2003–04—Continued

ID	Variable name	Category	Description	Derived from	Restricted-use file values	Public-use file values
84	N6PBLK	Class	Percent of blacks in science class—child-level data	N6BLACK (SCI Q5), N6CLSZ (composite)	0–100	Recoded to the following: 1=Less than 1%, 2=1% to less than 5%, 3=5% to less than 10%, 4=10% to less than 25%, 5=25% or more
85	G6PHIS	Class	Percent of Hispanics in reading class—child-level data	G6HISP (RDG Q16), G6CLSZ (composite) *Note: There are now separate "PHIS" variables for reading, math, and science.*	0–100	Recoded to the following: 1=Less than 1%, 2=1% to less than 5%, 3=5% to less than 10%, 4=10% to less than 25%, 5=25% or more
86	M6PHIS	Class	Percent of Hispanics in mathematics class—child-level data	M6HISP (MTH Q6), M6CLSZ (composite)	0–100	Recoded to the following: 1=Less than 1%, 2=1% to less than 5%, 3=5% to less than 10%, 4=10% to less than 25%, 5=25% or more
87	N6PHIS	Class	Percent of Hispanics in science class—child-level data	N6HISP (SCI Q5), N6CLSZ (composite)	0–100	Recoded to the following: 1=Less than 1%, 2=1% to less than 5%, 3=5% to less than 10%, 4=10% to less than 25%, 5=25% or more
88	G6PMIN	Class	Percent of minorities in reading class—child-level data	G6ASIAN, G6HISP, G6BLACK, G6AMRIN, G6RACEO (RDG Q16), G6CLSZ (composite) *Note: There are now separate "PMIN" variables for reading, math, and science.*	0–100	Recoded to the following: 1=Less than 10%, 2=10% to less than 25%, 3=25% to less than 50%, 4=50% to less than 75%, 5=75% or more

See note at end of table.

Table 7-15. Spring-fifth grade composite variables: School year 2003–04—Continued

ID	Variable name	Category	Description	Derived from	Restricted-use file values	Public-use file values
89	M6PMIN	Class	Percent of minorities in mathematics class—child-level data	M6ASIAN, M6HISP, M6BLACK, M6AMRIN, M6RACEO (MTH Q6), M6CLSZ (composite)	0–100	Recoded to the following: 1=Less than 10%, 2=10% to less than 25%, 3=25% to less than 50%, 4=50% to less than 75%, 5=75% or more
90	N6PMIN	Class	Percent of minorities in science class—child-level data	N6ASIAN, N6HISP, N6BLACK, N6AMRIN, N6RACEO (SCI Q5), N6CLSZ (composite)	0–100	Recoded to the following: 1=Less than 10%, 2=10% to less than 25%, 3=25% to less than 50%, 4=50% to less than 75%, 5=75% or more
91	J61TQUEX	Teacher flag	Presence of spring-fifth grade reading teacher data	Receipted reading teacher questionnaires in the FTS *Note: There are now separate "TQUEX" flags for reading and math/science.*	0=False, 1=True	0=False, 1=True
92	J62TQUEX	Teacher flag	Presence of spring-fifth grade mathematics or science teacher data	Receipted mathematics or science teacher questionnaires in the FTS	0=False, 1=True	0=False, 1=True
93	F6MTHSCI	Teacher flag	Whether child is linked to a mathematics or science teacher	Receipted mathematics or science teacher questionnaires in the FTS *Note: This is a new flag.*	1=Math, 2=Science	1=Math, 2=Science
94	T6SAMTCH	Teacher flag	Whether reading and mathematics teacher linked to the child is the same person	J61T_ID (reading teacher ID) and J62T_ID (mathematics or science teacher ID) *Note: This is a new flag*	0=False, 1=True	0=False, 1=True

See note at end of table.

Table 7-15. Spring-fifth grade composite variables: School year 2003–04—Continued

ID	Variable name	Category	Description	Derived from	Restricted-use file values	Public-use file values
95	G6TQUEX	Teacher flag	Presence of child-level spring-fifth grade reading teacher data	Receipted reading teacher questionnaires in the FTS *Note: There are now separate flags for reading, math, and science.*	0=False, 1=True	0=False, 1=True
96	M6TQUEX	Teacher flag	Presence of child-level spring-fifth grade mathematics teacher data	Receipted mathematics teacher questionnaires in the FTS	0=False, 1=True	0=False, 1=True
97	N6TQUEX	Teacher flag	Presence of child-level spring-fifth grade science teacher data	Receipted science teacher questionnaires in the FTS	0=False, 1=True	0=False, 1=True
98	D6SETQA	Teacher flag	Presence or Absence of Special Ed A data	Receipted special education instrument A in the FTS	0 =False, 1=True	Suppressed variable
99	E6SETQB	Teacher flag	Presence or Absence of Special Ed B data	Receipted special education instrument B in the FTS	0 =False, 1=True	Suppressed variable
100	R6REGION	School	Indicates the geographic region of the child's school	CREGION, R3REGION, R4REGION, R6REGION (composites), CCP and PSS files	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, HA, OR, WA	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, HA, OR, WA

See note at end of table.

Table 7-15. Spring-fifth grade composite variables: School year 2003–04—Continued

ID	Variable name	Category	Description	Derived from	Restricted-use file values	Public-use file values
101	R6URBAN	School	Locality type for school—7 category version	KURBAN, R3URBAN, R4URBAN, R6URBAN (composites), CCD and PSS files	1=Large city – a central city of Consolidated Metropolitan Statistical Area (CMSA) with a pop. Greater to or equal to 250,000; 2=Mid-size city – a central city of a CMSA or Metropolitan Statistical Area (MSA) with a pop. Less than 250,000; 3= Large suburb; urban fringe of large city – any incorporated place, Census Designated Place, or nonplace territory within a CMSA or MSA of a large city and defined as urban by the U.S. Census Bureau; 4 = Mid-size suburb; urban fringe of mid-size city – any incorporated place, Census Designated Place, or nonplace territory within a CMSA or MSA of a mid-size city and defined as urban by the U.S. Census Bureau; 5= Large town – an incorporated place or Census Designated Place with a pop. Greater than	1=Large city – a central city of Consolidated Metropolitan Statistical Area (CMSA) with a pop. Greater to or equal to 250,000; 2=Mid-size city – a central city of a CMSA or Metropolitan Statistical Area (MSA) with a pop. Less than 250,000; 3= Large suburb; urban fringe of large city – any incorporated place, Census Designated Place, or nonplace territory within a CMSA or MSA of a large city and defined as urban by the U.S. Census Bureau; 4 = Mid-size suburb; urban fringe of mid-size city – any incorporated place, Census Designated Place, or nonplace territory within a CMSA or MSA of a mid-size city and defined as urban by the U.S. Census Bureau; 5= Large town – an incorporated place or Census Designated Place with a pop. Greater than

See note at end of table.

Table 7-15. Spring-fifth grade composite variables: School year 2003–04—Continued

ID	Variable name	Category	Description	Derived from	Restricted-use file values	Public-use file values
101	R6URBAN (continued)	School	Locality type for school—7 category version	KURBAN, R3URBAN, R4URBAN, R6URBAN (composites), CCD and PSS files	or equal to 25,000 and located outside a CMSA or MSA; 6=Small town – an incorporated place or Census Designated Place with a pop. Less than 25,000 and greater than 2,500 – located outside a CMSA or MSA; 7=Rural – any incorporated place, Census Designated Place, or nonplace territory	or equal to 25,000 and located outside a CMSA or MSA; 6=Small town – an incorporated place or Census Designated Place with a pop. Less than 25,000 and greater than 2,500 – located outside a CMSA or MSA; 7=Rural – any incorporated place, Census Designated Place, or nonplace territory
102	R6LOCALE	School	Locality type for school—8 category version	R3LOCALE, R4LOCALE, R6LOCALE (composites), PSS and CCD files	1=Large city – a central city of Consolidated Metropolitan Statistical Area (CMSA) with a pop. Greater to or equal to 250,000; 2=Mid-size city – a central city of a CMSA or Metropolitan Statistical Area (MSA) with a pop. Less than 250,000; 3= Large suburb; urban fringe of large city – any incorporated place, Census Designated Place, or nonplace territory within a CMSA or MSA of a large city and defined as urban by the U.S.	Suppressed variable

See note at end of table.

Table 7-15. Spring-fifth grade composite variables: School year 2003–04—Continued

ID	Variable name	Category	Description	Derived from	Restricted-use file values	Public-use file values
102	R6LOCALE (continued)	School	Locality type for school—8 category version	R3LOCALE, R4LOCALE, R6LOCALE (composites), PSS and CCD files	Census Bureau; 4 = Mid-size suburb; urban fringe of mid-size city – any incorporated place, Census Designated Place, or nonplace territory within a CMSA or MSA of a mid-size city and defined as urban by the U.S. Census Bureau; 5= Large town – an incorporated place or Census Designated Place with a pop. Greater than or equal to 25,000 and located outside a CMSA or MSA; 6=Small town – an incorporated place or Census Designated Place with a pop. Less than 25,000 and greater than 2,500 – located outside a CMSA or MSA; 7 = non-MSA Rural – any incorporated place, Census Designated Place, or nonplace territory designated as rural by the U.S. Census Bureau that is not within a MSA; 8 = MSA Rural – any incorporated place, Census Designated Place,	Suppressed variable

See note at end of table.

Table 7-15. Spring-fifth grade composite variables: School year 2003–04—Continued

ID	Variable name	Category	Description	Derived from	Restricted-use file values	Public-use file values
102	R6LOCALE (continued)	School	Locality type for school—8 category version	R3LOCALE, R4LOCALE, R6LOCALE (composites), PSS and CCD files	or nonplace territory designated as rural by the U.S. Census Bureau that is within a MSA	Suppressed variable
103	S6SCTYP	School	School type from the school administrator questionnaire	S6PUBLIC (SAQ Q5), S6CATHOL, S6OTHREL, S6NAISKL, S6OTHPRI, S6PVTSPD, S6PVTEAR (all SAQ Q7), CS_TYPE2, S5SCTYP, S4SCTYP, S3SCTYP, S2KSCTYP (composites), SCHL_TYP (School Master file variable derived from PSS/CCD, not on file)	1=Catholic, 2=Other Religious, 3=Other Private, 4=Public	1=Catholic, 2=Other Religious, 3=Other Private, 4=Public
104	S6PUPRI	School	Public or private school	S6SCTYP (composite)	1=Public, 2=Private	1=Public, 2=Private
105	S6ENRL5	School	Total school fifth-grade enrollment	PSS and CCD data	Continuous	Recoded to the following: 1=0-20, 2=21-40, 3=41-60, 4=61-80, 5=81-100, 6=101-120, 7=121-140, 8=141-160 9=161-180, 10=181 or more

See note at end of table.

Table 7-15. Spring-fifth grade composite variables: School year 2003–04—Continued

ID	Variable name	Category	Description	Derived from	Restricted-use file values	Public-use file values
106	S6ENRLS	School	Total school enrollment	S6ANUMCH (SAQ Q3), PSS and CCD data	1=0–149 students, 2=150–299 students, 3=300–499 students, 4=500–749 students, 5= 750 and above students	1=0–149 students, 2=150–299 students, 3=300–499 students, 4=500–749 students, 5=750 and above students
107	S6MINOR	School	Percentage of minority students in school	PMINOR (School Master File variable derived from PSS/CCD, not on file), S6ASNPCT, S6HSPPCT, S6BLKPCT, S6INDPCT, S6OTHPCT (all from SAQ Q8)	Continuous	Recoded to the following: 1=Less than 10%, 2=10% to less than 25%, 3=25% to less than 50%, 4=50% to less than 75%, 5=75% or more
108	S6FLCH_I	School	Percentage of students eligible for free lunch in school	S6ELILNC (SAQ Q20), S6ANUMCH (SAQ Q3), CCD data	Continuous	Recoded to 0–95
109	S6RLCH_I	School	Percent of students eligible for reduced price lunch in school	S6ELIRED (SAQ Q20), S6ANUMCH (SAQ Q3), CCD data	Continuous	Recoded to the following: 1=Less than 1%, 2=1% to less than 5%, 3=5% to less than 10%, 4=10% to less than 25%, 5=25% or more
110	S6SCLVL	School	School instructional level	S6PRKNDR, S6KINDER, S6GRADE1, S6SECOND, S6THIRD, S6FOURTH, S6FIFTH, S6SIXTH, S67TH, S68TH, S6NINTH, S6TENTH, S611TH, S612TH (all from SAQ Q4); S5SCLVL, S4SCLVL, S2SCLVL, GRSPAN (School Master file variable derived from PSS/CCD, not on file)	1=Less than first grade; 2=Primary school, 3=Elementary school, 4=Combined school	1=Less than first grade; 2=Primary school, 3=Elementary school, 4=Combined school

See note at end of table.

Table 7-15. Spring-fifth grade composite variables: School year 2003–04—Continued

ID	Variable name	Category	Description	Derived from	Restricted-use file values	Public-use file values
112	S6SCHBMM	School	School Year Starting Date, Month	S6SYRSMM (SAQ Q10), FMS (variable not on file)	1–12	Suppressed variable
113	S6SCHBYY	School	School Year Starting Date, Year	Hard coded to 2003 in the questionnaire	2003	Suppressed variable
114	S6SCHEDD	School	School Year Ending Date, Day	S6SYREDD (SAQ Q10), FMS (variable not on file)	1–31	Suppressed variable
115	S6SCHEMM	School	School Year Ending Date, Month	S6SYREMM (SAQ Q10), FMS (variable not on file)	1–12	Suppressed variable
116	S6SCHEYY	School	School Year Ending Date, Year	Hard coded to 2004 in the questionnaire	2004	2004
117	F6YRRND	School	Year round school	S_YRRNDFLG (FMS variable not on file)	1 = Year round school, 2 = Not year round school	1 = Year round school, 2 = Not year round school
118	K6INFAC	School flag	Presence or absence of facilities checklist data	Receipted facilities checklists in the FTS	0 = False, 1 = True	0 = False, 1 = True
119	S6INSAQ	School flag	Presence or absence of SAQ data	Receipted SAQs in the FTS	0 = False, 1 = True	0 = False, 1 = True
120	U6SRABS	School flag	Presence of spring-fifth grade SRA data	Receipted student record abstracts in the FTS	0 = False, 1 = True	Suppressed variable

SOURCE: U.S. Department of Education, National Center for Education Statistics, Early Childhood Longitudinal Study, Kindergarten Class of 1998–99 (ECLS-K), spring 2004.

Table 7-16. Recoded and suppressed data on the ECLS-K Fifth-Grade Public-Use Data File

Field ID	Variable	Field label	Comment
9	D6T_ID	SPRING 2004 SPECIAL ED TEACHER ID NUMBER	These data suppressed for respondent confidentiality.
20	CS_TYPE2	TYPE OF SCHOOL IN BASE YEAR SAMPLE FRAME	These data recoded for respondent confidentiality.
21	R6REGION	R6 CENSUS REGION	These data recoded for respondent confidentiality.
22	R6URBAN	R6 LOCATION TYPE - 7 CATEGORIES	These data recoded for respondent confidentiality.
23	R6LOCALE	R6 LOCATION TYPE - 8 CATEGORIES	These data suppressed for respondent confidentiality
25	R6FIPSST	R6 SCHOOL FIPS STATE CODE	These data suppressed for respondent confidentiality
26	R6FIPSCT	R6 SCHOOL FIPS COUNTY CODE	These data suppressed for respondent confidentiality
27	R6CCDLEA	R6 CCD LEA\SCHOOL DIST ID (PUBLIC)	These data suppressed for respondent confidentiality
28	R6CCDSID	R6 CCD SCHOOL ID (PUBLIC)	These data suppressed for respondent confidentiality
29	R6STSID	R6 STATE SCHOOL ID (PUBLIC)	These data suppressed for respondent confidentiality
30	R6SCHZIP	R6 SCHOOL ZIP CODE	These data suppressed for respondent confidentiality
31	R6SCHPIN	R6 SCHOOL PIN (PRIVATE)	These data suppressed for respondent confidentiality
34	R6DOBYY	R6 CHILD COMPOSITE DOB YEAR	These data recoded to a maximum value of 1993 and a minimum value of 1992 for respondent confidentiality.
37	R6AGE	R6 COMPOSITE CHILD ASSESSMENT AGE(MNTHS)	These data recoded for respondent confidentiality.
69	D6SETQA	D6 SP ED PART A QUESTIONNAIRE COMPLETED	These data suppressed for respondent confidentiality.
70	E6SETQB	E6 SP ED PART B QUESTIONNAIRE COMPLETED	These data suppressed for respondent confidentiality.
71	U6SRABS	U6 STUDENT RECORDS ABSTRACT COMPLETED	These data suppressed for respondent confidentiality.
323	G6CLSZ	G6 NUMBER OF STUDENTS IN CLASS	These data recoded to a maximum value of 35 and a minimum value of 10 for respondent confidentiality.
324	G6PBLK	G6 PERCENT OF BLACKS IN CLASS	These data recoded for respondent confidentiality.
325	G6PHIS	G6 PERCENT OF HISPANICS IN CLASS	These data recoded for respondent confidentiality.
326	G6PMIN	G6 PERCENT OF MINORITIES IN CLASS	These data recoded for respondent confidentiality.
327	G6PLEP	G6 PERCENT OF LEP STUDENTS IN CLASS	These data recoded for respondent confidentiality.
328	M6CLSZ	M6 NUMBER OF STUDENTS IN CLASS	These data recoded to a maximum value of 35 and a minimum value of 10 for respondent confidentiality.
329	M6PBLK	M6 PERCENT OF BLACKS IN CLASS	These data recoded for respondent confidentiality.
330	M6PHIS	M6 PERCENT OF HISPANICS IN CLASS	These data recoded for respondent confidentiality.
331	M6PMIN	M6 PERCENT OF MINORITIES IN CLASS	These data recoded for respondent confidentiality.
332	M6PLEP	M6 PERCENT OF LEP STUDENTS IN CLASS	These data recoded for respondent confidentiality.
333	N6CLSZ	N6 NUMBER OF STUDENTS IN CLASS	These data recoded to a maximum value of 35 and a minimum value of 10 for respondent confidentiality.
334	N6PBLK	N6 PERCENT OF BLACKS IN CLASS	These data recoded for respondent confidentiality.
335	N6PHIS	N6 PERCENT OF HISPANICS IN CLASS	These data recoded for respondent confidentiality.
336	N6PMIN	N6 PERCENT OF MINORITIES IN CLASS	These data recoded for respondent confidentiality.
337	N6PLEP	N6 PERCENT OF LEP STUDENTS IN CLASS	These data recoded for respondent confidentiality.
340	S6SCHBMM	S6 SCHOOL YEAR BEGINNING DATE MONTH	These data suppressed for respondent confidentiality.
341	S6SCHBDD	S6 SCHOOL YEAR BEGINNING DATE DAY	These data suppressed for respondent confidentiality.
342	S6SCHBYY	S6 SCHOOL YEAR BEGINNING DATE YEAR	These data suppressed for respondent confidentiality.
343	S6SCHEMM	S6 SCHOOL YEAR ENDING DATE MONTH	These data suppressed for respondent confidentiality.
344	S6SCHEDD	S6 SCHOOL YEAR ENDING DATE DAY	These data suppressed for respondent confidentiality.
347	S6ENRL5	S6 TOTAL SCHOOL FIFTH GRADE ENROLLMENT	These data recoded for respondent confidentiality.
348	S6ENRLS	S6 TOTAL SCHOOL ENROLLMENT	These data recoded for respondent confidentiality.
349	S6MINOR	S6 PERCENT MINORITY STUDENTS	These data recoded for respondent confidentiality.
352	S6FLCH_I	S6 IMPUTED % FREE LUNCH ELIGIBLE	These data recoded for respondent confidentiality.
354	S6RLCH_I	S6 IMPUTED % REDUCED LUNCH ELIGIBLE	These data recoded for respondent confidentiality.

See note at end of table.

Table 7-16. Recoded and suppressed data on the ECLS-K Fifth-Grade Public-Use Data File—Continued

Field ID	Variable	Field label	Comment
358	C6INGRAD	C6 AIQ030 GRADE CHILD REPORTED	These data recoded for respondent confidentiality.
378	C6FIFTH	C6 ACQ005 CHILD IN FIFTH GRADE	These data suppressed for respondent confidentiality.
389	C6SPECAC	C6 ACQ270 SPECIAL ACCOMMODATION LISTED	These data suppressed for respondent confidentiality.
405	C6ATTMPT	C6 CHILD ASSESSMENT NUMBER OF ATTEMPTS	These data recoded to a maximum value of 3 for respondent confidentiality.
406	P6HOMZIP	P6 HOME ZIP CODE	These data suppressed for respondent confidentiality.
971	P6MTHBTH	P6 COQ005 COUNTRY BIOLOGCAL MTHR WAS BRN	These data recoded for respondent confidentiality.
974	P6FTBRTH	P6 COQ020 COUNTRY BIOLOGCAL FTHR WAS BRN	These data recoded for respondent confidentiality.
982	P6DGNATT	P6 CHQ060 1ST DIAGNOSIS-LEARNING ABILITY	These data suppressed for respondent confidentiality.
983	P6YYDIAG	P6 CHQ075 YR AT 1ST DIAGNOSIS-LRN ABLTY	These data suppressed for respondent confidentiality.
989	P6PROFFD	P6 CHQ110 IF ACTIVITY PROBLEM DIAGNOSED	These data suppressed for respondent confidentiality.
990	P6DGNACT	P6 CHQ120 WHAT 1ST DIAGNOSIS - ACTIVITY	These data suppressed for respondent confidentiality.
991	P6YYDIA2	P6 CHQ135 YR AT 1ST DIAGNOSIS-ACTIVITY	These data suppressed for respondent confidentiality.
997	P6YYDIA4	P6 CHQ185 YEAR AT 1ST DIAGNOSIS-SPEECH	These data suppressed for respondent confidentiality.
1003	P6YYDIA5	P6 CHQ225 YR AT 1ST DIAGNOSIS-HEARING	These data suppressed for respondent confidentiality.
1006	P6HEARS	P6 CHQ230 DEGREE OF CHILD’S DEAFNESS	These data suppressed for respondent confidentiality.
1007	P6HEARAI	P6 CHQ240 IF CHILD WEARS HEARING AID	These data suppressed for respondent confidentiality.
1008	P6COCHLE	P6 CHQ250 IF CHILD HAS COCHLEAR IMPLANTS	These data suppressed for respondent confidentiality.
1009	P6IMPLNT	P6 CHQ251 YEAR OF IMPLANT	These data suppressed for respondent confidentiality.
1010	P6IMPT02	P6 CHQ252 WAS IT BEFORE 2002	These data suppressed for respondent confidentiality.
1011	P6IMPELM	P6 CHQ253 WAS IT BEFORE ELEM SCHOOL	These data suppressed for respondent confidentiality.
1012	P6CLRUSE	P6 CHQ254 USE OF COCHLEAR IMPLANT IN SCH	These data suppressed for respondent confidentiality.
1013	P6HEARS2	P6 CHQ260 DEVICE EFFECT ON CHD’S HEARING	These data suppressed for respondent confidentiality.
1017	P6DIA6YY	P6 CHQ313 YR AT 1ST DIAGNOSIS-VISION	These data suppressed for respondent confidentiality.
1020	P6CORREC	P6 CHQ316 IF CHD’S VISION IS CORRECTABLE	These data suppressed for respondent confidentiality.
1021	P6BESTEY	P6 CHQ320 WHAT CAN CHILD BEST SEE	These data suppressed for respondent confidentiality.
1025	P6DIABEH	P6 CHQ335 BEHAVIOR PROBLEM DIAGNOSED	These data suppressed for respondent confidentiality.
1026	P6DGNBEH	P6 CHQ337 1ST DIAGNOSIS-BEHAVIOR	These data suppressed for respondent confidentiality.
1027	P6DGBEYY	P6 CHQ345 YR AT 1ST DIAGNOSIS-BEHAVIOR	These data suppressed for respondent confidentiality.
1038	P6SPECIL	P6 CHQ510 IF CHD USES SPECIAL EQUIPMENT	These data suppressed for respondent confidentiality.
1042	P6SERVRV	P6 CHQ536 SERVICES RCVD BEFORE 2002	These data suppressed for respondent confidentiality.
1043	P6SRVRCV	P6 CHQ537 SRVCS RCVD BEFORE ELEM SCHOOL	These data suppressed for respondent confidentiality.
1045	P6SPECND	P6 CHQ545 CHILD SPECIAL NEEDS/EDUCATION	These data suppressed for respondent confidentiality.
1097	P6FMTHRS	P6 CHQ780 REASON FOR FAMILY THERAPY	These data suppressed for respondent confidentiality.
1194	P6RECFRE	P6 WPQ215 DOES CHILD REC FREE REDUCED BF	These data suppressed for respondent confidentiality.
1195	P6FRERED	P6 WPQ216 FREE OR REDUCED BREAKFAST	These data suppressed for respondent confidentiality.
1201	P6HOWPAY	P6 PAQ137 HOW MUCH PAID IN TUITION ($)	These data recoded for respondent confidentiality.
1247	J61INET	J61 Q8A # OF COMPUTERS WITH INTERNET	These data recoded to a maximum value of 8 for respondent confidentiality.
1248	J61COMUS	J61 Q8B # OF COMPUTERS CHILDREN USE	These data recoded to a maximum value of 8 for respondent confidentiality.
1329	J61TGEND	J61 Q28 TEACHER’S GENDER	These data suppressed for respondent confidentiality.
1330	J61YRBOR	J61 Q29 TEACHER’S YEAR OF BIRTH	These data recoded to a maximum value of 1981 and a minimum value of 1940 for respondent confidentiality.
1331	J61HISP	J61 Q30 HISPANIC OR LATINO	These data suppressed for respondent confidentiality.
1332	J61RACE1	J61 Q31A AMERICAN INDIAN / ALASKA NATIVE	These data suppressed for respondent confidentiality.

See note at end of table.

Table 7-16. Recoded and suppressed data on the ECLS-K Fifth-Grade Public-Use Data File—Continued

Field ID	Variable	Field label	Comment
1333	J61RACE2	J61 Q31B ASIAN	These data suppressed for respondent confidentiality.
1334	J61RACE3	J61 Q31C BLACK OR AFRICAN AMERICAN	These data suppressed for respondent confidentiality.
1335	J61RACE4	J61 Q31D NATIVE HAWAIIAN OR OTHER PAC IS	These data suppressed for respondent confidentiality.
1337	J61YRSTC	J61 Q32 NUMBER YEARS BEEN SCHOOL TEACHER	These data recoded to a maximum value of 35 and a minimum value of 1 for respondent confidentiality.
1338	J61YRSGR	J61 Q33 YEARS TAUGHT THIS GRADE	These data recoded to a maximum value of 27 and a minimum value of 1 for respondent confidentiality.
1339	J61YRSCH	J61 Q34 YRS TCHR TAUGHT AT THIS SCHOOL	These data recoded to a maximum value of 30 and a minimum value of 1 for respondent confidentiality.
1340	J61HGHST	J61 Q35 HIGHEST ED LVL TEACHER ACHIEVED	These data recoded for respondent confidentiality.
1390	J61MASSI	J61 Q41 MAIN ASSIGNMENT AT SCHOOL	These data recoded for respondent confidentiality.
1391	J61CLORG	J61 Q42 HOW CLASSES ARE ORGANIZED	These data recoded for respondent confidentiality.
1423	J62INET	J62 Q8A # OF COMPUTERS WITH INTERNET	These data recoded to a maximum value of 8 for respondent confidentiality.
1424	J62COMUS	J62 Q8B # OF COMPUTERS CHILDREN USE	These data recoded to a maximum value of 8 for respondent confidentiality.
1505	J62TGEND	J62 Q28 TEACHER’S GENDER	These data suppressed for respondent confidentiality.
1506	J62YRBOR	J62 Q29 TEACHER’S YEAR OF BIRTH	These data recoded to a maximum value of 1981 and a minimum value of 1940 for respondent confidentiality.
1507	J62HISP	J62 Q30 HISPANIC OR LATINO	These data suppressed for respondent confidentiality.
1508	J62RACE1	J62 Q31A AMERICAN INDIAN / ALASKA NATIVE	These data suppressed for respondent confidentiality.
1509	J62RACE2	J62 Q31B ASIAN	These data suppressed for respondent confidentiality.
1510	J62RACE3	J62 Q31C BLACK OR AFRICAN AMERICAN	These data suppressed for respondent confidentiality.
1511	J62RACE4	J62 Q31D NATIVE HAWAIIAN OR OTHER PAC IS	These data suppressed for respondent confidentiality.
1513	J62YRSTC	J62 Q32 NUMBER YEARS BEEN SCHOOL TEACHER	These data recoded to a maximum value of 35 and a minimum value of 1 for respondent confidentiality.
1514	J62YRSGR	J62 Q33 YEARS TAUGHT THIS GRADE	These data recoded to a maximum value of 27 and a minimum value of 1 for respondent confidentiality.
1515	J62YRSCH	J62 Q34 YRS TCHR TAUGHT AT THIS SCHOOL	These data recoded to a maximum value of 30 and a minimum value of 1 for respondent confidentiality.
1516	J62HGHST	J62 Q35 HIGHEST ED LVL TEACHER ACHIEVED	These data recoded for respondent confidentiality.
1566	J62MASSI	J62 Q41 MAIN ASSIGNMENT AT SCHOOL	These data recoded for respondent confidentiality.
1567	J62CLORG	J62 Q42 HOW CLASSES ARE ORGANIZED	These data recoded for respondent confidentiality.
1580	G6GRENRL	G6 Q1 GRADE CHILD IS ENROLLED	These data suppressed for respondent confidentiality.
1583	G6TT1ENG	G6 Q3B TITLE 1 ENGLISH/LANGUAGE ARTS	These data suppressed for respondent confidentiality.
1584	G6TT1CMB	G6 Q3C TITLE 1 COMBINED SUBJECTS	These data suppressed for respondent confidentiality.
1585	G6TT1ES	G6 Q3D TITLE 1 ESL/BILINGUAL	These data suppressed for respondent confidentiality.
1586	G6TT1SP	G6 Q3E TITLE 1 HANDICAPPED/SPECIAL ED	These data suppressed for respondent confidentiality.
1589	G6PLLESL	G6 Q4C PULL-OUT ESL PROGRAM	These data suppressed for respondent confidentiality.
1590	G6INCESL	G6 Q4D IN-CLASS ESL	These data suppressed for respondent confidentiality.
1592	G6GFTRD	G6 Q4F GIFTED PROGRAM IN READING	These data suppressed for respondent confidentiality.
1595	G6MENTOR	G6 Q4I MEET W/MENTOR NOT PROF PSYCH	These data suppressed for respondent confidentiality.

See note at end of table.

Table 7-16. Recoded and suppressed data on the ECLS-K Fifth-Grade Public-Use Data File—Continued

Field ID	Variable	Field label	Comment
1599	G6ACCOM	G6 Q8 SPECIAL TEST ACCOMMODATIONS	These data suppressed for respondent confidentiality.
1605	G6PROMOT	G6 Q10 RECOMMEND PROMOTION/YR END	These data suppressed for respondent confidentiality.
1608	G6LNGTM	G6 Q13 LENGTH OF TIME IN READ CLASS	These data suppressed for respondent confidentiality.
1611	G6ASIAN	G6 Q16A # ASIAN/PACIFIC ISLANDERS READ	These data recoded to a maximum value of 7 for respondent confidentiality.
1612	G6HISP	G6 Q16B # HISPANICS (ALL RACES) READ	These data recoded to a maximum value of 11 for respondent confidentiality.
1613	G6BLACK	G6 Q16C # NON-HISPANIC BLACKS READ	These data recoded to a maximum value of 16 for respondent confidentiality.
1614	G6WHITE	G6 Q16D # NON-HISPANIC WHITES READ	These data recoded to a maximum value of 27 for respondent confidentiality.
1615	G6AMRIN	G6 Q16E # AMERICAN INDIANS READ	These data recoded to a maximum value of 2 for respondent confidentiality.
1616	G6RACEO	G6 Q16F # OF STUDENTS OTHER RACES READ	These data recoded to a maximum value of 2 for respondent confidentiality.
1617	G6TOTRA	G6 Q16G TOTAL ENROLLMENT (RACES) READ	These data recoded to a maximum value of 36 and a minimum value of 10 for respondent confidentiality.
1618	G6BOYS	G6 Q17A NUMBER OF BOYS IN READ CLASS	These data recoded to a maximum value of 19 and a minimum value of 4 for respondent confidentiality.
1619	G6GIRLS	G6 Q17B NUMBER OF GIRLS IN READ CLASS	These data recoded to a maximum value of 20 and a minimum value of 4 for respondent confidentiality.
1620	G6TOTGEN	G6 Q17C TOTAL ENROLLMENT (GENDER) READ	These data recoded to a maximum value of 36 and a minimum value of 10 for respondent confidentiality.
1621	G6GIFT	G6 Q18A # GIFTED/TALENTED IN READ CLASS	These data recoded to a maximum value of 9 for respondent confidentiality.
1622	G6NUMLE	G6 Q18B # LEP STUDENTS IN READ CLASS	These data recoded to a maximum value of 9 for respondent confidentiality.
1623	G6DISAB	G6 Q18C NUMBER WITH DISABILITIES READ	These data recoded to a maximum value of 9 for respondent confidentiality.
1669	G6IENG	G6 Q32A READ INSTRUCTION- ENGLISH	These data suppressed for respondent confidentiality.
1670	G6ISPNH	G6 Q32B READ INSTRUCTION - SPANISH	These data suppressed for respondent confidentiality.
1671	G6IASN	G6 Q32C READ INSTRUCTION - ASIAN LNG	These data suppressed for respondent confidentiality.
1672	G6IOTLN	G6 Q32D READ INSTRUCTION - OTHER LNG	These data suppressed for respondent confidentiality.
1673	G6LNGOS	G6 Q32D LANGUAGE OF READ INSTRUCTION	These data suppressed for respondent confidentiality.
1689	M6GFTMTH	M6 Q1C GIFTED PROGRAM IN MATHEMATICS	These data suppressed for respondent confidentiality.
1693	M6LNGTM	M6 Q4 LENGTH OF TIME IN MATHEMATICS CLASS	These data suppressed for respondent confidentiality.
1694	M6GRMTH	M6 Q5 GRADE OF CHILDREN IN MATHEMATICS CLASS	These data suppressed for respondent confidentiality.
1695	M6ASIAN	M6 Q6A # ASIAN/PACIFIC ISLANDERS MATH	These data recoded to a maximum value of 7 for respondent confidentiality.
1696	M6HISP	M6 Q6B # HISPANICS (ALL RACES) MATH	These data recoded to a maximum value of 11 for respondent confidentiality.
1697	M6BLACK	M6 Q6C # NON-HISPANIC BLACKS MATH	These data recoded to a maximum value of 16 for respondent confidentiality.
1698	M6WHITE	M6 Q6D # NON-HISPANIC WHITES MATH	These data recoded for to a maximum value of 27 respondent confidentiality.

See note at end of table.

Table 7-16. Recoded and suppressed data on the ECLS-K Fifth-Grade Public-Use Data File—Continued

Field ID	Variable	Field label	Comment
1699	M6AMRIN	M6 Q6E # AMERICAN INDIANS MATH	These data recoded to a maximum value of 2 for respondent confidentiality.
1700	M6RACEO	M6 Q6F # OF STUDENTS OTHER RACES MATH	These data recoded to a maximum value of 2 for respondent confidentiality.
1701	M6TOTRA	M6 Q6G TOTAL ENROLLMENT (RACES) MATH	These data recoded to a maximum value of 36 and minimum value of 10 for respondent confidentiality.
1702	M6BOYS	M6 Q7A NUMBER OF BOYS IN MATHEMATICS CLASS	These data recoded to a maximum value of 19 and minimum value of 4 for respondent confidentiality.
1703	M6GIRLS	M6 Q7B NUMBER OF GIRLS IN MATHEMATICS CLASS	These data recoded to a maximum value of 20 and minimum value of 4 for respondent confidentiality.
1704	M6TOTGEN	M6 Q7C TOTAL ENROLLMENT (GENDER) MATH	These data recoded to a maximum value of 36 and minimum value of 10 for respondent confidentiality.
1705	M6GIFT	M6 Q8A # GIFTED/TALENTED IN MATHEMATICS CLASS	These data recoded to a maximum value of 9 for respondent confidentiality.
1706	M6NUMLE	M6 Q8B # LEP STUDENTS IN MATHEMATICS CLASS	These data recoded to a maximum value of 9 for respondent confidentiality.
1707	M6DISAB	M6 Q8C NUMBER WITH DISABILITIES MATH	These data recoded to a maximum value of 9 for respondent confidentiality.
1743	M6IENG	M6 Q17A MATHEMATICS INSTRUCTION- ENGLISH	These data suppressed for respondent confidentiality.
1744	M6ISPNH	M6 Q17B MATHEMATICS INSTRUCTION - SPANISH	These data suppressed for respondent confidentiality.
1745	M6IASN	M6 Q17C MATHEMATICS INSTRUCTION - ASIAN LNG	These data suppressed for respondent confidentiality.
1746	M6IOTLN	M6 Q17D MATHEMATICS INSTRUCTION - OTHER LNG	These data suppressed for respondent confidentiality.
1747	M6LNGOS	M6 Q17D LANGUAGE OF MATHEMATICS INSTRUCTION	These data suppressed for respondent confidentiality.
1760	N6LNGTM	N6 Q3 LENGTH OF TIME IN SCIE CLASS	These data suppressed for respondent confidentiality.
1761	N6GRDSCI	N6 Q4 GRADE OF CHILDREN IN SCIENCE CLASS	These data suppressed for respondent confidentiality.
1762	N6ASIAN	N6 Q5A # ASIAN/PACIFIC ISLANDERS SCIE	These data recoded to a maximum value of 7 for respondent confidentiality.
1763	N6HISP	N6 Q5B # HISPANICS (ALL RACES) SCIE	These data recoded to a maximum value of 11 for respondent confidentiality.
1764	N6BLACK	N6 Q5C # NON-HISPANIC BLACKS SCIE	These data recoded to a maximum value of 16 for respondent confidentiality.
1765	N6WHITE	N6 Q5D # NON-HISPANIC WHITES SCIE	These data recoded to a maximum value of 27 for respondent confidentiality.
1766	N6AMRIN	N6 Q5E # AMERICAN INDIANS SCIE	These data recoded to a maximum value of 2 for respondent confidentiality.
1767	N6RACEO	N6 Q5F # OF STUDENTS OTHER RACES SCIE	These data recoded to a maximum value of 2 for respondent confidentiality.
1768	N6TOTRA	N6 Q5G TOTAL ENROLLMENT (RACES) SCIE	These data recoded to a maximum value of 36 and a minimum value of 10 for respondent confidentiality.
1769	N6BOYS	N6 Q6A NUMBER OF BOYS IN SCIE CLASS	These data recoded to a maximum value of 19 and a minimum value of 4 for respondent confidentiality.
1770	N6GIRLS	N6 Q6B NUMBER OF GIRLS IN SCIE CLASS	These data recoded to a maximum value of 20 and a minimum value of 4 for respondent confidentiality.
1771	N6TOTGEN	N6 Q6C TOTAL ENROLLMENT (GENDER) SCIE	These data recoded to a maximum value of 36 and a minimum value of 10 for respondent confidentiality.

See note at end of table.

Table 7-16. Recoded and suppressed data on the ECLS-K Fifth-Grade Public-Use Data File—Continued

Field ID	Variable	Field label	Comment
1772	N6GIFT	N6 Q7A # GIFTED/TALENTED IN SCIE CLASS	These data recoded to a maximum value of 9 for respondent confidentiality.
1773	N6NUMLE	N6 Q7B # LEP STUDENTS IN SCIE CLASS	These data recoded to a maximum value of 9 for respondent confidentiality.
1774	N6DISAB	N6 Q7C NUMBER WITH DISABILITIES SCIE	These data recoded to a maximum value of 9 for respondent confidentiality.
1804	N6IENG	N6 Q13A SCIE INSTRUCTION- ENGLISH	These data suppressed for respondent confidentiality.
1805	N6ISPNH	N6 Q13B SCIE INSTRUCTION - SPANISH	These data suppressed for respondent confidentiality.
1806	N6IASN	N6 Q13C SCIE INSTRUCTION - ASIAN LNG	These data suppressed for respondent confidentiality.
1807	N6IOTLN	N6 Q13D SCIE INSTRUCTION - OTHER LNG	These data suppressed for respondent confidentiality.
1808	N6LNGOS	N6 Q13D LANGUAGE OF SCIE INSTRUCTION	These data suppressed for respondent confidentiality.
1812	S6DAYSYR	S6 Q1 DAYS IN SCH YR	These data recoded for respondent confidentiality.
1813	S6ADA	S6 Q2 % AVERAGE DAILY ATTENDANCE FOR YR	These data recoded for respondent confidentiality.
1815	S6ANUMCH	S6 Q3A # ENROLLED AROUND 10/1/2003	These data recoded for respondent confidentiality.
1816	S6BNUMCH	S6 Q3B # ENROLLED SINCE 10/1/2003	These data recoded for respondent confidentiality.
1817	S6CNUMCH	S6 Q3C # WHO LEFT - DIDN'T RETURN	These data recoded to a maximum value of 150 for respondent confidentiality.
1818	S6UNGRAD	S6 Q4A GRADE LEVEL-UNGRADED	These data suppressed for respondent confidentiality.
1821	S6KINDER	S6 Q4D GRADE LEVEL-KINDERGARTEN	These data suppressed for respondent confidentiality.
1822	S6GRADE1	S6 Q4E GRADE LEVEL-FIRST GRADE	These data suppressed for respondent confidentiality.
1823	S6SECOND	S6 Q4F GRADE LEVEL-SECOND GRADE	These data suppressed for respondent confidentiality.
1824	S6THIRD	S6 Q4G GRADE LEVEL-THIRD GRADE	These data suppressed for respondent confidentiality.
1825	S6FOURTH	S6 Q4H GRADE LEVEL-FOURTH GRADE	These data suppressed for respondent confidentiality.
1830	S6NINTH	S6 Q4M GRADE LEVEL-NINTH GRADE	These data suppressed for respondent confidentiality.
1831	S6TENTH	S6 Q4N GRADE LEVEL-TENTH GRADE	These data suppressed for respondent confidentiality.
1832	S611TH	S6 Q4O GRADE LEVEL-ELEVENTH GRADE	These data suppressed for respondent confidentiality.
1833	S612TH	S6 Q4P GRADE LEVEL-TWELFTH GRADE	These data suppressed for respondent confidentiality.
1835	S6REGSKL	S6 Q6A IS IT REGULAR PUBLIC SCHOOL	These data suppressed for respondent confidentiality.
1836	S6MAGSKL	S6 Q6B IS IT A MAGNET SCHOOL	These data suppressed for respondent confidentiality.
1838	S6BIASKL	S6 Q6D IS IT A TRIBAL SCHOOL	These data suppressed for respondent confidentiality.
1839	S6SPEDSK	S6 Q6E IS IT A SPECIAL ED SCHOOL	These data suppressed for respondent confidentiality.
1844	S6PRIVRD	S6 Q7A4 IS IT A PRIVATE ORDER	These data suppressed for respondent confidentiality.
1846	S6NAISKL	S6 Q7C PRIVATE-ACCREDITED BY NAIS	These data suppressed for respondent confidentiality.
1847	S6OTHPRI	S6 Q7D IS IT OTHER PRIVATE	These data suppressed for respondent confidentiality.
1848	S6PVTSPD	S6 Q7E IS IT SPECIAL EDUCATION	These data suppressed for respondent confidentiality.
1850	S6ASNPCT	S6 Q8A PERCENT OF ASIAN STUDENTS	These data suppressed for respondent confidentiality.
1852	S6HSPPCT	S6 Q8B PERCENT OF HISPANIC STUDENTS	These data recoded for respondent confidentiality.
1854	S6BLKPCT	S6 Q8C PERCENT OF BLACK STUDENTS	These data recoded for respondent confidentiality.
1856	S6WHTPCT	S6 Q8D PERCENT OF WHITE STUDENTS	These data suppressed for respondent confidentiality.
1858	S6INDPCT	S6 Q8E PERCENT OF AMERICAN INDIANS	These data suppressed for respondent confidentiality.
1860	S6OTHPCT	S6 Q8F PERCENT OF OTHER STUDENTS	These data recoded for respondent confidentiality.
1862	S6LEPSCH	S6 Q9A PERCENT OF LEP CHILDREN	These data recoded for respondent confidentiality.
1863	S6LEPFIF	S6 Q9B % LEP IN FIFTH GRADE	These data recoded for respondent confidentiality.
1864	S6SYRSMM	S6 Q10A SCH START MONTH	These data suppressed for respondent confidentiality.
1865	S6SYRSDD	S6 Q10B SCH START DAY	These data suppressed for respondent confidentiality.
1866	S6SYREMM	S6 Q10D SCH END MONTH	These data suppressed for respondent confidentiality.

See note at end of table.

Table 7-16. Recoded and suppressed data on the ECLS-K Fifth-Grade Public-Use Data File—Continued

Field ID	Variable	Field label	Comment
1867	S6SYREDD	S6 Q10E SCH END DAY	These data suppressed for respondent confidentiality.
1868	S6AMBUSF	S6 Q11 TIME FIRST BUS AM	These data recoded for respondent confidentiality.
1869	S6AMBUSL	S6 Q12 TIME LAST BUS AM	These data recoded for respondent confidentiality.
1870	S6STRTAM	S6 Q13 OFFICIAL SCHOOL START TIME AM	These data recoded for respondent confidentiality.
1878	S6BRKSTR	S6 Q16A TIME BREAKFAST START	These data recoded for respondent confidentiality.
1879	S6BRKEND	S6 Q16B TIME BREAKFAST END	These data recoded for respondent confidentiality.
1882	S6PRABRK	S6 Q19A2 PARTICIPATE ANY SCH BREAKFAST	These data suppressed for respondent confidentiality.
1883	S6ELIBRK	S6 Q19B1 ELIGIBLE FOR FREE BREAKFAST	These data suppressed for respondent confidentiality.
1884	S6PARBRK	S6 Q19B2 PARTICIPATES IN BREAKFAST	These data suppressed for respondent confidentiality.
1885	S6ELRPBK	S6 Q19C1 ELIGIBLE RED-PRICE BREAKFAST	These data suppressed for respondent confidentiality.
1886	S6PARPBK	S6 Q19C2 PARTICIPATE RED-PRICE BREAKFAST	These data suppressed for respondent confidentiality.
1887	S6PAALUN	S6 Q20A2 PARTICIPATE ANY SCH LUNCH	These data suppressed for respondent confidentiality.
1888	S6ELILNC	S6 Q20B1 ELIGIBLE FOR FREE LUNCH	These data suppressed for respondent confidentiality.
1889	S6PARLNC	S6 Q20B2 PARTICIPATES IN FREE LUNCH	These data suppressed for respondent confidentiality.
1890	S6ELIRED	S6 Q20C1 ELIGIBLE IN REDUCED-PRICE LUNCH	These data suppressed for respondent confidentiality.
1891	S6PARRED	S6 Q20C2 PARTICIPATES IN RED-PRICE LUNCH	These data suppressed for respondent confidentiality.
1904	S6LBRYOK	S6 Q24C DOES LIBRARY MEET NEEDS	These data suppressed for respondent confidentiality.
1908	S6PLAYOK	S6 Q24G DOES PLAYGROUND MEET NEEDS	These data suppressed for respondent confidentiality.
1912	S6PORTBL	S6 Q25 # PORTABLE CLASSROOMS	These data recoded to a maximum value of 15 for respondent confidentiality.
1959	S6NOTEST	S6 Q34 NO GRADE TESTED	These data suppressed for respondent confidentiality.
1971	S6RTCHFL	S6 Q38A1 # REG CLASSROOM TCHR-FULL	These data recoded for respondent confidentiality.
1972	S6RTCHPT	S6 Q38A2 # REG CLASSROOM TCHR-PART	These data recoded to a maximum value of 5 for respondent confidentiality.
1973	S6MSARFL	S6 Q38B1 # GYM DRAMA MUSIC ART TCHR-FULL	These data recoded to a maximum value of 6 for respondent confidentiality.
1974	S6MSARPT	S6 Q38B2 # GYM DRAMA MUSIC ART TCHR-PART	These data recoded to a maximum value of 6 for respondent confidentiality.
1975	S6SPEDFL	S6 Q38C1 # SPECIAL ED TCHR-FULL	These data recoded to a maximum value of 11 for respondent confidentiality.
1976	S6SPEDPT	S6 Q38C2 # SPECIAL ED TCHR-PART	These data recoded to a maximum value of 6 for respondent confidentiality.
1977	S6ESLFL	S6 Q38D1 # ESL/BILINGUAL TCHR-FULL	These data recoded to a maximum value of 8 for respondent confidentiality.
1978	S6ESLPT	S6 Q38D2 # ESL/BILINGUAL TCHR-PART	These data recoded to a maximum value of 2 for respondent confidentiality.
1979	S6READFL	S6 Q38E1 # READING TCHR/SPECIALIST-FULL	These data recoded to a maximum value of 5 for respondent confidentiality.
1980	S6READPT	S6 Q38E2 # READING TCHR/SPECIALIST-PART	These data recoded to a maximum value of 4 for respondent confidentiality.
1981	S6GIFTFL	S6 Q38F1 # GIFTED/TALENTED TCHR-FULL	These data recoded to a maximum value of 5 for respondent confidentiality.
1982	S6GIFTPT	S6 Q38F2 # GIFTED/TALENTED TCHR-PART	These data recoded to a maximum value of 2 for respondent confidentiality.

See note at end of table.

Table 7-16. Recoded and suppressed data on the ECLS-K Fifth-Grade Public-Use Data File—Continued

Field ID	Variable	Field label	Comment
1983	S6NURSFL	S6 Q38G1 # SCH NURSE HEALTH PROF-FULL	These data recoded to a maximum value of 2 for respondent confidentiality.
1984	S6NURSPT	S6 Q38G2 # SCH NURSE HEALTH PROF-PART	These data recoded to a maximum value of 2 for respondent confidentiality.
1985	S6PSYCFL	S6 Q38H1 SCH PSYCH/ SOCIAL WORKER-FULL	These data recoded to a maximum value of 2 for respondent confidentiality.
1986	S6PSYCPT	S6 Q38H2 SCH PSYCH/SOCIAL WORKER-PART	These data recoded to a maximum value of 2 for respondent confidentiality.
1987	S6PARAFL	S6 Q38I1 # PARAPROFESSIONALS-FULL	These data recoded to a maximum value of 20 for respondent confidentiality.
1988	S6PARAPT	S6 Q38I2 # PARAPROFESSIONALS-PART	These data recoded to a maximum value of 14 for respondent confidentiality.
1989	S6LIBRFL	S6 Q38J1 # LIBRARIANS-FULL	These data recoded to a maximum value of 2 for respondent confidentiality.
1990	S6LIBRPT	S6 Q38J2 # LIBRARIANS-PART	These data recoded to a maximum value of 2 for respondent confidentiality.
1992	S6HWLONG	S6 Q39E RESP (NOT PRINCIPAL) YR AT SCH	These data recoded to a maximum value of 15 and a minimum of 1 for respondent confidentiality.
2012	S6BRTHYR	S6 Q43 YEAR PRINCIPAL WAS BORN	These data recoded to a maximum value of 1971 and a minimum of 1941 for respondent confidentiality.
2013	S6ORIGIN	S6 Q44 PRINCIPAL IS HISPANIC/LATINO	These data suppressed for respondent confidentiality.
2014	S6RACE1	S6 Q45A PRINCIPAL IS AMERICAN INDIAN	These data suppressed for respondent confidentiality.
2015	S6RACE2	S6 Q45B PRINCIPAL IS ASIAN	These data suppressed for respondent confidentiality.
2016	S6RACE3	S6 Q45C PRINCIPAL IS BLACK	These data suppressed for respondent confidentiality.
2017	S6RACE4	S6 Q45D PRINCIPAL IS HAWAIIAN OR PAC IS	These data suppressed for respondent confidentiality.
2018	S6RACE5	S6 Q45E PRINCIPAL IS WHITE	These data suppressed for respondent confidentiality.
2019	S6YSTCH	S6 Q46A NUMBER OF YRS TEACHING	These data recoded to a maximum value of 26 and a minimum of 2 for respondent confidentiality.
2020	S6TOTPRI	S6 Q46B NUMBER OF YRS AS PRINCIPAL	These data recoded to a maximum value of 25 and a minimum of 1 for respondent confidentiality.
2021	S6PRINHR	S6 Q46C NUMBER YRS A PRINCIPAL HERE	These data recoded to a maximum value of 15 and a minimum of 1 for respondent confidentiality.
2022	S6EDLVL	S6 Q47 HIGHEST LEVEL OF EDUCATION	These data recoded for respondent confidentiality.
2023	S6MAJOR	S6 Q48 MAJOR FIELD HIGHEST ED LEVEL	These data recoded for respondent confidentiality.
2027	K6GUARDO	K6 Q1A1 OBSERVED SECURITY GUARD	These data suppressed for respondent confidentiality.
2028	K6GUARDS	K6 Q1A2 SCH CNFRM SECURITY GUARD	These data suppressed for respondent confidentiality.
2029	K6METDTO	K6 Q1B1 OBSERVED METAL DETECTORS	These data suppressed for respondent confidentiality.
2030	K6METDTS	K6 Q1B2 SCH CNFRM METAL DETECTORS	These data suppressed for respondent confidentiality.
2032	K6SCAMS	K6 Q1C2 SCH CNFRM SECURITY CAMERAS	These data suppressed for respondent confidentiality.
2034	K6BARSS	K6 Q1D2 SCH CNFRM WINDOW AND DOOR BARS	These data suppressed for respondent confidentiality.
2038	K6FENCES	K6 Q1F2 SCH CNFRM FENCING AROUND SCHOOL	These data suppressed for respondent confidentiality.
2043	K6INTCMO	K6 Q1I1 OBSERVED INTERCOMS	These data suppressed for respondent confidentiality.
2045	K6ALARMO	K6 Q1J1 OBSERVED FIRE ALARMS	These data suppressed for respondent confidentiality.
2046	K6ALARMS	K6 Q1J2 SCH CNFRM FIRE ALARMS	These data suppressed for respondent confidentiality.
2047	K6FREXTO	K6 Q1K1 OBSERVED FIRE EXTINGUISHERS	These data suppressed for respondent confidentiality.

See note at end of table.

Table 7-16. Recoded and suppressed data on the ECLS-K Fifth-Grade Public-Use Data File—Continued

Field ID	Variable	Field label	Comment
2058	D6GENDER	D6 Q1 TEACHER'S GENDER	These data suppressed for respondent confidentiality.
2059	D6YRBORN	D6 Q2 TEACHER'S YEAR OF BIRTH	These data suppressed for respondent confidentiality.
2060	D6HISP	D6 Q3 HISPANIC OR LATINO	These data suppressed for respondent confidentiality.
2061	D6RACE1	D6 Q4 AMERICAN INDIAN OR ALASKA NATIVE	These data suppressed for respondent confidentiality.
2062	D6RACE2	D6 Q4 ASIAN	These data suppressed for respondent confidentiality.
2063	D6RACE3	D6 Q4 BLACK OR AFRICAN AMERICAN	These data suppressed for respondent confidentiality.
2064	D6RACE4	D6 Q4 NATIVE HAWAIIAN OR OTHER PAC ISL	These data suppressed for respondent confidentiality.
2065	D6RACE5	D6 Q4 WHITE	These data suppressed for respondent confidentiality.
2066	D6SCHLYR	D6 Q5 YEARS AT THIS SCHOOL	These data suppressed for respondent confidentiality.
2067	D6SPLYRS	D6 Q6 YEARS WITH SPECIAL ED STUDENTS	These data suppressed for respondent confidentiality.
2068	D6YRSTCH	D6 Q7 TOTAL YEARS TEACHING	These data suppressed for respondent confidentiality.
2069	D6ASSIGN	D6 Q8 TEACHER'S MAIN ASSIGNMENT	These data suppressed for respondent confidentiality.
2070	D6HGHSTD	D6 Q9 HIGHEST ED LEVEL TEACHER ACHIEVED	These data suppressed for respondent confidentiality.
2071	D6EMRGN	D6 Q10A EMERGENCY CREDENTIAL	These data suppressed for respondent confidentiality.
2072	D6PRVSN	D6 Q10B PROVISIONAL CREDENTIAL	These data suppressed for respondent confidentiality.
2073	D6DISSPE	D6 Q10C DISABILITY-SPECIFIC CREDENTIAL	These data suppressed for respondent confidentiality.
2074	D6SPED	D6 Q10D SPECIAL EDUCATION CREDENTIAL	These data suppressed for respondent confidentiality.
2075	D6GNED	D6 Q10E GENERAL EDUCATION CREDENTIAL	These data suppressed for respondent confidentiality.
2076	D6SPCH	D6 Q10F SPEECH/LANGUAGE LICENSE	These data suppressed for respondent confidentiality.
2077	D6PHYST	D6 Q10G PHYSICAL THERAPY LICENSE	These data suppressed for respondent confidentiality.
2078	D6OCCPT	D6 Q10H OCCUPATIONAL THERAPY LICENSE	These data suppressed for respondent confidentiality.
2079	D6CTCLIN	D6 Q10I CERTIF OF CLINICAL COMPETENCE	These data suppressed for respondent confidentiality.
2080	D6OTHPRF	D6 Q10J OTHER PROFESSIONAL LICENSE	These data suppressed for respondent confidentiality.
2081	D6NOCRED	D6 Q10K NO CREDENTIALS/ENDORSEMENTS	These data suppressed for respondent confidentiality.
2082	D6EARLY	D6 Q11A TEACHER'S EARLY EDUCATION COURSE	These data suppressed for respondent confidentiality.
2083	D6ERLSPE	D6 Q11B EARLY CHDHD SPECIAL ED COURSE	These data suppressed for respondent confidentiality.
2084	D6ELEM	D6 Q11C TEACHER'S ELEMENTARY ED COURSES	These data suppressed for respondent confidentiality.
2085	D6SECED	D6 Q11D SECONDARY EDUCATION COURSE	These data suppressed for respondent confidentiality.
2086	D6ESL	D6 Q11E TEACHER'S ESL COLLEGE COURSES	These data suppressed for respondent confidentiality.
2087	D6BILED	D6 Q11F BILINGUAL EDUCATION COURSE	These data suppressed for respondent confidentiality.
2088	D6SPECED	D6 Q11G TEACHER'S SPECIAL ED COURSES	These data suppressed for respondent confidentiality.
2089	D6LRNDIS	D6 Q11H LEARNING DISABILITIES COURSE	These data suppressed for respondent confidentiality.
2090	D6MNTL	D6 Q11I MENTAL RETARDATION COURSE	These data suppressed for respondent confidentiality.
2091	D6ORTHPD	D6 Q11J ORTHOPEDIC IMPAIRMNTS COURSE	These data suppressed for respondent confidentiality.
2092	D6EMTNL	D6 Q11K EMOTIONAL DISTURBAN COURSE	These data suppressed for respondent confidentiality.
2093	D6DEAF	D6 Q11L DEAFNESS COURSE	These data suppressed for respondent confidentiality.
2094	D6BLIND	D6 Q11M BLINDNESS COURSE	These data suppressed for respondent confidentiality.
2095	D6COMDIS	D6 Q11N COMMNCTN DISORDERS COURSE	These data suppressed for respondent confidentiality.
2096	D6INFNT	D6 Q11O DISABLD INFANTS/TODLRS COURSE	These data suppressed for respondent confidentiality.
2097	D6PHYSTH	D6 Q11P PHYSICAL THERAPY COURSE	These data suppressed for respondent confidentiality.
2098	D6OCCTH	D6 Q11Q OCCUPATIONAL THERAPY COURSE	These data suppressed for respondent confidentiality.
2099	D6SCHPSY	D6 Q11R SCHOOL PSYCHOLOGY COURSE	These data suppressed for respondent confidentiality.
2100	D6CLMGMT	D6 Q11S CLASSROOM MANAGEMENT COURSE	These data suppressed for respondent confidentiality.
2101	D6CRPOS2	D6 Q12 CURRENT POSITION IN SCHOOL (R)	These data suppressed for respondent confidentiality.
2102	D6GENED	D6 Q13A WORK IN GENERAL ED ROOM	These data suppressed for respondent confidentiality.

See note at end of table.

Table 7-16. Recoded and suppressed data on the ECLS-K Fifth-Grade Public-Use Data File—Continued

Field ID	Variable	Field label	Comment
2103	D6SPEDRM	D6 Q13B WORK IN A SPECIAL ED ROOM	These data suppressed for respondent confidentiality.
2104	D6NCLSS	D6 Q13C WORK IN NON-CLASSROOM SPACE	These data suppressed for respondent confidentiality.
2105	D6OTHRM	D6 Q13D WORK IN OTHER TYPE OF ROOM	These data suppressed for respondent confidentiality.
2106	D6NODIR	D6 Q13E DON'T WORK W/STUDENT DIRECTLY	These data suppressed for respondent confidentiality.
2107	D6NOSTDN	D6 Q14 NUMBER OF STUDENTS W/ IEPS	These data suppressed for respondent confidentiality.
2108	D6MMCOM	D6 Q15 MONTH QUESTIONNAIRE COMPLETED	These data suppressed for respondent confidentiality.
2109	D6DDCOM	D6 Q15 DAY QUESTIONNAIRE COMPLETED	These data suppressed for respondent confidentiality.
2110	D6YYCOM	D6 Q15 YEAR QUESTIONNAIRE COMPLETED	These data suppressed for respondent confidentiality.
2111	E6SPEIEP	E6 Q1 CURRENT SP ED SERVICE THROUGH IEP	These data suppressed for respondent confidentiality.
2112	E6ENRGR	E6 Q2 CHILD ENROLLMENT GRADE	These data suppressed for respondent confidentiality.
2113	E6FIRIEP	E6 Q3 WHEN DID CHILD FIRST HAVE IEP	These data suppressed for respondent confidentiality.
2114	E6RVRCRD	E6 Q4 REVIEWED CHILD'S SP ED RECORD	These data suppressed for respondent confidentiality.
2115	E6PRMDIS	E6 Q5 STUDENT'S MAIN DISABILITY CATEGORY	These data suppressed for respondent confidentiality.
2116	E6LRNDIS	E6 Q6A SPECIAL ED/LEARNING DISABILITY	These data suppressed for respondent confidentiality.
2117	E6EMTPRB	E6 Q6B SPECIAL ED/EMOTIONAL PROBLEM	These data suppressed for respondent confidentiality.
2118	E6SPCHLN	E6 Q6C SPECIAL ED /SPEECH IMPAIRMENT	These data suppressed for respondent confidentiality.
2119	E6MNTRTR	E6 Q6D SPECIAL ED/MENTAL RETARDATION	These data suppressed for respondent confidentiality.
2120	E6BLNVSL	E6 Q6E SPECIAL ED/VISUAL IMPAIRMENT	These data suppressed for respondent confidentiality.
2121	E6DEAFHH	E6 Q6F SPECIAL ED/HARD OF HEARING	These data suppressed for respondent confidentiality.
2122	E6HLTHIM	E6 Q6G SPECIAL ED/HEALTH IMPAIRMENT	These data suppressed for respondent confidentiality.
2123	E6PHYSIM	E6 Q6H SPECIAL ED/PHYSICAL IMPAIRMNT	These data suppressed for respondent confidentiality.
2124	E6MLTIM	E6 Q6I SPECIAL ED/MULTIPLE IMPAIRMENT	These data suppressed for respondent confidentiality.
2125	E6DFBLND	E6 Q6J SPECIAL ED/DEAF-BLIND	These data suppressed for respondent confidentiality.
2126	E6DEVDLY	E6 Q6K SPECIAL ED/DEV DELAY	These data suppressed for respondent confidentiality.
2127	E6AUTISM	E6 Q6L SPECIAL ED/AUTISM	These data suppressed for respondent confidentiality.
2128	E6BRAIN	E6 Q6M SPECIAL ED/BRAIN INJURY	These data suppressed for respondent confidentiality.
2129	E6SPED	E6 Q7 RECEIVING SP ED OR RELATED SERVCS	These data suppressed for respondent confidentiality.
2130	E6IEPRDG	E6 Q8A IEP GOAL-READING	These data suppressed for respondent confidentiality.
2131	E6IEPMTH	E6 Q8B IEP GOAL-MATHEMATICS	These data suppressed for respondent confidentiality.
2132	E6IEPLNG	E6 Q8C IEP GOAL-LANGUAGE ARTS	These data suppressed for respondent confidentiality.
2133	E6IEPSCI	E6 Q8D IEP GOAL-SCIENCE	These data suppressed for respondent confidentiality.
2134	E6IEPADT	E6 Q8E IEP GOAL-AUDITORY PROCESSING	These data suppressed for respondent confidentiality.
2135	E6IEPLST	E6 Q8F IEP GOAL-LISTENING COMPREHENSION	These data suppressed for respondent confidentiality.
2136	E6IEPORL	E6 Q8G IEP GOAL-ORAL EXPRESSION	These data suppressed for respondent confidentiality.
2137	E6IEPVOC	E6 Q8H IEP GOAL-VOICE/SPEECH ARTICULATN	These data suppressed for respondent confidentiality.
2138	E6IEPLP	E6 Q8I IEP GOAL-LANGUAGE PRAGMATICS	These data suppressed for respondent confidentiality.
2139	E6IEPSOC	E6 Q8J IEP GOAL-SOCIAL SKILLS	These data suppressed for respondent confidentiality.
2140	E6IEPADP	E6 Q8K IEP GOAL-ADAPTIVE BEHAVIOR	These data suppressed for respondent confidentiality.
2141	E6IEPFMS	E6 Q8L IEP GOAL-FINE MOTOR SKILLS	These data suppressed for respondent confidentiality.
2142	E6IEPGMS	E6 Q8M IEP GOAL-GROSS MOTOR SKILLS	These data suppressed for respondent confidentiality.
2143	E6IEPMOB	E6 Q8N IEP GOAL-ORIENTATION+MOBILITY	These data suppressed for respondent confidentiality.
2144	E6IEPOTH	E6 Q8O IEP GOAL-OTHER SPECIFY	These data suppressed for respondent confidentiality.
2145	E6ADLGY	E6 Q9A AUDIOLOGY PROVIDED	These data suppressed for respondent confidentiality.
2146	E6CNSSER	E6 Q9B COUNSELING SERVICES PROVIDED	These data suppressed for respondent confidentiality.
2147	E6OCCTHR	E6 Q9C OCCUPATIONAL THERAPY PROVIDED	These data suppressed for respondent confidentiality.

See note at end of table.

Table 7-16. Recoded and suppressed data on the ECLS-K Fifth-Grade Public-Use Data File—Continued

Field ID	Variable	Field label	Comment
2148	E6PHYTHR	E6 Q9D PHYSICAL THERAPY PROVIDED	These data suppressed for respondent confidentiality.
2149	E6PSYTHR	E6 Q9E PSYCHOLOGICAL SERVICES PR0VIDED	These data suppressed for respondent confidentiality.
2150	E6SCHHLT	E6 Q9F SCHOOL HEALTH SERVICES PROVIDED	These data suppressed for respondent confidentiality.
2151	E6SOCWRK	E6 Q9G SOCIAL WORK SERVICES PROVIDED	These data suppressed for respondent confidentiality.
2152	E6TRNSPR	E6 Q9H SPECIAL TRANSPORT PROVIDED	These data suppressed for respondent confidentiality.
2153	E6LNGTHR	E6 Q9I LANGUAGE THERAPY PROVIDED	These data suppressed for respondent confidentiality.
2154	E6MOBILT	E6 Q9J ORIENTATION SERVICES PROVIDED	These data suppressed for respondent confidentiality.
2155	E6MIBILT	E6 Q9K MOBILITY SERVICES PROVIDED	These data suppressed for respondent confidentiality.
2156	E6REHAB	E6 Q9L REHABILITATION SERVICES PROVIDED	These data suppressed for respondent confidentiality.
2157	E6OTHSER	E6 Q9M OTHER SERVICE PROVIDED	These data suppressed for respondent confidentiality.
2158	E6HRSSPE	E6 Q10 HRS/WK SP ED SCHEDULED FOR CHILD	These data suppressed for respondent confidentiality.
2159	E6ADPPE	E6 Q11A ADAPTIVE PHYSICAL EDUCATION	These data suppressed for respondent confidentiality.
2160	E6CLSAD	E6 Q11B CLASSROOM AIDES	These data suppressed for respondent confidentiality.
2161	E6BRAILE	E6 Q11C INSTRUCTION IN BRAILLE	These data suppressed for respondent confidentiality.
2162	E6INTRPR	E6 Q11D INTERPRETER FOR THE DEAF	These data suppressed for respondent confidentiality.
2163	E6SGNLNG	E6 Q11E INSTRUCTN IN AMERCN SIGN LNG	These data suppressed for respondent confidentiality.
2164	E6MNLENG	E6 Q11F INSTRUCTN IN MANUAL ENGLISH	These data suppressed for respondent confidentiality.
2165	E6CUEDSP	E6 Q11G INSTRUCTION IN CUED SPEECH	These data suppressed for respondent confidentiality.
2166	E6USEBRA	E6 Q11H USE OF BRAILLE INSTRUCTION	These data suppressed for respondent confidentiality.
2167	E6USESGN	E6 Q11I USE OF AMERCN SIGN LNG INSTRUCT	These data suppressed for respondent confidentiality.
2168	E6USECUE	E6 Q11J USE OF MANUAL ENG INSTRUCTION	These data suppressed for respondent confidentiality.
2169	E6USECSP	E6 Q11K USE OF CUED SPEECH INSTRUCTION	These data suppressed for respondent confidentiality.
2170	E6PRMPLC	E6 Q12 PRIMARY PLACEMENT IN GEN ED CLSRM	These data suppressed for respondent confidentiality.
2171	E6SPEDOT	E6 Q13 % TIME SERV OUTSDE GN ED CLSRM	These data suppressed for respondent confidentiality.
2172	E6ONEON1	E6 Q14A ONE-ON-ONE INSTRUCTION	These data suppressed for respondent confidentiality.
2173	E6SMLGRP	E6 Q14B SMALL-GROUP INSTRUCTION	These data suppressed for respondent confidentiality.
2174	E6LRGGRP	E6 Q14C LARGE-GROUP INSTRUCTION	These data suppressed for respondent confidentiality.
2175	E6COPLRN	E6 Q14D COOPERATIVE LEARNING	These data suppressed for respondent confidentiality.
2176	E6PEERTR	E6 Q14E PEER TUTORING	These data suppressed for respondent confidentiality.
2177	E6CMPTR	E6 Q14F COMPUTER-BASED INSTRUCTION	These data suppressed for respondent confidentiality.
2178	E6DIRINS	E6 Q14G DIRECT INSTRUCTION	These data suppressed for respondent confidentiality.
2179	E6COGSTR	E6 Q14H COGNITIVE STRATEGIES	These data suppressed for respondent confidentiality.
2180	E6SMNGT	E6 Q14I SELF-MANAGEMENT	These data suppressed for respondent confidentiality.
2181	E6BMNGT	E6 Q14J BEHAVIOR MANAGEMENT	These data suppressed for respondent confidentiality.
2182	E6NOINS	E6 Q14K DID NOT DELIVER INSTRUCTION	These data suppressed for respondent confidentiality.
2183	E6SGNINT	E6 Q14L THROUGH SIGN INTERPRETER	These data suppressed for respondent confidentiality.
2184	E6DKMTHD	E6 Q14M DON'T KNOW METHODS USED	These data suppressed for respondent confidentiality.
2185	E6GENEDC	E6 Q15A CURRICULUM GENERAL ED CLASSROOM	These data suppressed for respondent confidentiality.
2186	E6SPEDC	E6 Q15B CURRICULUM SPECIAL ED CLASSROOM	These data suppressed for respondent confidentiality.
2187	E6ACHLVL	E6 Q16 GOALS CHILD EXPECTED TO ACHIEVE	These data suppressed for respondent confidentiality.
2188	E6NOTEC	E5 Q17 DID NOT USE ASSIST TECH	These data suppressed for respondent confidentiality.
2189	E6VANS	E6 Q17A VANS, VEHICLES	These data suppressed for respondent confidentiality.
2190	E6WHLCHR	E6 Q17B WHEELCHAIRS	These data suppressed for respondent confidentiality.
2191	E6WHTCN	E6 Q17C WHITE CANES	These data suppressed for respondent confidentiality.
2192	E6ELCTRN	E6 Q17D ELECTRONIC COMMUNICATION AID	These data suppressed for respondent confidentiality.

See note at end of table.

Table 7-16. Recoded and suppressed data on the ECLS-K Fifth-Grade Public-Use Data File—Continued

Field ID	Variable	Field label	Comment
2193	E6NOELC	E6 Q17E NONELECTRONIC COMMUNICATION AID	These data suppressed for respondent confidentiality.
2194	E6HAIDS	E6 Q17F HEARING AIDS	These data suppressed for respondent confidentiality.
2195	E6FMLOOP	E6 Q17G FM LOOPS	These data suppressed for respondent confidentiality.
2196	E6TTYS	E6 Q17H TTYS/TDDS	These data suppressed for respondent confidentiality.
2197	E6IMPLNT	E6 Q17I COCHLEAR IMPLANTS	These data suppressed for respondent confidentiality.
2198	E6CPTN	E6 Q17J REAL TIME CAPTIONING	These data suppressed for respondent confidentiality.
2199	E6BRATXT	E6 Q17K BRAILLE TEXTS	These data suppressed for respondent confidentiality.
2200	E6ELCBRA	E6 Q17L ELECTRONIC BRAILLE DEVICES	These data suppressed for respondent confidentiality.
2201	E6DIGTXT	E6 Q17M DIGITAL TEXTS	These data suppressed for respondent confidentiality.
2202	E6MGNFY	E6 Q17N MAGNIFYING DEVICES	These data suppressed for respondent confidentiality.
2203	E6CCTV	E6 Q17O CCTV	These data suppressed for respondent confidentiality.
2204	E6TAPERC	E6 Q17P TAPE RECORDERS	These data suppressed for respondent confidentiality.
2205	E6CALC	E6 Q17Q CALCULATORS	These data suppressed for respondent confidentiality.
2206	E6ELCSPL	E6 Q17R ELECTRONIC SPELLING DEVICES	These data suppressed for respondent confidentiality.
2207	E6CMPIND	E6 Q17S COMPUTER FOR SOLE USE OF CHILD	These data suppressed for respondent confidentiality.
2208	E6CMPSHR	E6 Q17T COMPUTER SHARED W/OTHR CHILDREN	These data suppressed for respondent confidentiality.
2209	E6CMPRDG	E6 Q17U READING SOFTWARE	These data suppressed for respondent confidentiality.
2210	E6CMPWRT	E6 Q17V WRITING SOFTWARE	These data suppressed for respondent confidentiality.
2211	E6CMPMTH	E6 Q17W MATHEMATICS SOFTWARE	These data suppressed for respondent confidentiality.
2212	E6ADPOTH	E6 Q17X OTHER ASSIST TECH SPCFY	These data suppressed for respondent confidentiality.
2213	E6COMPUT	E6 Q18 CHILD ASSIGNED FULL TIME COMPUTER	These data suppressed for respondent confidentiality.
2214	E6OFTGTC	E6 Q19 FREQ MEET WITH GENERAL ED TCHRS	These data suppressed for respondent confidentiality.
2215	E6LNGTHM	E6 Q20 LENGTH OF GENERAL ED TEACHER MTGS	These data suppressed for respondent confidentiality.
2216	E6OFTPAR	E6 Q21 FREQ COMMUNICATION WITH PARENTS	These data suppressed for respondent confidentiality.
2217	E6EVLPSY	E6 Q22A PSYCHOLOGICAL EVALUATION	These data suppressed for respondent confidentiality.
2218	E6EVLSPC	E6 Q22B SPEECH/LANGUAGE EVALUATION	These data suppressed for respondent confidentiality.
2219	E6EVLVSN	E6 Q22C VISION EVALUATION	These data suppressed for respondent confidentiality.
2220	E6EVLHR	E6 Q22D HEARING EVALUATION	These data suppressed for respondent confidentiality.
2221	E6EVLLD	E6 Q22E LEARNING/EDUCATIONAL EVALUATION	These data suppressed for respondent confidentiality.
2222	E6EVLMS	E6 Q22F MOTOR SKILLS EVALUATION	These data suppressed for respondent confidentiality.
2223	E6EVLAC	E6 Q22G ACADEMICS EVALUATION	These data suppressed for respondent confidentiality.
2224	E6EVLOTH	E6 Q22H OTHER EVALUATION	These data suppressed for respondent confidentiality.
2225	E6GOAL	E6 Q23 PERCENT OF IEP GOALS MET	These data suppressed for respondent confidentiality.
2226	E6IEPNXY	E6 Q24 IEP NEXT YEAR	These data suppressed for respondent confidentiality.
2227	E6MMCOM	E6 Q25 MONTH QUESTIONNAIRE COMPLETED	These data suppressed for respondent confidentiality.
2228	E6DDCOM	E6 Q25 DAY QUESTIONNAIRE COMPLETED	These data suppressed for respondent confidentiality.
2229	E6YYCOM	E6 Q25 YEAR QUESTIONNAIRE COMPLETED	These data suppressed for respondent confidentiality.
2230	U6ALLYR	U6 Q1 IN SCHOOL WHOLE YEAR	These data suppressed for respondent confidentiality.
2231	U6MMENTR	U6 Q2A MONTH ENTERED SCHOOL	These data suppressed for respondent confidentiality.
2232	U6DDENTR	U6 Q2B DAY ENTERED SCHOOL	These data suppressed for respondent confidentiality.
2233	U6YYENTR	U6 Q2C YEAR ENTERED SCHOOL	These data suppressed for respondent confidentiality.
2234	U6MMLEFT	U6 Q3A MONTH LEFT SCHOOL	These data suppressed for respondent confidentiality.
2235	U6DDLEFT	U6 Q3B DAY LEFT SCHOOL	These data suppressed for respondent confidentiality.
2236	U6YYLEFT	U6 Q3C YEAR LEFT SCHOOL	These data suppressed for respondent confidentiality.
2237	U6WHYLFT	U6 Q4 WHY CHILD LEFT SCHOOL	These data suppressed for respondent confidentiality.

See note at end of table.

Table 7-16. Recoded and suppressed data on the ECLS-K Fifth-Grade Public-Use Data File—Continued

Field ID	Variable	Field label	Comment
2238	U6ATNREC	U6 Q6 SCH KEEPS ATTENDANCE RECORDS	These data suppressed for respondent confidentiality.
2239	U6ABSTOT	U6 Q6A CHILD TOTAL ABSENCES	These data recoded for respondent confidentiality.
2240	U6AXABS	U6 Q6B CHILD EXCUSED ABSENCES	These data recoded for respondent confidentiality.
2241	U6AUXABS	U6 Q6C CHILD UNEXCUSED ABSENCES	These data recoded for respondent confidentiality.
2242	U6BTARD	U6 Q7A CHILD TOTAL TARDIES	These data recoded for respondent confidentiality.
2243	U6BXTARD	U6 Q7B CHILD EXCUSED TARDIES	These data recoded for respondent confidentiality.
2244	U6BUXTAR	U6 Q7C CHILD UNEXCUSED TARDIES	These data recoded for respondent confidentiality.
2245	U6IEP	U6 Q8 IEP/IFSP ON FILE	These data suppressed for respondent confidentiality.
2246	U6IEP03	U6 Q9A1 PRESENCE OF 2003-2004 IEP RECORD	These data suppressed for respondent confidentiality.
2247	U6IEP02	U6 Q9B1 PRESENCE OF 2002-2003 IEP RECORD	These data suppressed for respondent confidentiality.
2248	U6IEP01	U6 Q9C1 PRESENCE OF 2001-2002 IEP RECORD	These data suppressed for respondent confidentiality.
2249	U6MM03	U6 Q9A2 MONTH 2003-2004 IEP SIGNED	These data suppressed for respondent confidentiality.
2250	U6MM02	U6 Q9B2 MONTH 2002-2003 IEP SIGNED	These data suppressed for respondent confidentiality.
2251	U6MM01	U6 Q9C2 MONTH 2001-2002 IEP SIGNED	These data suppressed for respondent confidentiality.
2252	U6YY03	U6 Q9A3 YEAR 2003-2004 IEP SIGNED	These data suppressed for respondent confidentiality.
2253	U6YY02	U6 Q9B3 YEAR 2002-2003 IEP SIGNED	These data suppressed for respondent confidentiality.
2254	U6YY01	U6 Q9C3 YEAR 2001-2002 IEP SIGNED	These data suppressed for respondent confidentiality.
2255	U6LRNNG	U6 Q10A LEARNING DISABILITY	These data suppressed for respondent confidentiality.
2256	U6EMTNL	U6 Q10B SERIOUS EMOTIONAL DISTURBANCE	These data suppressed for respondent confidentiality.
2257	U6SPCH	U6 Q10C SPEECH OR LANGUAGE IMPAIRED	These data suppressed for respondent confidentiality.
2258	U6MNTL	U6 Q10D MENTAL RETARDATION	These data suppressed for respondent confidentiality.
2259	U6BLND	U6 Q10E VISUALLY IMPAIRED-BLIND	These data suppressed for respondent confidentiality.
2260	U6DEAF	U6 Q10F HEARING IMPAIRED-DEAF	These data suppressed for respondent confidentiality.
2261	U6HLTH	U6 Q10G HEALTH IMPAIRED	These data suppressed for respondent confidentiality.
2262	U6PHYSCL	U6 Q10H PHYSICALLY IMPAIRED	These data suppressed for respondent confidentiality.
2263	U6MLTIMP	U6 Q10I MULTIPLE IMPAIRMENTS	These data suppressed for respondent confidentiality.
2264	U6BLNDF	U6 Q10J DEAF AND BLIND	These data suppressed for respondent confidentiality.
2265	U6DEVDLY	U6 Q10K DEVELOPMENTAL DELAY	These data suppressed for respondent confidentiality.
2266	U6AUTISM	U6 Q10L AUTISM	These data suppressed for respondent confidentiality.
2267	U6BRAIN	U6 Q10M TRAUMATIC BRAIN INJURY	These data suppressed for respondent confidentiality.
2268	U6OTHDIS	U6 Q10N OTHER DISABILITY	These data suppressed for respondent confidentiality.
2269	U6PLCMNT	U6 Q11 PRIMARY PLACEMNT IN GNRL ED CLSRM	These data suppressed for respondent confidentiality.
2270	U6MMCOM	U6 Q12A MONTH QUESTIONNAIRE COMPLETED	These data suppressed for respondent confidentiality.
2271	U6DDCOM	U6 Q12B DAY QUESTIONNAIRE COMPLETED	These data suppressed for respondent confidentiality.
2272	U6YYCOM	U6 Q12C YEAR QUESTIONNAIRE COMPLETED	These data suppressed for respondent confidentiality.

SOURCE: U.S. Department of Education, National Center for Education Statistics, Early Childhood Longitudinal Study, Kindergarten Class of 1998–99 (ECLS-K), spring 2004.

8. ELECTRONIC CODEBOOK

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 Codebook (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 8.7 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 exactly match what the users see on their own screens.

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.

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, as well as file-specific information on the child catalog in appendix E. When needed, additional user's guides and 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 Codebook 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 codebook, 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.

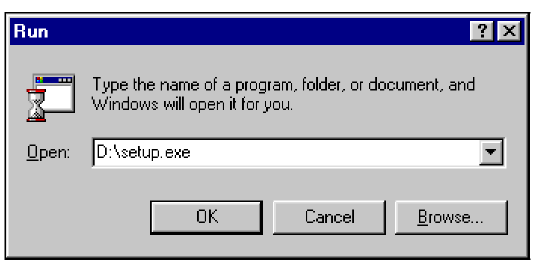
8.2.1 Installing the ECB Program on Your Personal Computer

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



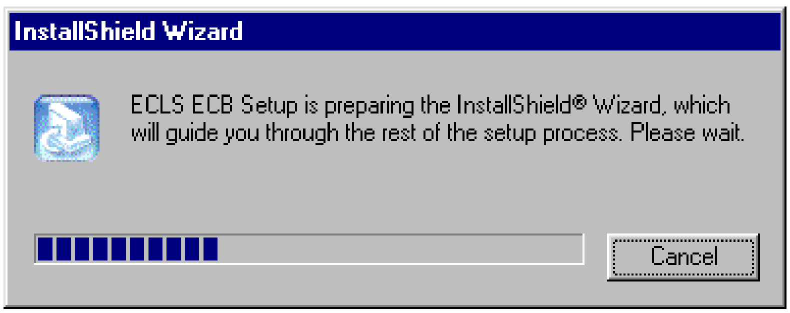
5. Click on the OK button to start the installation. You will now see several installation screens, some of which will prompt you for a response.

   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."

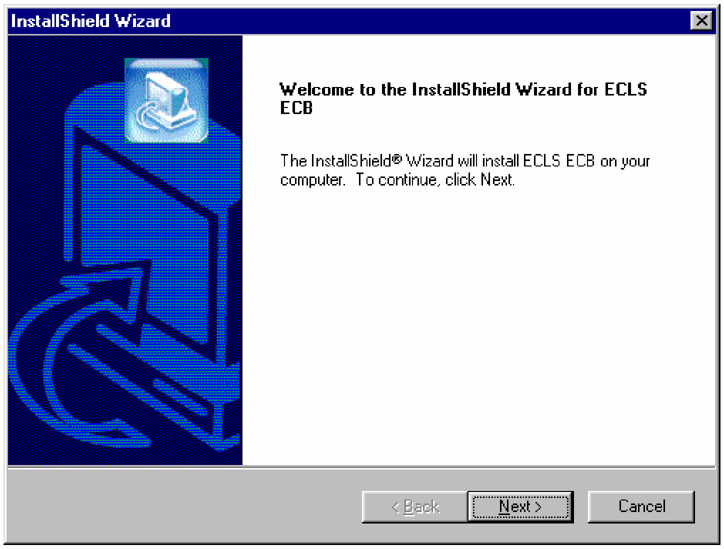
6. The screen shown in exhibit 8-2 indicates that the setup is being prepared.

Exhibit 8-2. InstallShield Wizard



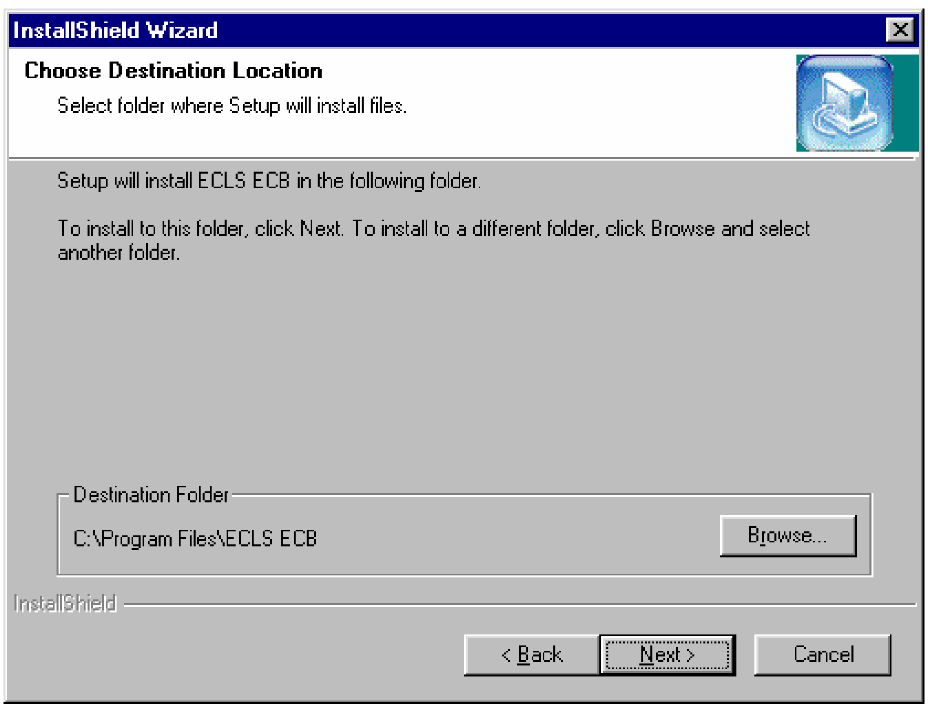
7. You will be prompted to continue with the installation in the Welcome window shown in exhibit 8-3. Click on the Next button to continue.

Exhibit 8-3. Welcome window



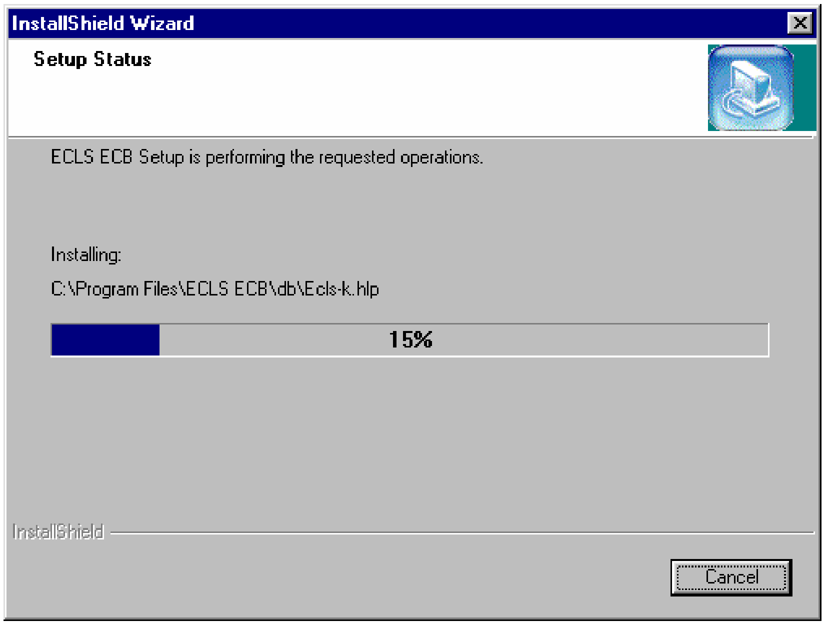
8. When you continue, you will be prompted to choose a destination location for the installation in the window shown in exhibit 8-4. If you wish to change the destination location, click on the Browse button to change the directory. Click on the Next button when the desirable destination folder is shown.

Exhibit 8-4. Choose Destination Location



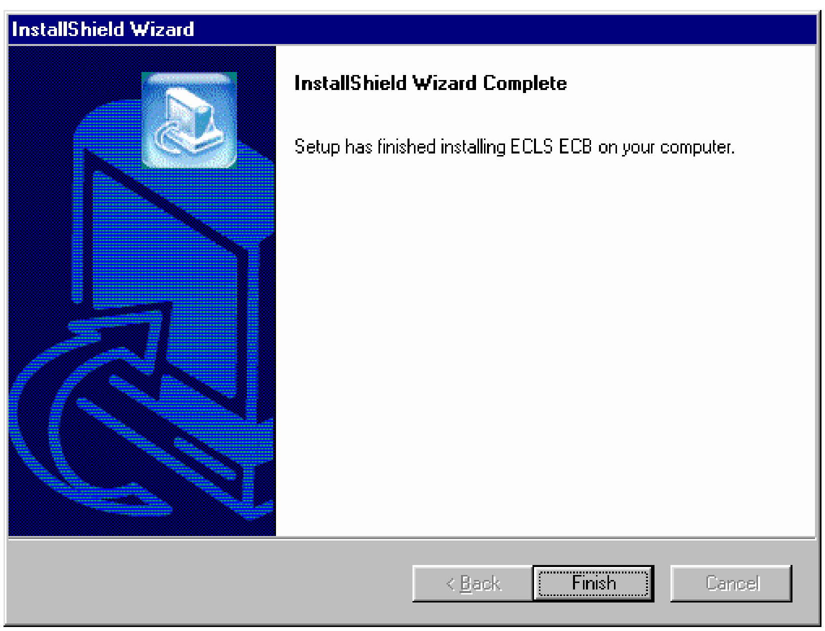
9. Setup will then start installing files. Exhibit 8-5 shows the setup status.

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 step 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, click on the ECB desktop icon (exhibit 8-7a) shown below to invoke the program. Alternatively, on the desktop screen, click on the Start button and then point to Programs (exhibit 8-7b). Click on the ECB title to invoke the program.

Exhibit 8-7a. Desktop icon

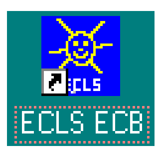
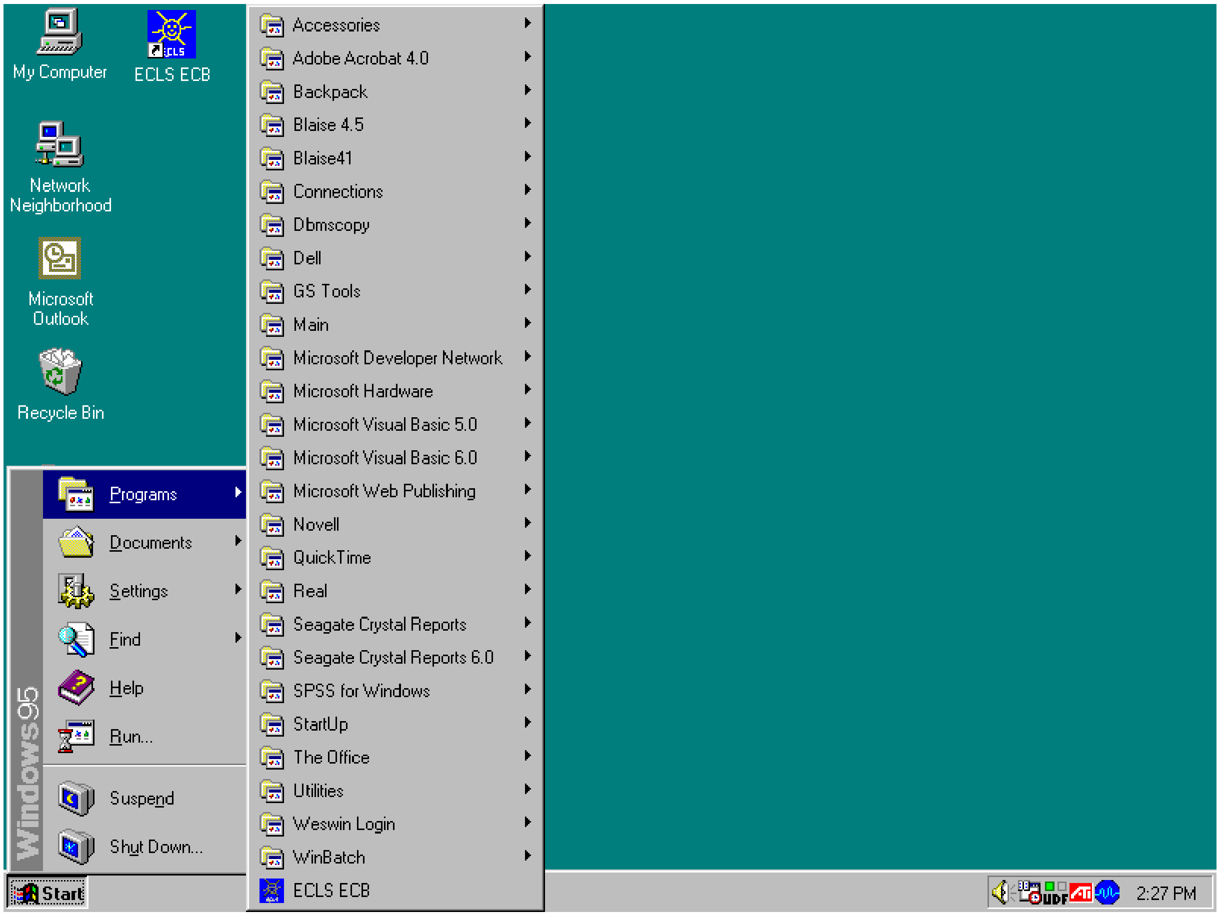


Exhibit 8-7b. Desktop screen—click start

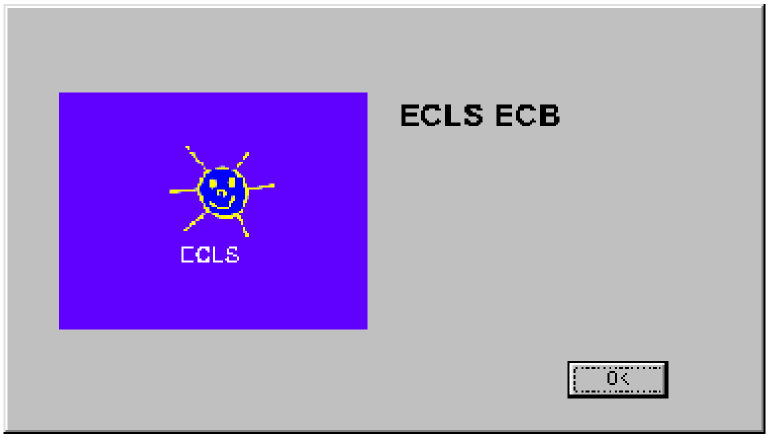


2. If you are a first-time user of the ECB, exhibit 8-8 will appear and ask if you are a new ECB user.

Exhibit 8-8. First-time user dialog box

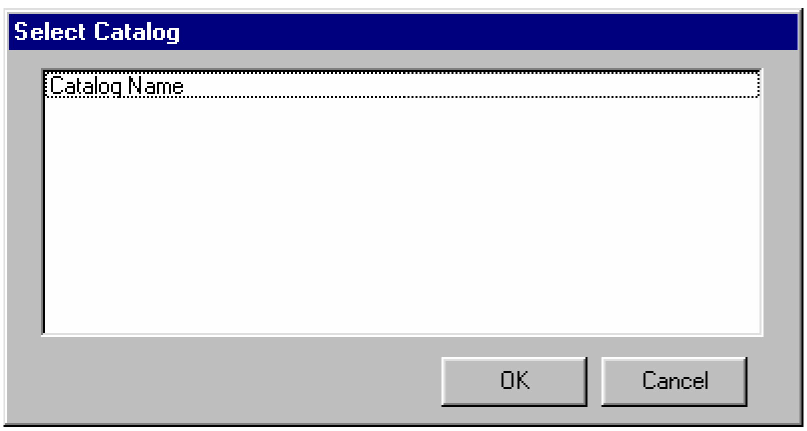
3. Click “Yes” if you are a first-time user. The ECB splash-screen shown in exhibit 8-9 will appear.

Exhibit 8-9. ECB splash screen



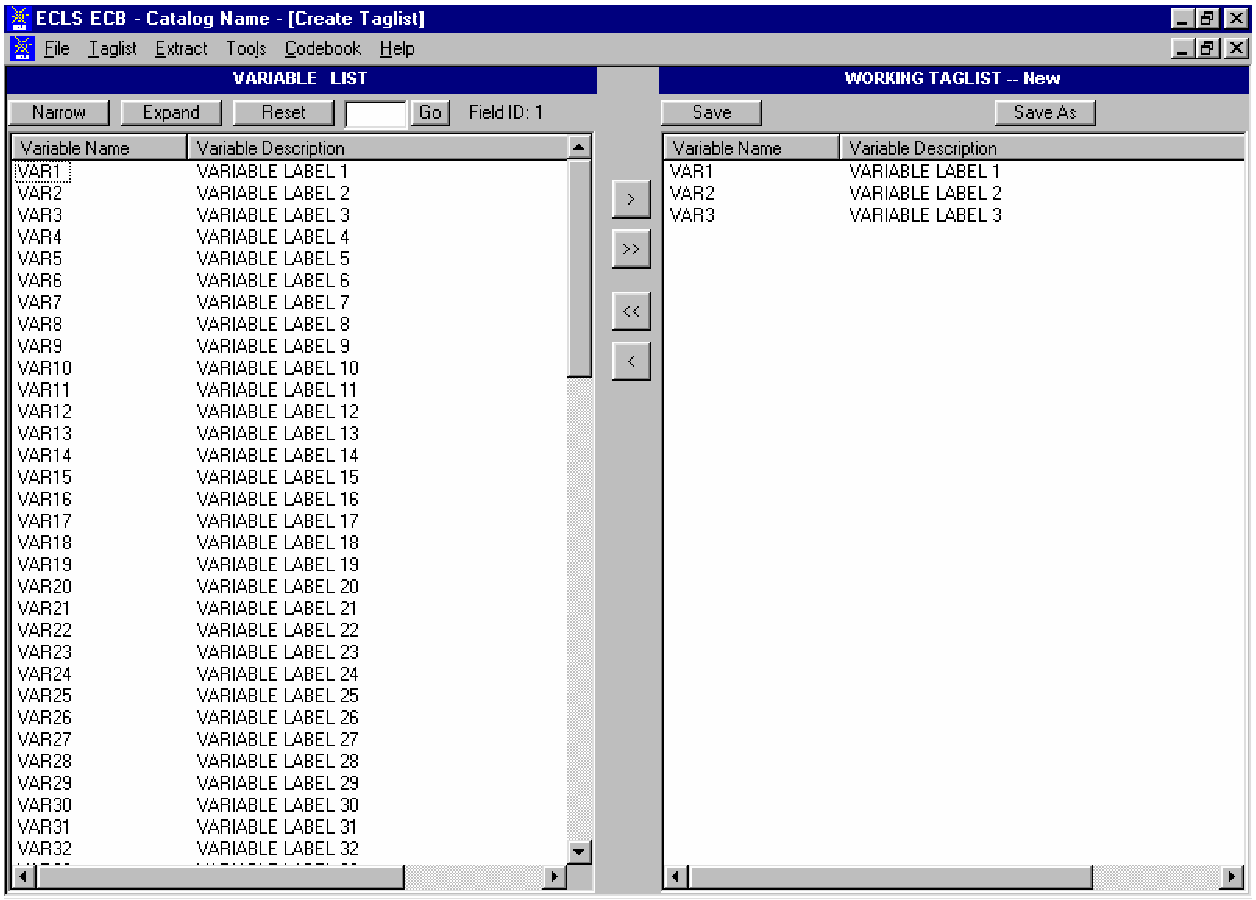
4. On the Select Catalog screen (exhibit 8-10), highlight the name of the catalog. (The fifth-grade ECB has only one catalog.)

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.

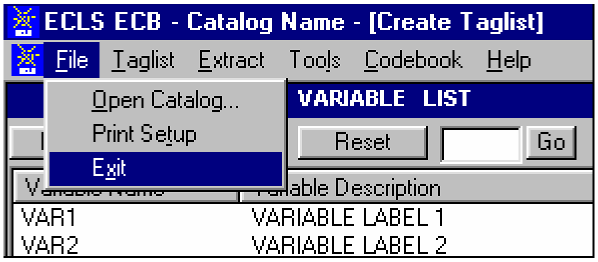
8.2.3 Exiting the ECB

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:

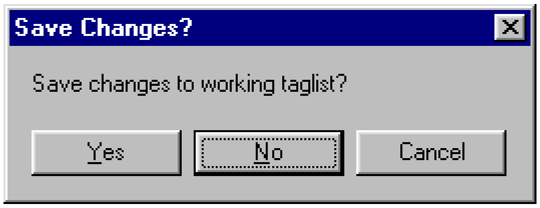
1. From the File menu, click on the Exit option as shown in exhibit 8-12.

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 8.7 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 underscore 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 underscored 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

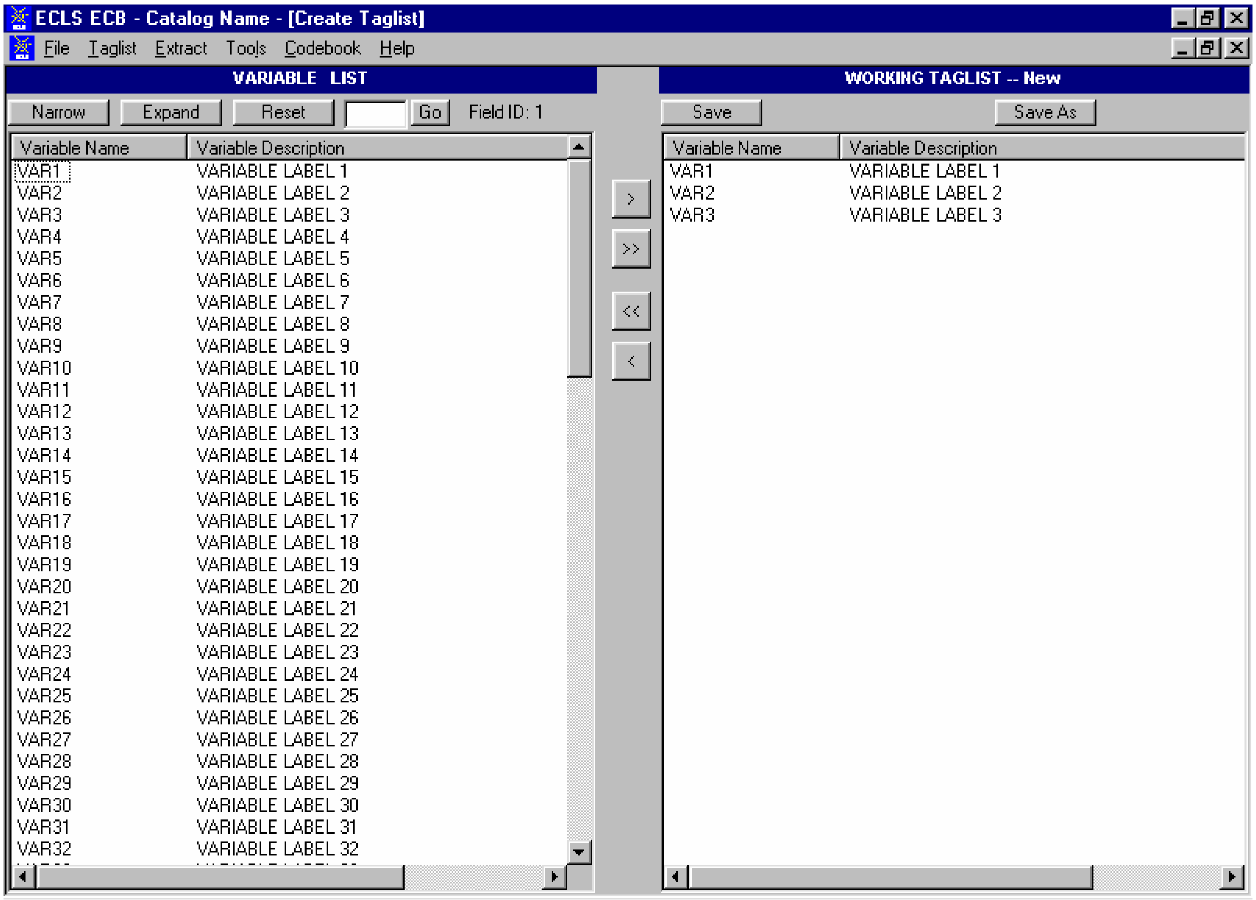
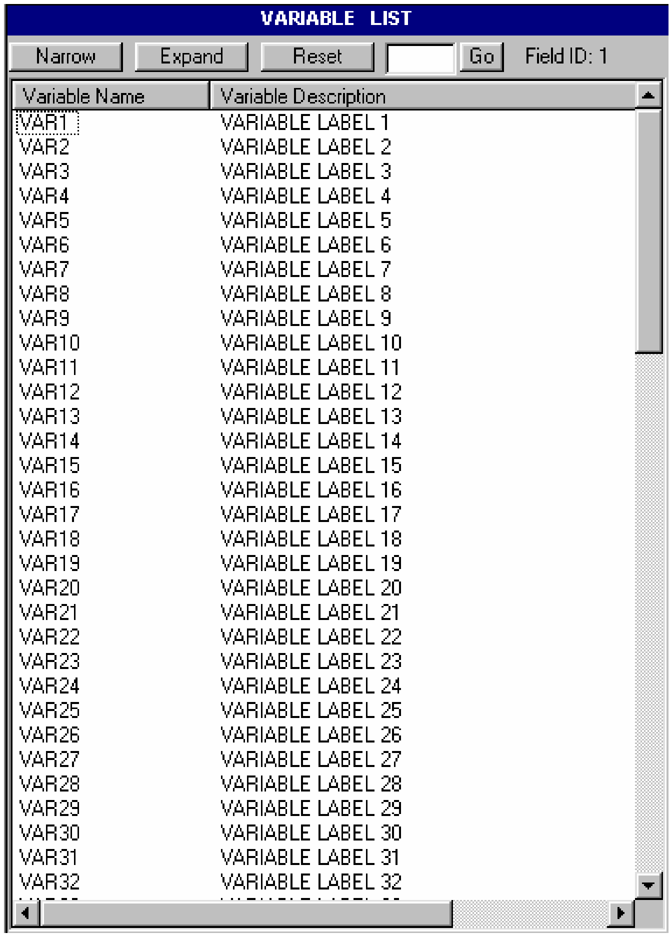


Exhibit 8-17. Variable List



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 codebook 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 Codebook 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 appendix E on the CD-ROM.

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:

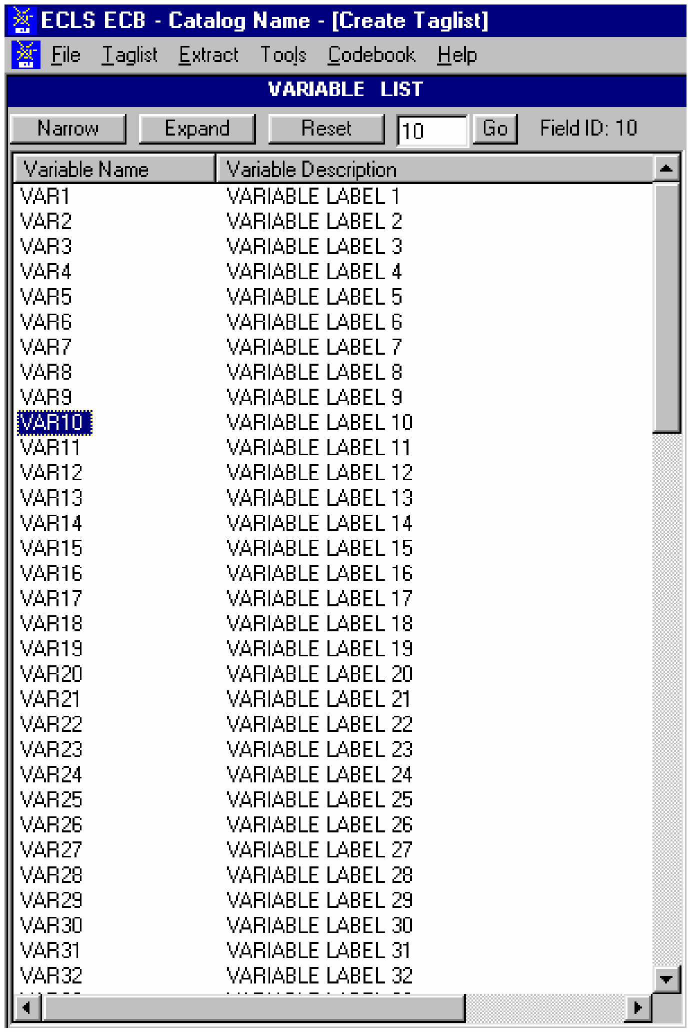
1. Type the field ID in the input box on the left of the Go button.
2. Click on the Go button.
3. The Variable List will then scroll down automatically to show the selected variable.
4. The selected variable is highlighted.
5. The field ID of the current variable selected is shown on the right of the 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 appendix E on the CD-ROM 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

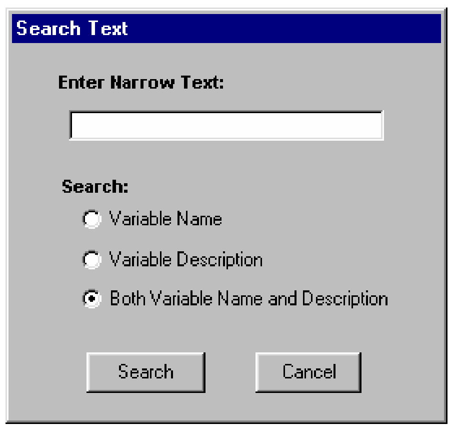
The Narrow Search function can be used to narrow the list of variables displayed in the Variable List. Since some catalogs have several thousand variables, this feature helps eliminate the variables that do not apply to your analysis. In performing the Narrow Search, you can enter key characters, words, or phrases as your criteria for searching the variable names, variable descriptions, or both. 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.

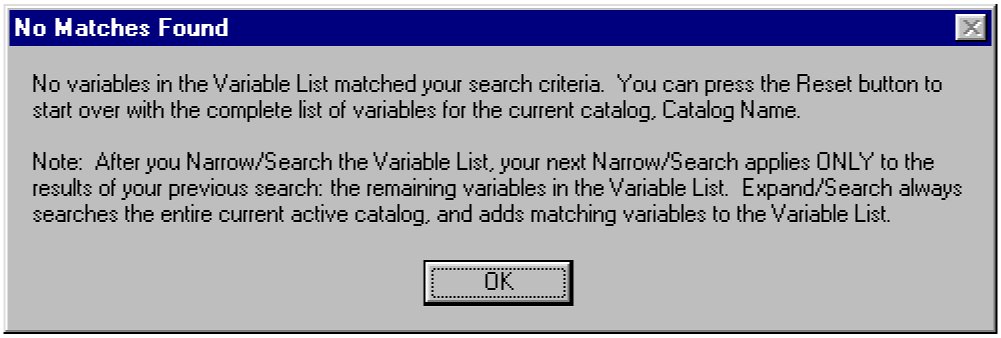
Exhibit 8-19. Narrow Search Text dialog box



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.

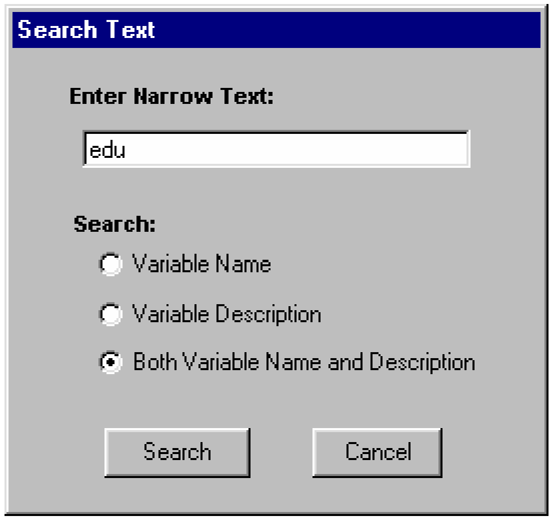
Example of Narrowing a Search

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.

The catalog-specific topical variable groupings can be found in appendix E on the CD-ROM. 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.

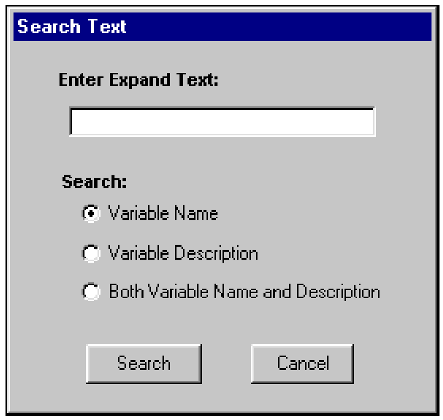
Performing the Expand Search function will only expand the variables listed in the Variable List window and will not affect those in the Working Taglist window.

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.

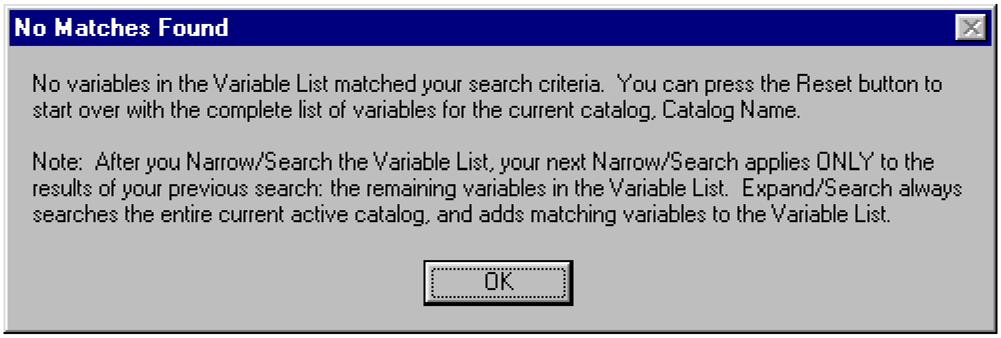
6. The variables meeting the specified criteria will be added to the variables already displayed in the Variables List column.

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



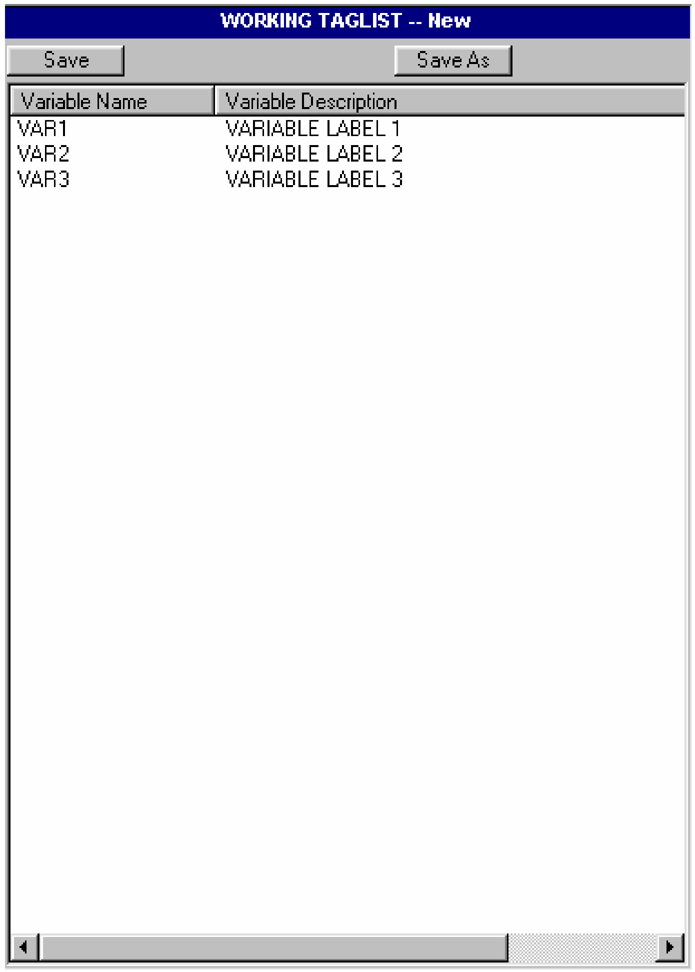
8.3.1.4 Resetting Your Variable List

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 appendix E on the CD-ROM 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.

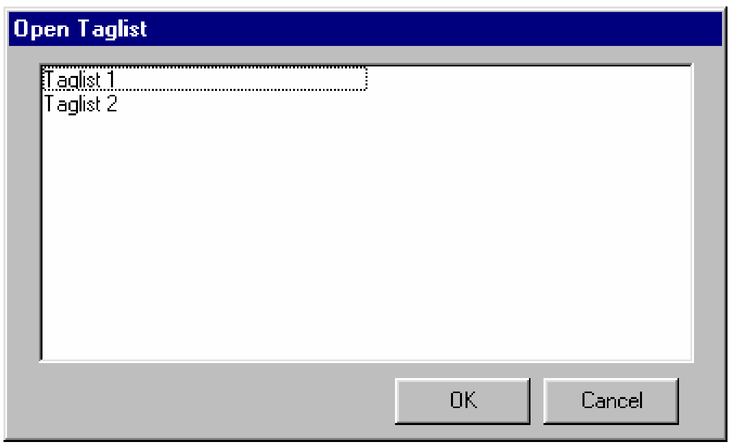
How To Open a Taglist:

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.

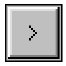
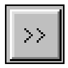
5. Click on the OK button.

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.

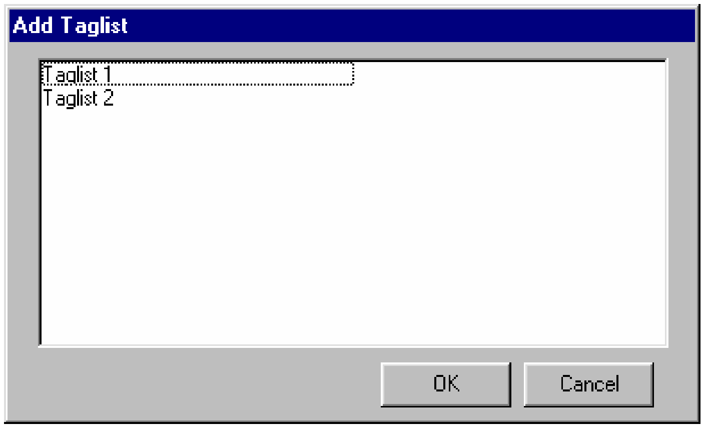
How To Add Variables to a Working Taglist:

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

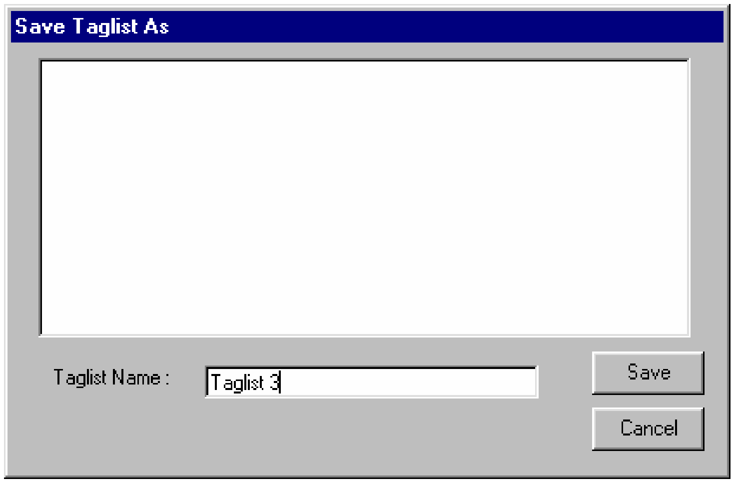
The ECB has the ability to save the newly created or modified taglist displayed in the Working Taglist column. Taglists can be saved either under the name already assigned or under a new name. If you have opened a new taglist and have not yet assigned it a name, you will be presented with the Save As 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.

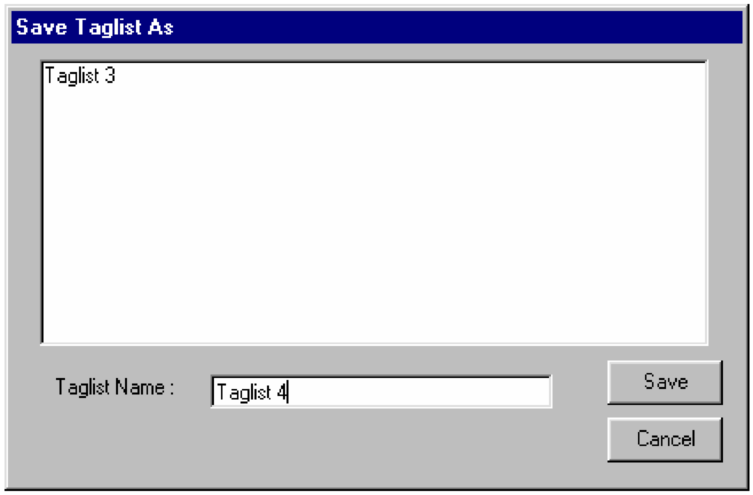
4. Enter the new name of the taglist in the Taglist Name field.

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" only 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 (exhibit 8-32).
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

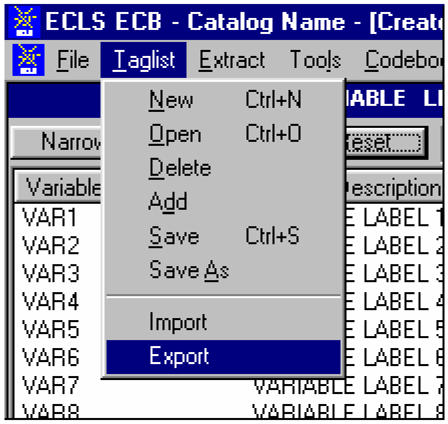
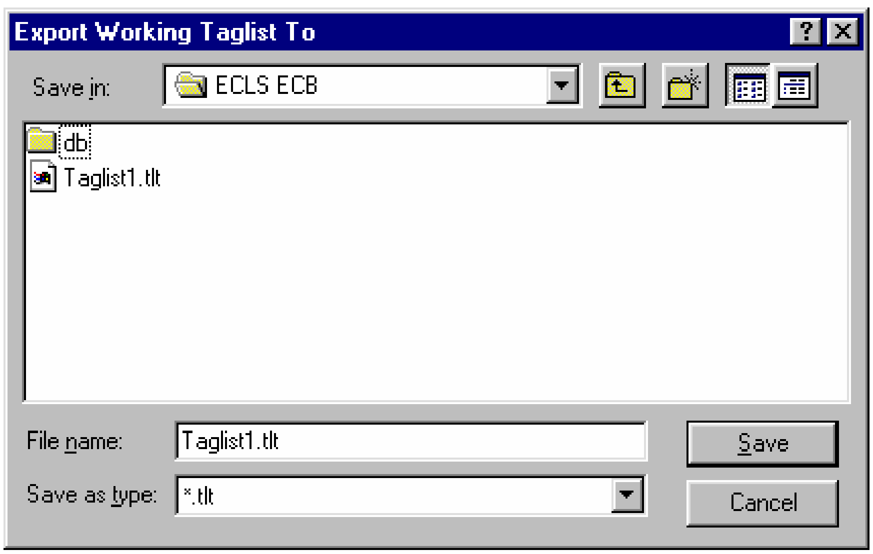


Exhibit 8-32.   Export Taglist dialog box



8.4.6 Importing Taglists

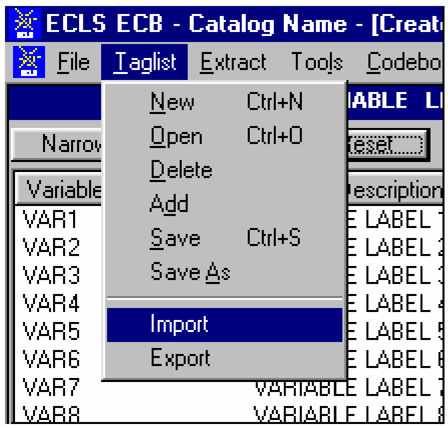
Taglists can be imported to the Working Taglist from external *.tlt files that are created by the ECB Taglist/Export function. Please note that only taglists exported from the same catalog of the same version ECB should be imported.

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 appears (exhibit 8-34).

5. Enter the file name for the taglist you want to import.

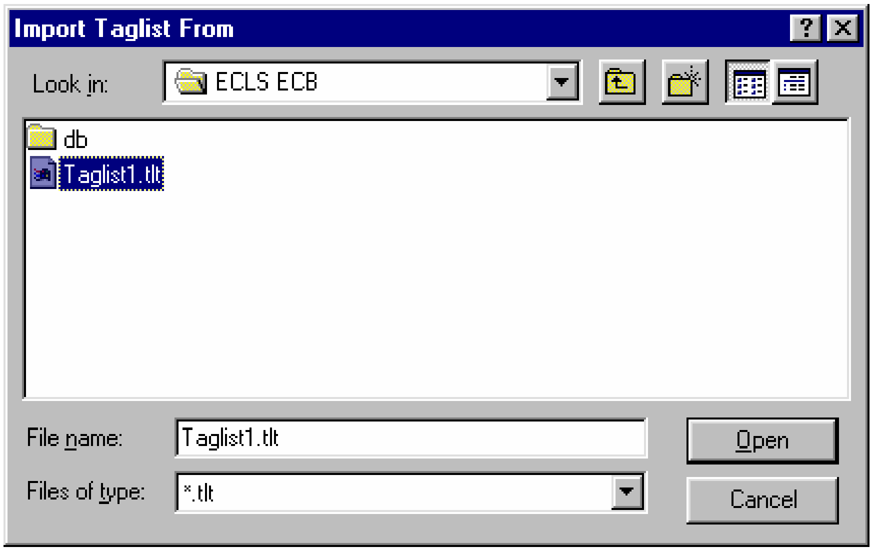
6. Click on the Open button.

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 appendix E on the CD-ROM for listings and descriptions of the catalog-specific predefined taglists.

Exhibit 8-34. Import Taglist dialog box



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 delete 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 Codebook and Variable Information

The codebook for a taglist displayed in the Working Taglist column can be created, viewed, and printed from the ECB main screen. The codebook displays several pieces of information about each variable that are described in exhibit 8-37.

Exhibit 8-37. Codebook 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 Codebook for Tagged Variables:

1. Complete any changes you wish to make to the displayed taglist.
2. Click on the Codebook pulldown menu and select the View option.
3. The codebook 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 codebook.

Exhibit 8-38. Codebook view

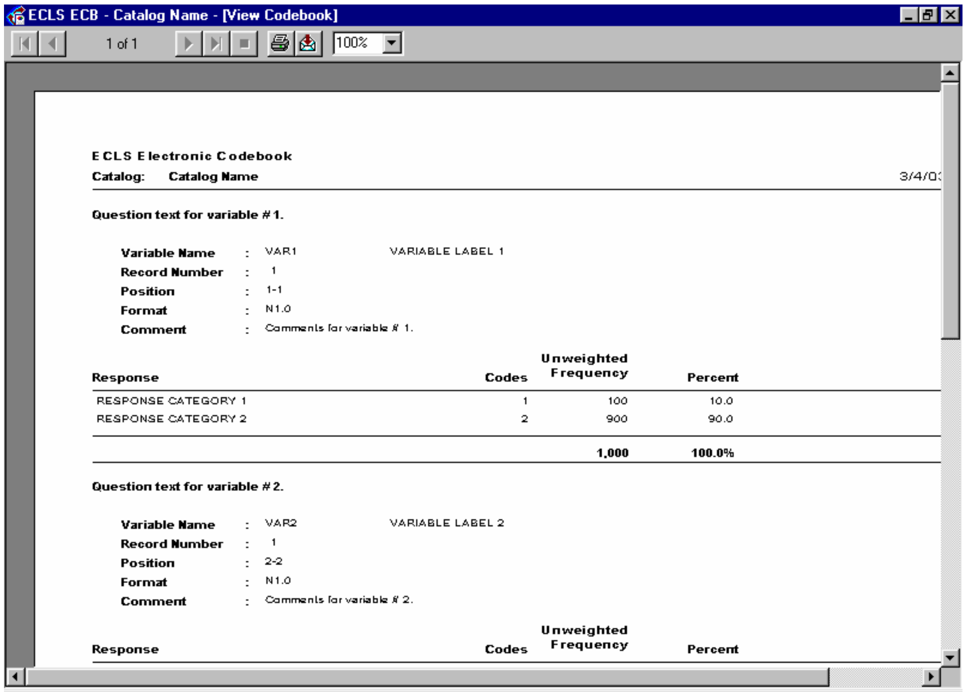
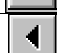
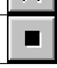
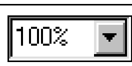


Exhibit 8-39. Navigation buttons

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 codebook. Refer to the procedure below for steps on printing the codebook.
	Click this button to export the codebook to a different destination and save it as a different file format. Refer to the procedure below for steps on exporting the codebook.
	Click the dropdown arrow to select a display magnification of the codebook.

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 codebook, 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 Codebook:

1. Complete any changes you wish to make to the displayed taglist.

2. Click on the Codebook pulldown menu and select the Print option.

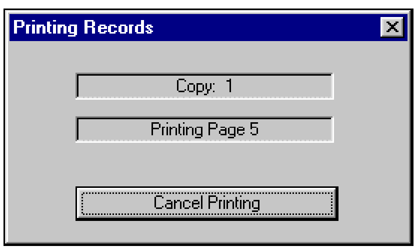
3. The Printing Status screen, shown in exhibit 8-40, appears, and the codebook prints on your PC's default printer.

How To Export the Codebook:

1. Complete any changes you wish to make to the displayed taglist.

2. Click on the Codebook pulldown menu and select the View option.

Exhibit 8-40. Printing status screen

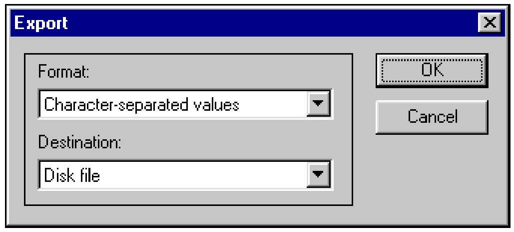


3. The codebook for the current taglist opens in a new window, similar to the one shown in exhibit 8-38.

4. Click on the Export codebook button: 

5. The Export codebook selection screen, shown in exhibit 8-41, appears.

Exhibit 8-41. Export codebook 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.

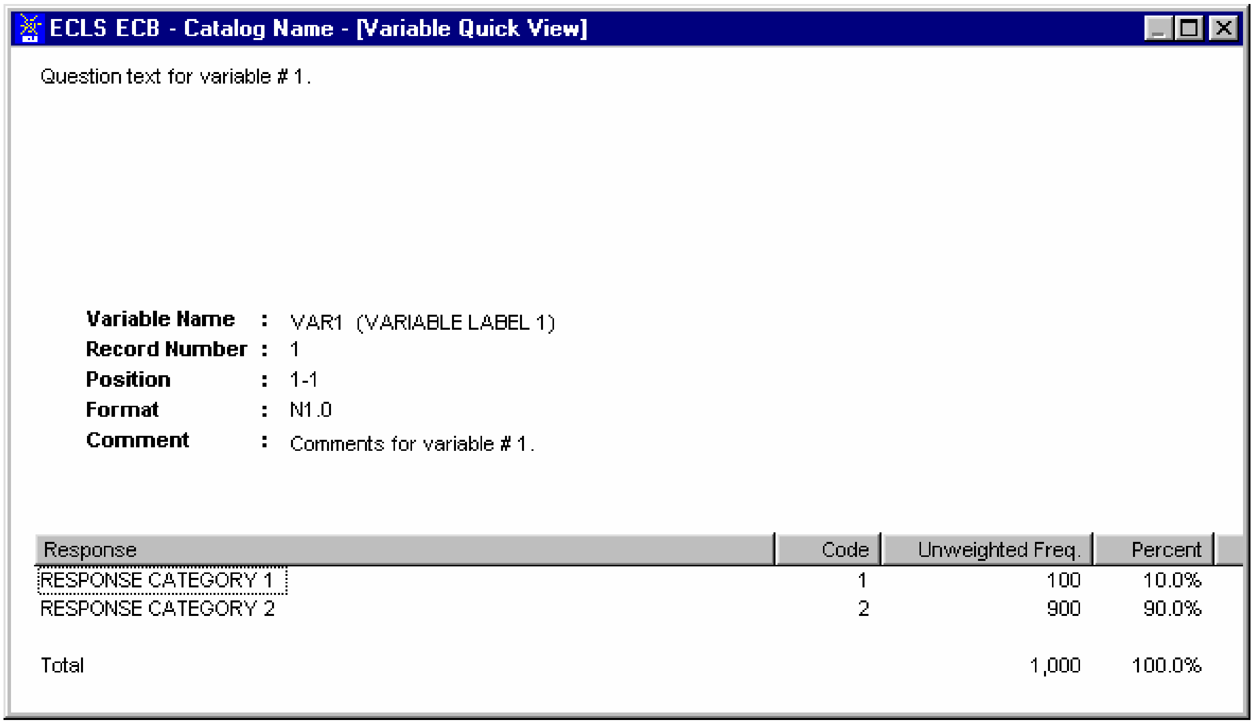
The codebook and its variables can be selected to display their information from either the Variable List or the Working Taglist. The information that can be displayed for a variable includes the variable name and label, the question wording associated with the variable, the position and format of the 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 codebook 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 codebook variables:

How To Display Information for a Single Codebook 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 codebook 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 Codebook 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.

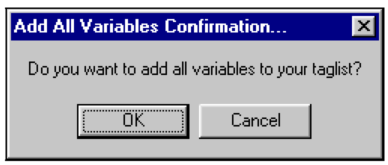
4. Print the image to the printer using the print function of the application that you are using.

How to Display and Print the Entire Codebook 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

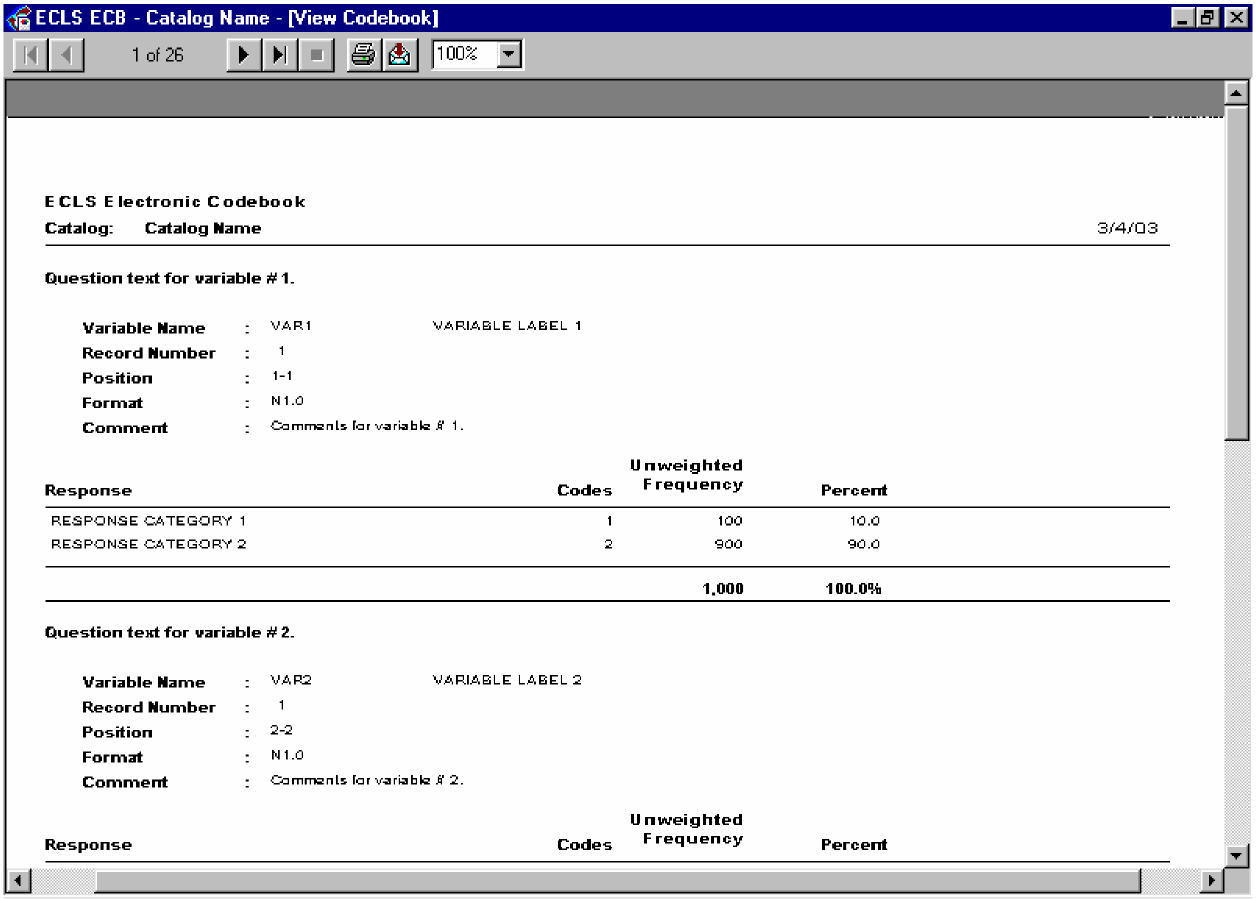


3. All of the variables listed in the Variable List are now displayed in the Working Taglist.

4. Select View from the Codebook pulldown menu.

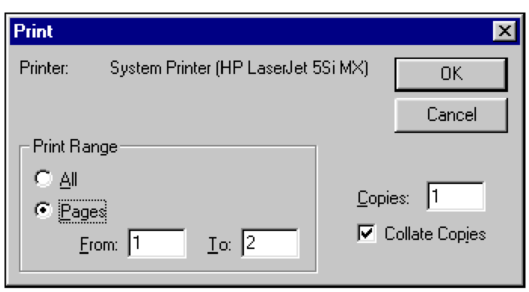
5. The entire codebook displays as shown in exhibit 8-44. Note that this view includes ALL variables in the catalog and can span more than 1000 pages depending on the size of the ECB. The page number is in the upper left corner of the window.

Exhibit 8-44. View of the entire codebook



6. To print the entire codebook, click on the printer icon displayed at the top of the codebook 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 codebook, 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 codebook, 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.

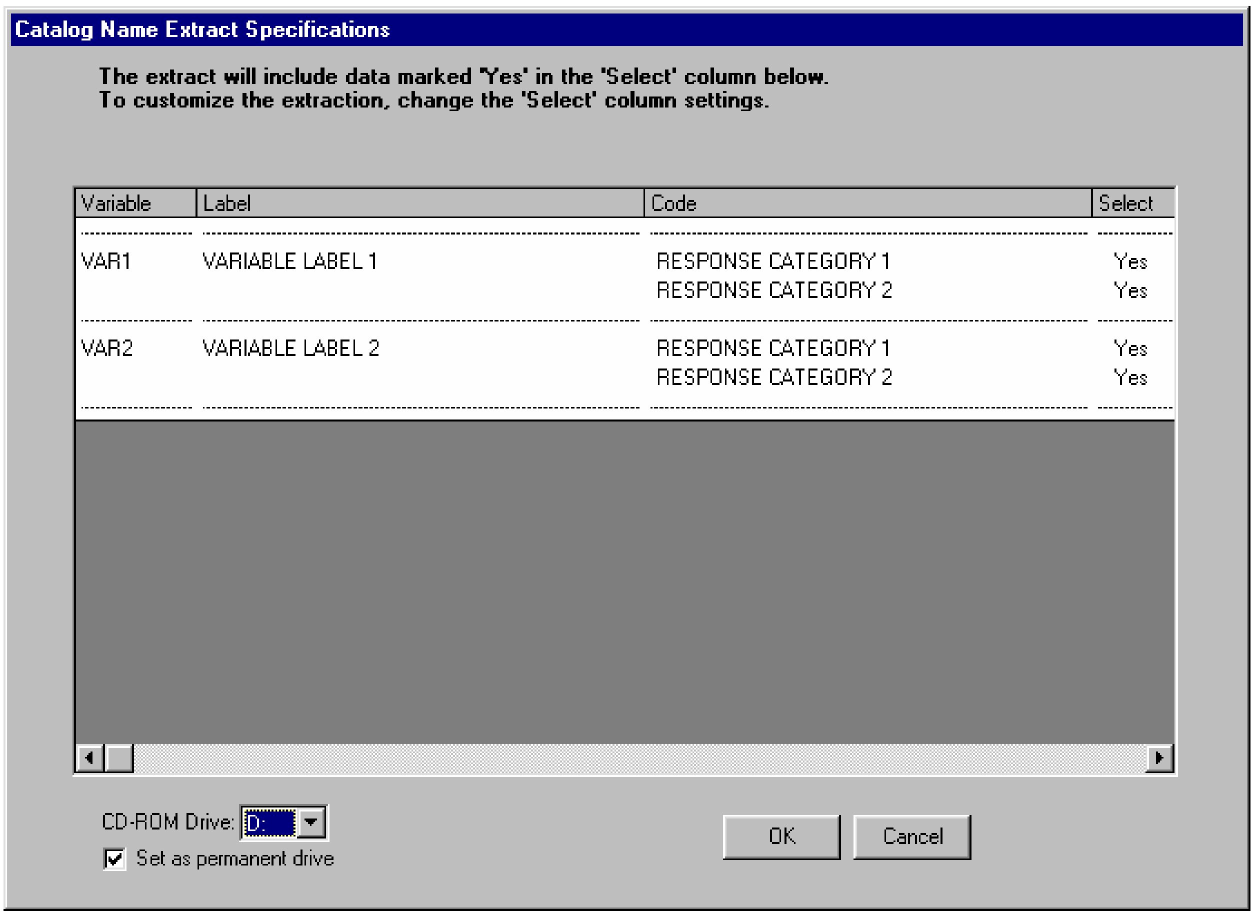
When extracting data to be used with either the SAS, SPSS for Windows, or Stata programs, a dialog box will be presented that allows the user to define the extract population through the Limiting Fields. See exhibit 8-46. The Limiting Fields include various subgroups of respondents that are typically 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 appendix E 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.

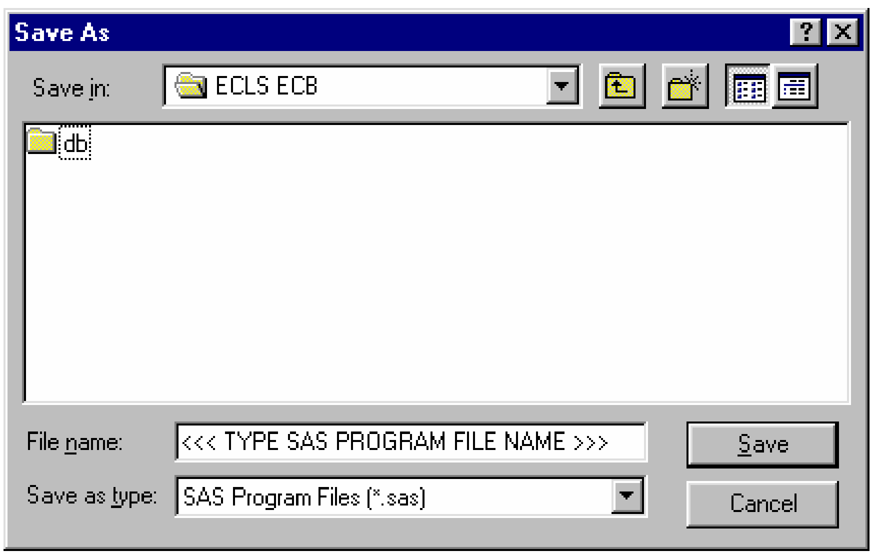
2. Click on the Extract pulldown menu and select the SAS option.

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.

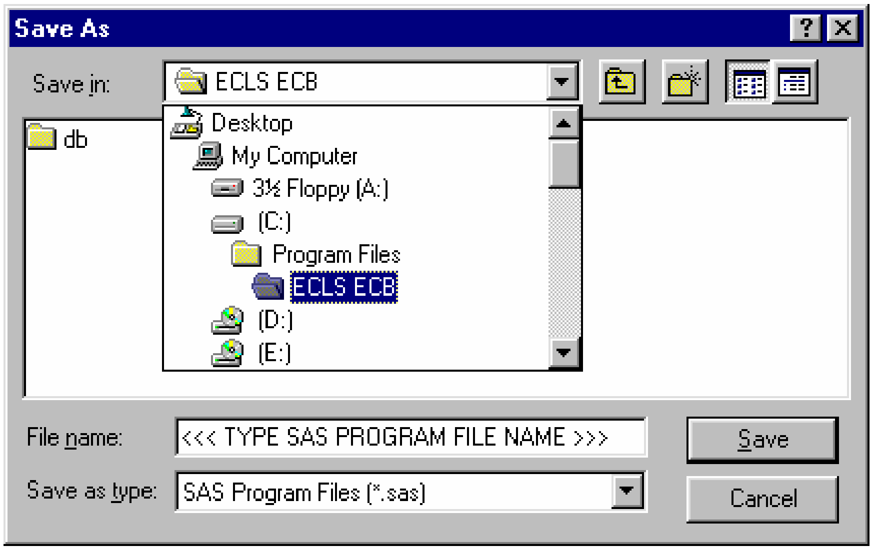
5. Type the desired name of the extract program file in the file name field of the screen shown in exhibit 8-47.

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.

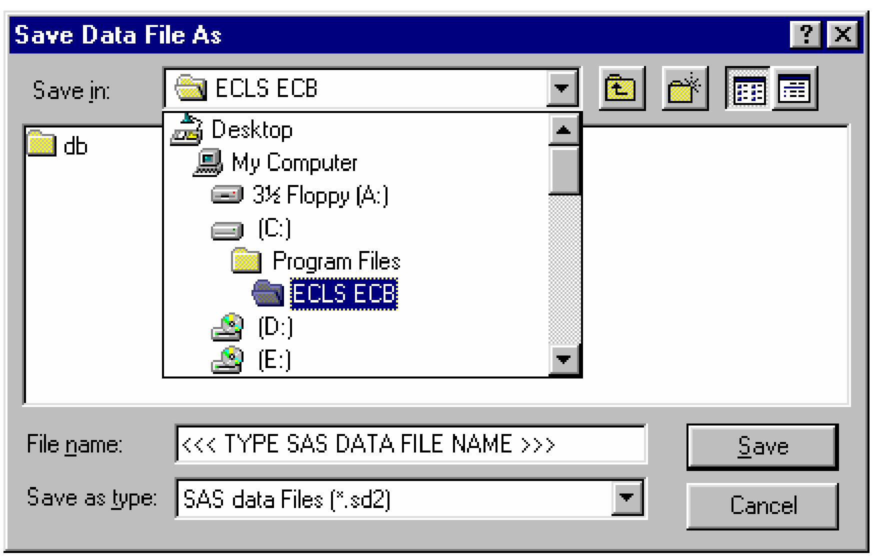
Exhibit 8-48.   Save SAS program file location browse screen



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

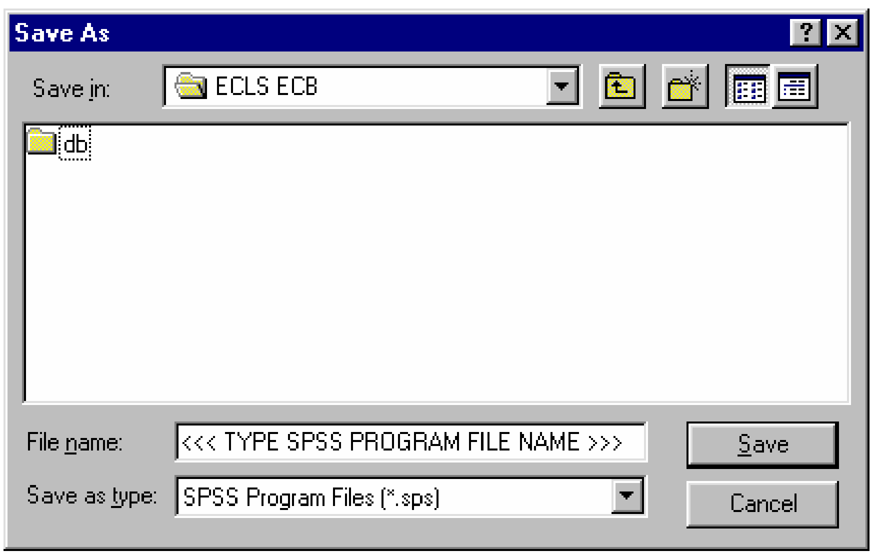
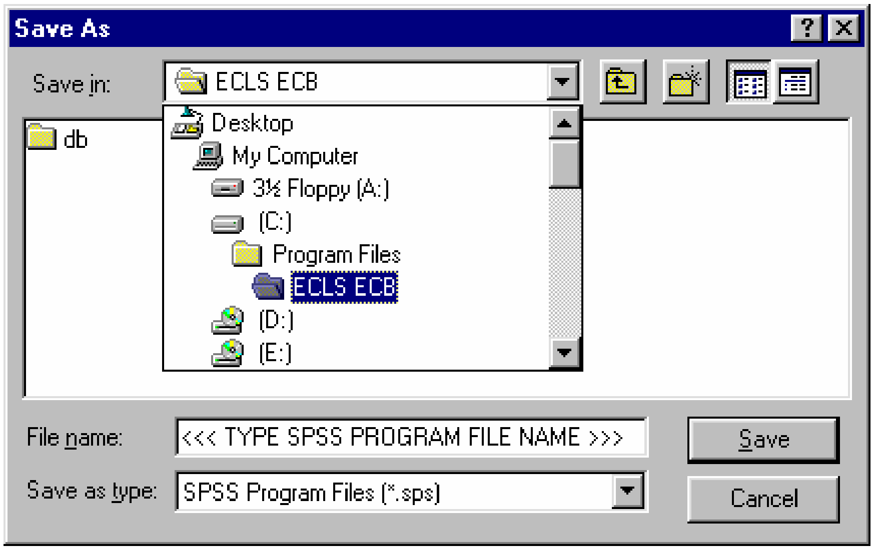


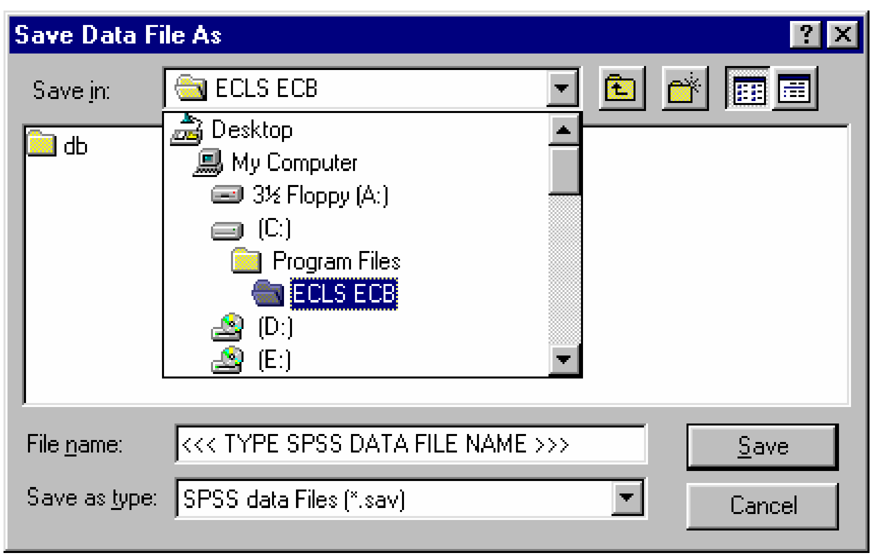
Exhibit 8-51. Save SPSS program file location browse screen



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

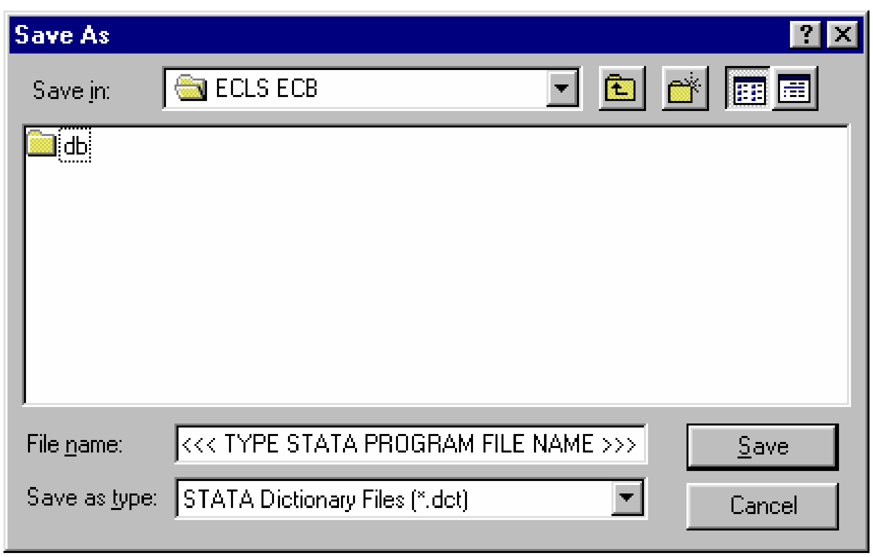


Exhibit 8-54. Save Stata program file location browse screen

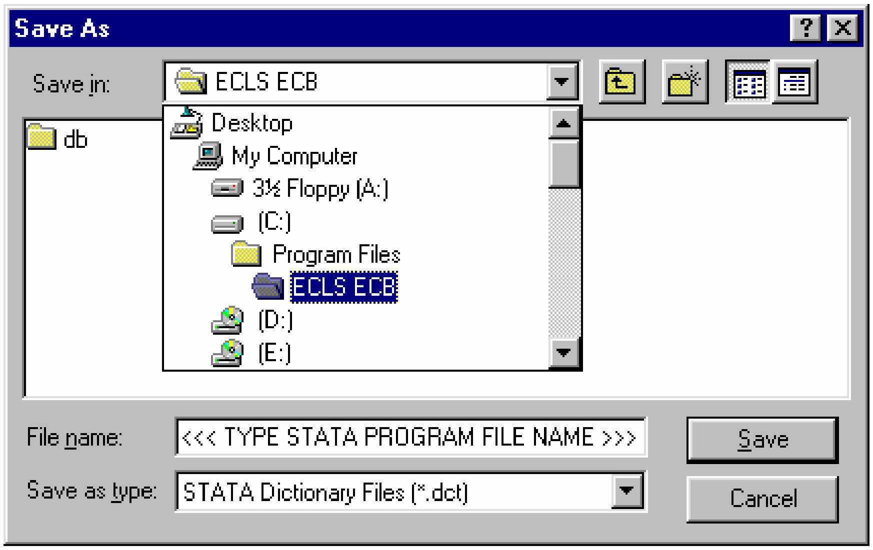
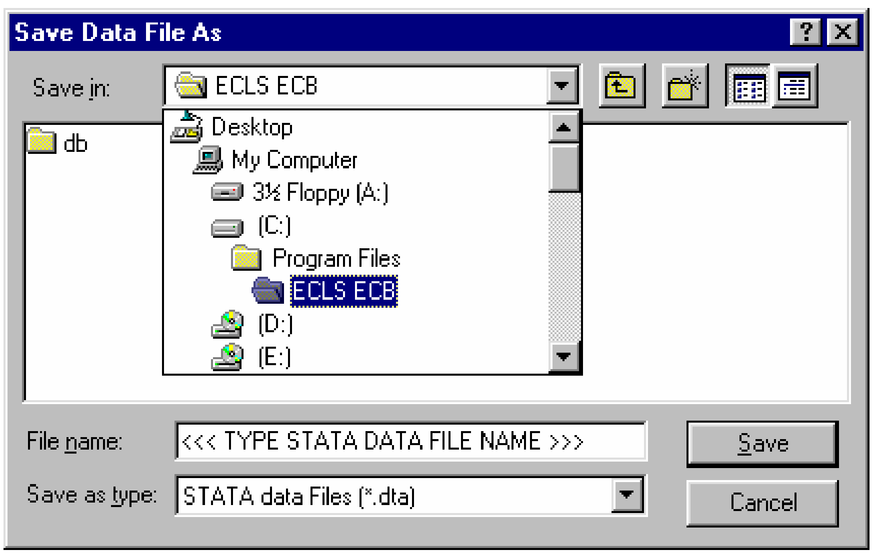


Exhibit 8-55. Save Stata data file dialog box



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 Changes to the On-Line Help Feature

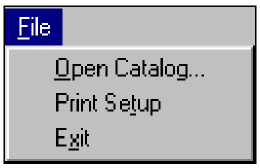
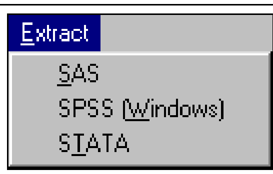
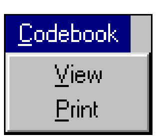
Experienced users of ECLS-K data will notice some differences in the Help feature in this version of the ECB:

- The Help feature is now formatted as an Adobe PDF (portable document format) file. Most users will already be familiar with PDF documents and the Adobe reader. It is highly recommended that ECB users download the latest version of the free Adobe reader from the Adobe web site (www.adobe.com). Version 7.0 of the Adobe reader is included on the ECB CD for users without Internet access.

- The "Search" feature is no longer contained in a tab on the left-side of the screen; it can now be accessed using the search button (identified by the binoculars icon in the toolbar).

8.7 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 Codebook menu contains the command for: ■ Viewing the entire codebook based on the working taglist; and ■ Printing the entire codebook based on the working taglist.
	The Help menu provides access to the detailed online help system.

This page is intentionally left blank.

9. LONGITUDINAL ANALYSES

► *Please note that this chapter is for users who conduct longitudinal analyses. The last section of this chapter is for users of the fifth-grade restricted-use or fifth-grade public-use files who wish to create their own longitudinal files using data from previous rounds of the ECLS-K. Users who intend to use the K–5 longitudinal data file that NCES releases should refer to chapter 10 for additional information.*

Longitudinal analyses with the Early Childhood Longitudinal Study, Kindergarten Class of 1998–99 (ECLS-K) can be conducted both "within school year" and "across school years." Examples of within-year analyses are those that look at children's growth in cognitive scores between fall and spring of kindergarten or between fall and spring of first grade. Such analyses do not require the combined use of kindergarten and first-grade data. They can be conducted using only the kindergarten base-year files or the first-grade files only. Therefore, within-school year analyses are not discussed in this chapter. Since data were only collected once for third grade and once for fifth grade, neither within-third grade nor within-fifth grade analyses are possible. Cross-year analyses, on the other hand, those that combine information from two or more of the kindergarten, first-grade, third-grade, or fifth-grade years, are the focus of this chapter.

This chapter describes how to combine (or merge) the kindergarten, first-grade, third-grade, and fifth-grade files to create cross-year files for K–5 longitudinal analyses. The information contained in this chapter applies to users of the base-year, first-grade, third-grade, and fifth-grade files. Users of the public-use files can consider using the public-use longitudinal file briefly described in chapter 1, which combines data from the base year, first, third, and fifth grades. It contains longitudinal weights so that analysts can examine children's growth and development between kindergarten and fifth grade. Although it is somewhat streamlined, it contains most of the variables in the restricted-use files. Most users will find it more convenient to use the K–5 longitudinal data file that NCES releases rather than creating their own longitudinal file (see chapter 10). However, if users are primarily interested in using data from two or three of the rounds of data collection, they may want to create their own file because they will have more cases included in their analyses. The K–5 longitudinal data file is restricted to a subset of cases that participated in multiple rounds of data collection (see section 10.2 for details on the individuals included on the K–5 longitudinal data file).

This chapter begins with a discussion of K–5 longitudinal analyses and the types of research questions that can be addressed with cross-year files. It then describes the K–5 longitudinal weights available on the cross-sectional files and merging procedures for users who wish to create their own longitudinal files.

9.1 Conducting Longitudinal Analyses

As described in chapter 1, one of the primary goals of the ECLS-K is to understand how children's early experiences influence their transition into kindergarten and their progression through the early elementary school years. A major strength of the ECLS-K design is that it captures important aspects of children's experiences as they occur. Thus, information about children's experiences in each grade is captured in that grade. Capturing this information as it occurs means that the information is not distorted by faulty memory or by revisions to memory based on subsequent experiences. In addition, information from earlier points in time can be included in multivariable models to assess whether they are associated with later events and experiences, thereby strengthening the ability of researchers to make causal inferences.

In conducting K–5 longitudinal analyses with the ECLS-K data, it is important to keep in mind the sample design described in chapter 4. Certain features of the design must be considered. First, because the first-, third-, and fifth-grade data are released only as child-based files, all analyses involving either first-grade, third-grade or fifth-grade data will, of necessity, be child-based. Second, the first-, third-, and fifth-grade data are not representative of all first-grade or third-grade or fifth-grade schools or of classrooms or teachers in the United States. Since the sample was freshened neither in third grade nor in fifth grade, the children are not representative of all children attending third grade in the 2001–02 school year and fifth grade in the 2003–04 school year. Children who started their schooling in the U.S. in second, third, fourth or fifth grade are not represented in the sample. Similarly, since the study follows a cohort, children who were in fifth grade in the 2003–04 school year because they were repeating that grade are not represented in the sample. Researchers conducting K–5 analyses should not attempt to use the data to describe the population of all third- or fifth-grade children, their classrooms, teachers, or schools. However, information about the schools can be used in the child-based analyses to examine, for example, the relationship of the school environment with children's learning or to describe the learning environments of the group of children who attended kindergarten 3 or 5 years earlier. Users may also examine the relationship of the kindergarten year school characteristics with children's later school experiences.

9.2 Examples of Research Questions

A variety of research questions can be examined using the K–5 longitudinal files. The following are some examples:

1. How much do children's reading and mathematics skills increase between the fall of kindergarten and the spring of fifth grade?

2. Do measures of school readiness at the beginning of kindergarten predict children's skill and knowledge levels at the end of fifth grade?

3. What family background characteristics (e.g., family poverty, parent education, maternal employment) are associated with children's later school outcomes?

4. Do children who adapted easily to a school setting in kindergarten do better in fifth grade than their peers who experienced more difficulty settling into school or are there any lingering effects of a slow adjustment to kindergarten?

5. Are there particular school or classroom characteristics that enhance growth rates in reading and mathematics skills between first and fifth grade or between third grade and fifth grade?

6. Are kindergartners' reading and mathematics growth over the first 4 years of school associated with their family's poverty status in kindergarten?

To study these and similar questions, researchers would combine information from two or more rounds of data collection, across the kindergarten, first-, third-, and fifth-grade years. For the first question, the researcher would need to examine differences between fall-kindergarten and spring-fifth grade assessment scores. To do this, one would combine fall-kindergarten data with spring-fifth grade data. Similarly, questions 2 and 3 (regarding the relationship between readiness at kindergarten entry—or maternal employment in that time frame—and fifth-grade outcomes) would be examined by combining data from the same two time points. Note that for question 3 one would need to include data from the parent interview in the base year.

To examine the relationship of children's kindergarten adjustment with their later grade performance, as in question 4, researchers might use data from several rounds (i.e., fall-kindergarten, spring-kindergarten, spring-first grade, spring-third grade, and spring-fifth grade). For example, one could create variables from fall-kindergarten and spring-kindergarten to measure adjustment during kindergarten and then relate those variables to outcomes in the spring of the third and fifth grades.

To be assigned a longitudinal weight for the K–5 data, a case must have participated in the base year (since only base-year respondents were included in subsequent rounds) and in both spring-third grade and spring-fifth grade. Thus, the K–5 longitudinal weights should not be used to examine questions that use data only from the base-year, the first-grade year, and the third-grade year. For such analyses, it is advisable to use the K–1 or K-3 longitudinal weights.

9.3 K–5 Longitudinal Weights

9.3.1 Types of K–5 Longitudinal Weights

K–5 longitudinal weights are used to analyze data in a K–5 file created by merging base-year, first-grade, third-grade, and fifth-grade data or by users of the K–5 created by NCES.[1] Cross-sectional weights, on the other hand, are used for analyses within one round of data collection. There are several sets of K–5 longitudinal weights computed for children with complete data from different combinations of rounds. All K–5 longitudinal weights are child-level weights. There are no K–5 longitudinal weights at the school or teacher level since school- and teacher-level weights are not computed for the first-grade, third-grade, or fifth-grade year. The K–5 longitudinal weights are defined as follows:

- C56CW0 is nonzero if assessment data are present for both spring-third grade and spring-fifth grade, or if the child was excluded from direct assessment in both of these rounds of data collection due to a disability;
- C56PW0 is nonzero if parent interview data are present for both spring-third grade and spring-fifth grade;
- C456CW0 is nonzero if assessment data are present for spring-first and spring-third grade and spring-fifth grade, or if the child was excluded from direct assessment in all of these three rounds of data collection due to a disability;

[1] Please note that the NCES-created file contains more longitudinal weights than are described here. See chapter 10 for details on these additional weights.

- C456PW0 is nonzero if parent interview data are present for spring-first grade and spring-third grade and spring-fifth grade;

- C2_6FC0 is nonzero if assessment data are present for four rounds of data collections involving the full sample of children (spring-kindergarten, spring-first grade, spring-third grade, and spring-fifth grade), or if the child was excluded from direct assessment in all four of these rounds of data collection due to a disability;

- C2_6FP0 is nonzero if parent interview data are present for four rounds of data collections involving the full sample of children (spring-kindergarten, spring-first grade, spring-third grade, and spring-fifth grade);

- C1_6FC0 is nonzero if assessment data are present for five rounds of data collections involving the full sample of children (fall-kindergarten, spring-kindergarten, spring-first grade, spring-third grade, and spring-fifth grade), or if the child was excluded from direct assessment in all five of these rounds of data collection due to a disability;

- C1_6FP0 is nonzero if parent interview data are present for five rounds of data collections involving the full sample of children (fall-kindergarten, spring-kindergarten, spring-first grade, spring-third grade, and spring-fifth grade);

- C1_6SC0 is nonzero if assessment data are present for all six rounds of data collection (fall-kindergarten, spring-kindergarten, fall-first grade, spring-first grade, spring-third grade, and spring-fifth grade), or if the child was excluded from direct assessment in all six rounds of data collection due to a disability; and

- C1_6SP0 is nonzero if parent interview data are present for all six rounds of data collection (fall-kindergarten, spring-kindergarten, fall-first grade, spring-first grade, spring-third grade, and spring-fifth grade).

The use of the K–5 longitudinal weights is described in exhibit 9-1. This exhibit is designed to help users choose appropriate weights for their analysis. First, decide which two or more points in time are the focus of the analysis. The analysis could pertain to two points in time (spring-third grade and spring-fifth grade), three points in time (spring-first grade, spring-third grade, and spring-fifth grade), four points in time (spring-kindergarten, spring-first grade, spring-third grade, and spring-fifth grade), five points in time (fall-kindergarten, spring-kindergarten, spring-first grade, spring-third grade, and spring-fifth grade), or six points in time (all six rounds of data collection). For example, if the analysis uses spring-third grade and spring-fifth grade data, then the appropriate weights would be those beginning with C56 (denoting child-level data from round 5, spring-third grade AND round 6, spring-fifth grade). Second, consider the source of the data, which also affects the choice of the weight. In exhibit 9-1, details under “to be used for analysis of …” provide guidance based on whether the data were collected through the child assessments, parent interviews, teacher questionnaires at the teacher level, or at the child level (reading, math, or science teacher questionnaire). For the same example noted earlier, the two weights available are C56CW0 and C56PW0. If parent data from spring-third grade and spring-fifth grade are needed for the analysis, then C56PW0 should be used.

Exhibit 9-1. ECLS-K: K–5 longitudinal weights, spring-fifth grade: School year 2003–04

Weight	to be used for analysis of ...
C56CW0	child direct assessment data from BOTH spring-third grade and spring-fifth grade, alone or in conjunction with any of the school, teacher, or classroom data, or a limited set of child characteristics (e.g., age, sex, and race/ethnicity).
C56PW0	parent interview data from BOTH spring-third grade and spring-fifth grade, alone or in conjunction with any of the child assessment, school, teacher, or classroom data.
C456CW0	child direct assessment data from spring-first grade AND spring-third grade AND spring-fifth grade, alone or in conjunction with any of the school, teacher, or classroom data, or a limited set of child characteristics (e.g., age, sex, and race/ethnicity).
C456PW0	parent interview data from spring-first grade AND spring-third grade AND spring-fifth grade, alone or in conjunction with any of the child assessment, school, teacher, or classroom data.
C2_6FC0	child direct assessment data from FOUR rounds of data collection involving the FULL sample of children (spring-kindergarten, spring-first grade, spring-third grade, and spring-fifth grade), alone or in conjunction with any of the school, teacher, or classroom data, or a limited set of child characteristics (e.g., age, sex, and race/ethnicity).
C2_6FP0	parent interview data from spring-kindergarten AND spring-first grade AND spring-third grade AND spring-fifth grade, alone or in conjunction with any of the child assessment, school, teacher, or classroom data.
C1_6FC0	child direct assessment data from FIVE rounds of data collections involving the FULL sample of children (fall-kindergarten, spring-kindergarten, spring-first grade, spring-third grade, and spring-fifth grade), alone or in conjunction with any of the school, teacher, or classroom data, or a limited set of child characteristics (e.g., age, sex, and race/ethnicity).
C1_6FP0	parent interview data from FIVE rounds of data collections involving the FULL sample of children (fall-kindergarten, spring-kindergarten, spring-first grade, spring-third grade, spring-fifth grade), alone or in conjunction with any of the child assessment, school, teacher, or classroom data.
C1_6SC0	child direct assessment data from ALL SIX rounds of data collection (fall-kindergarten, spring-kindergarten, fall-first grade, spring-first grade, spring-third grade, and spring-fifth grade), alone or in conjunction with any of the school, teacher, or classroom data, or a limited set of child characteristics (e.g., age, sex, and race/ethnicity).
C1_6SP0	parent interview data from ALL SIX rounds of data collection (fall-kindergarten, spring-kindergarten, fall-first grade, spring-first grade, spring-third grade, and spring-fifth grade), alone or in conjunction with any of the child assessment, school, teacher, or classroom data.

SOURCE: U.S. Department of Education, National Center for Education Statistics, Early Childhood Longitudinal Study, Kindergarten Class of 1998–99 (ECLS-K), spring 2004.

Base year longitudinal weights for the analysis of the base-year data (within the kindergarten year) alone are described in the base-year user's manuals. First-grade longitudinal weights for the analysis of the first-grade data (within the first-grade year) alone, and of the combined kindergarten/first grade data are described in the first-grade user's manuals. Third-grade longitudinal weights for the analysis of the third-grade data alone, and of the combined kindergarten/first grade/third grade data, are described in the third-grade user's manuals.

K–5 longitudinal weights are used to produce estimates of differences between two or more rounds of data collection spanning across kindergarten, first grade, third grade, and fifth grade. Simple examples involving two rounds of data collection are the differences in children's mean assessment scores between spring-third grade and spring-fifth grade using the C56CW0 weight and the difference in the total number of persons in the household size using C56PW0. K–5 longitudinal weights are also used to study the characteristics of children who were assessed in two or more rounds of data collection. For example, one can study how family background characteristics of children in kindergarten are related to assessment scores in spring-fifth grade for children who were assessed in spring-first grade, spring-third grade, and spring-fifth grade. In this case, C456PW0 is used to study the characteristics of the children as reported by their parents, and C456CW0 is used to estimate the difference in assessment scores between spring-first grade and spring-fifth grade. As noted earlier, any longitudinal analysis that uses data from fall-first grade will be limited to a 27 percent subsample of children.

There may be combinations of data for which no weights were developed. For further advice on which weights to use when analyzing a complex combination of data, contact NCES at ECLS@ed.gov.

9.3.2 Weighting Procedures

In this section we discuss the statistical procedures used to produce the K–5 longitudinal weights. These procedures are nearly identical to the procedures used for the cross-sectional weights (see chapter 4). The differences are primarily in how mover status and eligible respondents are defined, and in how adjustment cells are created. For example, in computing weight C56CW0, a child was identified as a mover if the child moved in either spring-third grade or spring-fifth grade; a respondent was defined as a child for whom both cross-sectional weights, C5CW0 and C6CW0, are nonzero. A child with a nonzero C56CW0 had both spring-third grade and spring-fifth grade scorable cognitive assessment data, or was excluded from the cognitive assessments because he or she was a child with disabilities. Longitudinal weights involving the fall-first grade collections were computed differently to adjust for the fact that only a subsample of children was included in fall-first grade.

9.3.2.1 Longitudinal Weights Not Involving the Fall-First Grade Data

The first stage of weighting was to compute an initial child weight that reflected the following:

- Adjustment of the school base weight for base year school-level nonresponse;
- Adjustment of the child weights for base year child-level nonresponse; and
- Adjustment of the base year child weight for subsampling of schools for freshening in first grade (for children sampled in first grade only).

The second stage of weighting was to adjust the initial child weight computed in the first stage for the following:

- Subsampling of movers and
- Child-level nonresponse.

In the adjustment for subsampling of movers, mover status was created so that it was specific to each panel. For example, for the spring-third grade/spring-fifth grade panel (longitudinal weights C56CW0 and C56PW0), a child was a mover if he had been identified as a mover in spring-third grade *or* spring-fifth grade, i.e., in either round he attended a school that was not the school where he had been sampled in kindergarten. The adjustment factor for subsampling movers was computed within cells created using the following characteristics: whether children were sampled in kindergarten or first grade and whether they were language minority children. A small number of children with large weights had their weights trimmed. However, the weights were not redistributed because the total sum of weights was reestablished in the raking procedure that came later. In both steps of the nonresponse adjustment, separate nonresponse classes were created for movers and nonmovers using school affiliation, various combinations of response status of child assessments and parent interviews in the base year, whether children belonged to the language minority group, and the type of household collected from the parent interviews.

The third and last stage was to rake the weights adjusted in the second stage to sample-based control totals. The raking factor was computed separately within raking cells as the sample-based control total for the raking cell over the sum of the nonresponse adjusted weights for children in the same cell. Raking cells (also known as raking dimensions) were created using school and child characteristics collected in the base-year or first-grade year: school affiliation, region, type of locale, sex, age, race/ethnicity, socioeconomic status (SES), language minority status, whether sampled in kindergarten or first grade and, if sampled in kindergarten, mover status.

9.3.2.2 Longitudinal Weights Involving the Fall-First Grade Data

For the longitudinal weights involving the fall-first grade data collection where children were part of a subsample of the ECLS-K full sample, the initial weights were from fall-first grade. These were the base year child adjusted weights (as described in section 4.7.3.2 for base year respondents), incorporating the school subsampling factor appropriate for fall-first grade. These weights were also trimmed to reduce the weight of all the children in one private school that had a large school weight.

The adjustments for subsampling movers and for child nonresponse are identical to those for the other longitudinal weights. The adjustment factor for subsampling movers was computed within cells by whether they belonged in the language minority group. A small number of children with large weights had their weights trimmed. However, the weights were not redistributed because the total sum of weights was reestablished in the raking procedure that came later. In both steps of the nonresponse adjustment, separate nonresponse classes were created for movers and nonmovers using the type of household collected from the parent interviews, whether children belonged to the language minority group, and school affiliation.

The raking dimensions are the same as those for the other longitudinal weights. After the first raking, a small number of children had their weights trimmed; then all the weights were raked again.

9.3.3 Characteristics of Longitudinal Weights

The statistical characteristics of the longitudinal weights are presented in table 9-1. For each weight, the number of cases with nonzero values is presented together with the mean weight, the standard deviation, the coefficient of variation (i.e., the standard deviation as a percentage of the mean weight), the minimum value of the weight, the maximum value of the weight, the skewness, the kurtosis, and the sum of weights.

Table 9-1. Characteristics of child-level K–5 longitudinal weights, spring-third grade: School year 2003–04

Variable name	Number of cases	Mean	Standard deviation	CV[1] (× 100)	Minimum	Maximum	Skewness	Kurtosis	Sum
C56CW0	11,136	353.53	546.33	154.54	1.85	6,088.46	4.23	22.14	3,936,880
C56PW0	10,079	390.45	552.94	141.62	1.87	6,635.16	3.81	19.01	3,935,347
C456CW0	10,852	362.33	588.43	162.40	1.78	6,681.37	4.13	20.98	3,932,020
C456PW0	9,568	410.86	582.33	141.73	2.18	5,941.85	3.68	16.93	3,931,097
C2_6FC0	10,673	359.60	596.79	165.96	1.75	6,360.58	4.25	22.07	3,838,004
C2_6FP0	9,267	414.05	585.96	141.52	2.19	5,945.74	3.59	15.69	3,836,967
C1_6FC0	9,796	391.72	651.89	166.41	1.62	6,867.64	4.21	21.76	3,837,337
C1_6FP0	8,370	458.36	646.59	141.06	2.16	6,801.76	3.62	16.27	3,836,496
C1_6SC0	3,000	1,274.18	1,841.67	144.54	58.68	11,913.28	3.28	11.10	3,822,526
C1_6SP0	2,566	1,490.10	1,835.53	123.18	86.76	10,279.37	2.71	7.31	3,823,589

[1] Coefficient of variation.
SOURCE: U.S. Department of Education, National Center for Education Statistics, Early Childhood Longitudinal Study, Kindergarten Class of 1998–99 (ECLS-K), spring 2004.

The difference in the estimate of the population of students (sum of weights) between the different panels of students and types of weights results from a combination of factors, among them: (1) the number of base-year respondents who became ineligible (due to death, leaving the country, or being a nonsampled mover) after the base year, (2) the adjustment of the weights for the children of unknown eligibility, and (3) the difference in the number of records used to construct sample-based control totals. Of the ten longitudinal weights computed, only the first four (C56CW0, C56PW0, C456CW0, and C456PW0) involve children sampled in first grade. For these four weights, the child records included in the file used for computing the control totals are records of base year respondents and records of eligible children sampled in first grade. The sums of all other longitudinal weights are smaller because records of children sampled in first grade were not included in the file.

9.3.4 Variance Estimation

For each K–5 full sample weight listed in exhibit 9-1, a set of replicate weights was calculated. Replicate weights are used in the jackknife replication method to estimate the standard errors of survey estimates. Any adjustments done to the full sample weights were repeated for the replicate weights.

For longitudinal weights not involving the fall-first grade data, there are 90 replicate weights. For a description of how the replicates were formed, see chapter 4, section 4.7. For the two longitudinal weights involving fall-first grade (C1_6SC0 and C1_6SP0), there are 40 replicate weights. The reason for the smaller number of replicates is that only a subsample of schools was included in the fall-first grade sample. The weights associated with the fall-first grade data do not account for the Durbin method of selecting primary sampling units (PSUs), since it did not apply. Rather, they reflect the fact that only one of the two sampled PSUs in the non-self-representing (NSR) strata was kept in the subsample. To account for this feature, pairs of similar NSR PSUs were collapsed into 19 variance strata. The self-representing (SR) PSUs account for the remaining 21 variance strata.

Each replicate weight variable name has the same weight prefix as for the full sample weight variable name. For example, the replicate weights for C1_6FC0 are C1_6FC1 through C1_6FC90; the replicate weights for C1_6SC0 are C1_6SC1 through C1_6SC40.

Stratum and first-stage unit identifiers used with the Taylor Series method are provided for each of the K–5 longitudinal weights in the file. They are described in exhibit 9-2. For a description of the Taylor Series method, see section 4.8.2.

Specifications for computing standard errors are given in table 9-2. For each type of analysis described in table 9-2, users can choose between the replication method and the Taylor Series method for computing standard errors.

Exhibit 9-2. ECLS-K Taylor Series stratum and first-stage unit identifiers, spring-fifth grade: School year 2003–04

Variable name	Description
C56CSTR	Sampling stratum—spring-third grade/spring-fifth grade longitudinal C-weights
C56CPSU	First-stage primary sampling unit within stratum—spring-third grade/spring-fifth grade longitudinal C-weights
C56PSTR	Sampling stratum—spring-first third/spring-fifth grade longitudinal P-weights
C56PPSU	First-stage primary sampling unit within stratum—spring-third grade/spring-fifth grade longitudinal P-weights
C456CSTR	Sampling stratum—spring-first/spring-third grade/spring-fifth grade longitudinal C-weights
C456CPSU	First-stage primary sampling unit within stratum—spring-first/spring-third grade/spring-fifth grade longitudinal C-weights
C456PSTR	Sampling stratum—spring-first/spring-third grade/spring-fifth grade longitudinal P-weights
C456PPSU	First-stage primary sampling unit within stratum—spring-first/spring-third grade/spring-fifth grade longitudinal P-weights
C26FCSTR	Sampling stratum—spring-kindergarten/spring-first/spring-third grade/spring-fifth grade longitudinal C-weights
C26FCPSU	First-stage primary sampling unit within stratum—spring-kindergarten/spring-first/spring-third grade/spring-fifth grade longitudinal C-weights
C26FPSTR	Sampling stratum—spring-kindergarten/spring-first/spring-third grade/spring-fifth grade longitudinal P-weights
C26FPPSU	First-stage primary sampling unit within stratum—spring-kindergarten/spring-first/spring-third grade/spring-fifth grade longitudinal P-weights
C16FCSTR	Sampling stratum—fall-kindergarten/spring-kindergarten/spring-first grade/spring-third grade/spring-fifth grade longitudinal C-weights
C16FCPSU	First-stage primary sampling unit within stratum—fall-kindergarten/spring-kindergarten/spring-first grade/spring-third grade/spring-fifth grade longitudinal C-weights
C16FPSTR	Sampling stratum—fall-kindergarten/spring-kindergarten/spring-first grade/spring-third grade/spring-fifth grade longitudinal P-weights
C16FPPSU	First-stage primary sampling unit within stratum—fall-kindergarten/spring-kindergarten/spring-first grade/spring-third grade/spring-fifth grade longitudinal P-weights
C16SCSTR	Sampling stratum—longitudinal C-weights covering all six rounds of data collection
C16SCPSU	First-stage primary sampling unit within stratum—longitudinal C-weights covering all six rounds of data collection
C16SPSTR	Sampling stratum—longitudinal P-weights covering all six rounds of data collection
C16SPPSU	First-stage primary sampling unit within stratum—longitudinal P-weights covering all six rounds of data collection

SOURCE: U.S. Department of Education, National Center for Education Statistics, Early Childhood Longitudinal Study, Kindergarten Class of 1998–99 (ECLS-K), spring 2004.

Table 9-2. Specifications for computing standard errors, spring-fifth grade: School year 2003–04

Type of analysis	Full sample weight	Computing standard errors: Replication method (WesVar, SUDAAN or AM): ID	Replicate weights	Jackknife method	Taylor Series method (SUDAAN, Stata, SAS, SPSS or AM): Sample design[1]	Nesting variables	Approximating sampling errors: DEFT (Average root design effect)
Spring-third grade/ spring-fifth grade longitudinal	C56CW0	CHILDID	C56CW1-C56CW90	JK2	WR	C56CSTR-C56CPSU	
	C56PW0	CHILDID	C56PW1-C56PW90	JK2	WR	C56PSTR-C56PPSU	2.037
Spring-first grade/ spring-third grade/ spring-fifth grade longitudinal	C456CW0	CHILDID	C456CW1-C456CW90	JK2	WR	C456CSTR-C456CPSU	
	C456PW0	CHILDID	C456PW1-C456PW90	JK2	WR	C456PSTR-C456PPSU	2.045
Fall-kindergarten/ spring-kindergarten/ spring-first grade/ spring-third grade/ spring-fifth grade longitudinal	C2_6FC0	CHILDID	C2_6FC1-C2_6FC90	JK2	WR	C26FCSTR-C26FCPSU	
	C2_6FP0	CHILDID	C2_6FP1-C2_6FP90	JK2	WR	C26FPSTR-C26FPPSU	2.018
Fall-kindergarten/ spring-kindergarten/ spring-first grade/ spring-third grade/ spring-fifth grade longitudinal	C1_6FC0	CHILDID	C1_6FC1-C1_6FC90	JK2	WR	C16FCSTR-C16FCPSU	
	C1_6FP0	CHILDID	C1_6FP1-C1_6FP90	JK2	WR	C16FPSTR-C16FPPSU	2.028
All six rounds longitudinal	C1_6SC0	CHILDID	C1_6SC1-C1_6SC40	JK2	WR	C16SCSTR-C16SCPSU	
	C1_6SP0	CHILDID	C1_6SP1-C1_6SP40	JK2	WR	C16SPSTR-C16SPPSU	1.787

[1]WR = with replacement, specified only if using SUDAAN. WR is the only option available if using SAS, Stata, SPSS, or AM.
SOURCE: U.S. Department of Education, National Center for Education Statistics, Early Childhood Longitudinal Study, Kindergarten Class of 1998–99 (ECLS-K), spring 2004.

For the replication method using WesVar or AM, the full sample weight, the replicate weights, and the method of replication are required parameters. Variance estimation using the ECLS-K data should be done using the paired jackknife method (JK2). As an example, to compute the mean difference in reading scores between spring-third and spring-fifth grade and their standard errors, users need to specify C56CW0 as the full sample weight, C56CW1 to C56CW90 as the replicate weights, and JK2 as the method of replication.

For the Taylor Series method using SUDAAN, SAS, Stata, SPSS, or AM the full sample weight, the sample design, the nesting stratum, and PSU variables are required. For the same example cited earlier, the full sample weight (C56CW0), the stratum variable (C56CSTR), and the PSU variable (C56CPSU) must be specified. The "with replacement" sample design option, WR, must also be specified if using SUDAAN.

9.3.5 Design Effects

An important analytic device compares the statistical efficiency of survey estimates with what would have been obtained in a hypothetical and usually impractical simple random sample (SRS) of the same size. For a discussion of design effects and their use, see section 4.9. In this section, design effects are presented for selected illustrative estimates produced using longitudinal weights. The tables that follow show estimates, standard errors, and design effects for selected means and proportions based on the ECLS-K child and parent data. For each survey item, the tables present the number of cases, the estimate, the standard error taking into account the actual sample design (Design SE), the standard error assuming SRS (SRS SE), the root design effect (DEFT), and the design effect (DEFF). Standard errors (Design SE) were produced using JK2.

Standard errors and design effects are presented in tables 9-3 to 9-7. Data items are from the direct child assessment, the parent interview, and the child-level teacher questionnaires. Full sample weights were used to compute the estimates; then the corresponding replicate weights were used to compute standard errors and design effects.

Table 9-3. ECLS-K, spring-third grade/spring-fifth grade panel: standard errors and design effects using C56CW0-C56CW90 and C56PW0-C56PW90, by selected child and parent variables: School years 2001–02 and 2003–04

Survey item	Variable name	Number of cases	Estimate	Design SE[1]	SRS SE[2]	DEFT[3]	DEFF[4]
Difference between spring-third grade and spring-fifth grade scores (mean)							
Reading scale score	C6R3RSCL-C5R3RSCL	10,974	20.95	0.228	0.122	1.868	3.490
Mathematics scale score	C6R3MSCL-C5R3MSCL	11,036	20.77	0.254	0.104	2.443	5.966
Science scale score	C6R1SSCL-C5R1SSCL	11,024	12.73	0.172	0.075	2.295	5.269
Self-described : Externalizing problems	C6SDQEXT-C5SDQEXT	11,043	-0.14	0.013	0.007	1.861	3.465
Self-described : Internalizing problems	C6SDQINT-C5SDQINT	11,043	-0.16	0.014	0.007	2.012	4.050
Self-described : Competence in math	C6SDQMTC-C5SDQMTC	11,043	-0.23	0.015	0.008	1.793	3.216
Self-described : Competence in peer relation	C6SDQPRC-C5SDQPRC	11,042	-0.04	0.012	0.006	1.871	3.502
Self-described : Competence in reading	C6SDQRDC-C5SDQRDC	11,043	-0.26	0.015	0.007	2.108	4.443
Self-described : Competence in all subjects	C6SDQSBC-C5SDQSBC	11,043	-0.21	0.013	0.007	1.893	3.582
Approaches to learning-Teacher	T6LEARN-T5LEARN	9,106	0.04	0.013	0.007	1.941	3.769
Self-control-Teacher	T6CONTRO-T5CONTRO	8,944	0.05	0.014	0.007	1.965	3.862
Interpersonal-Teacher	T6INTERP-T5INTERP	8,818	0.00	0.014	0.007	1.881	3.540
Externalizing problems-Teacher	T6EXTERN-T5EXTERN	9,031	-0.07	0.012	0.006	1.972	3.889
Internalizing problems-Teacher	T6INTERN-T5INTERN	8,877	-0.01	0.016	0.007	2.277	5.187
Other differences							
Child’s Body Mass Index (BMI)	C6BMI-C5BMI	10,591	1.97	0.045	0.021	2.136	4.564
Child’s height	C6HEIGHT-C5HEIGHT	10,621	4.57	0.035	0.015	2.303	5.304
Child’s weight	C6WEIGHT-C5WEIGHT	10,635	23.13	0.262	0.117	2.247	5.047
Household size	P6HTOTAL-P5HTOTAL	10,079	-0.01	0.014	0.008	1.702	2.897
Number of hours watched TV after dinner	P6TVAFDH-P5TVAFDH	9,890	0.21	0.017	0.010	1.702	2.896
Child and parent characteristics from parent interview (percent)							
Lived in single parent family	P6HFAMIL	10,079	26.66	0.853	0.440	1.937	3.752
Lived in two-parent family	P6HFAMIL	10,079	70.72	0.942	0.454	2.077	4.316
Mom worked 35 hours+/week	P6HMEMP	7,539	67.00	0.919	0.542	1.697	2.881
Primary care is center-based	P6PRIMNW	3,238	29.52	1.400	0.802	1.746	3.049
Primary care is home-based	P6PRIMNW	3,238	70.48	1.400	0.802	1.746	3.049
Parents had high school or less	W5PARED	10,079	31.83	0.793	0.464	1.710	2.923
Household income category below median	W5INCCAT	10,079	47.47	1.073	0.497	2.157	4.652
Parent attended PTA	P6ATTENP	10,064	40.25	1.280	0.489	2.618	6.852
Visited library	P6LIBRAR	10,057	49.25	1.085	0.499	2.176	4.734
Used computer 1-2 times per week	P6HOMECM						
	P6COMPWK	8,634	34.78	0.983	0.513	1.917	3.676
Had Internet access	P6HOMECM						
	P6INTACC	8,443	88.79	0.576	0.343	1.677	2.814
Used computer 1-2 times/week for homework	P6HOMECM						
	P6CMPEDU	8,433	55.85	0.922	0.541	1.705	2.907
Had family rule for TV	P6TVHOME						
	P6TVRULE	10,014	89.18	0.594	0.311	1.912	3.656
Have someone help with reading homework	P6HELPR	9,941	97.71	0.308	0.150	2.058	4.235
Talked to child about day at school daily	P6OFTTLK	10,048	82.96	0.642	0.375	1.711	2.929
Talked to child about smoking 3+times/year	P6TLKSMK	10,049	72.86	0.774	0.444	1.744	3.043
Talked to child about alcohol 3+times/year	P6TLKALC	10,048	65.34	0.884	0.475	1.862	3.466

See notes at end of table.

Table 9-3. ECLS-K, spring-third grade/spring-fifth grade panel: standard errors and design effects using C56CW0-C56CW90 and C56PW0-C56PW90, by selected child and parent variables: School years 2001–02 and 2003–04—Continued

Survey item	Variable name	Number of cases	Estimate	Design SE[1]	SRS SE[2]	DEFT[3]	DEFF[4]
Child and parent characteristics from parent interview (percent)—Continued							
Took away privilege when child angry	P6HITPRV	9,950	69.14	1.187	0.463	2.564	6.574
Self-reported in very good health	P6HEALTH	9,840	87.48	0.748	0.334	2.242	5.025
HH received food stamps in last 12 months	P6FSTAMP	10,012	15.47	0.806	0.361	2.231	4.977
Child characteristics from teacher questionnaire (percent)							
Child was in fifth grade	T6GLVL	11,136	85.55	0.966	0.333	2.898	8.400
Participated fully in grade-level assessment	G6ASSMT	10,228	86.49	0.942	0.338	2.786	7.763
Parents attended regular conferences	G6REGCON	10,109	84.07	0.758	0.364	2.082	4.336
Child usually worked best ability in reading	G6ABIL	10,587	56.84	1.008	0.481	2.094	4.385
Child was average in language skills	G6RTLANG	10,576	71.49	0.895	0.439	2.039	4.159
Child was in reading class entire school year	G6LNGTM	10,590	82.95	0.800	0.365	2.189	4.793
Child characteristics (mean)							
Age of child in months	R6AGE	11,078	134.81	0.113	0.045	2.508	6.289
Child's BMI	C6BMI	10,876	20.65	0.079	0.045	1.748	3.056
Child's household size	P6HTOTAL	10,079	4.55	0.027	0.014	1.907	3.637
Number of children <18 in child's HH	P6LESS18	10,079	2.51	0.026	0.012	2.144	4.596
Number of siblings in HH	P6NUMSIB	10,079	1.56	0.021	0.011	1.868	3.490
Number of hours watched TV after dinner	P6TVAFDH	10,002	1.10	0.016	0.009	1.850	3.423
Median						1.965	3.862
Mean						2.037	4.231
Standard deviation						0.290	1.271
Coefficient of variation						0.142	0.300
Minimum						1.677	2.814
Maximum						2.898	8.400

[1] Design SE is the standard error under the ECLS-K sample design. For an explanation of this statistic, see section 4.9.
[2] SRS SE is the standard error assuming simple random sample. For an explanation of this statistic, see section 4.9.
[3] DEFT is the root design effect. For an explanation of DEFT, see section 4.9.
[4] DEFF is the design effect. For an explanation of DEFF, see section 4.9.
SOURCE: U.S. Department of Education, National Center for Education Statistics, Early Childhood Longitudinal Study, Kindergarten Class of 1998–99 (ECLS-K), spring 2002 and spring 2004.

Table 9-4. ECLS-K, spring-first grade/spring-third grade/spring-fifth grade panel: standard errors and design effects using C456CW0-C456CW90 and C456PW0-C456PW90, by selected child and parent variables: School years 1999–2000, 2001–02, and 2003–04

Survey item	Variable name	Number of cases	Estimate	Design SE[1]	SRS SE[2]	DEFT[3]	DEFF[4]
Difference between spring-third grade and spring-fifth grade scores (mean)							
Reading scale score	C6R3RSCL-C5R3RSCL	10,702	20.88	0.232	0.124	1.868	3.489
Mathematics scale score	C6R3MSCL-C5R3MSCL	10,764	20.79	0.263	0.105	2.504	6.270
Science scale score	C6R1SSCL-C5R1SSCL	10,754	12.70	0.165	0.076	2.181	4.755
Self-described : Externalizing problems	C6SDQEXT-C5SDQEXT	10,769	-0.13	0.013	0.007	1.960	3.841
Self-described : Internalizing problems	C6SDQINT-C5SDQINT	10,769	-0.15	0.015	0.007	2.153	4.634
Self-described : Competence in math	C6SDQMTC-C5SDQMTC	10,769	-0.23	0.015	0.008	1.780	3.170
Self-described : Competence in peer relation	C6SDQPRC-C5SDQPRC	10,768	-0.04	0.013	0.007	1.961	3.844
Self-described : Competence in reading	C6SDQRDC-C5SDQRDC	10,769	-0.27	0.015	0.007	2.113	4.465
Self-described : Competence in all subjects	C6SDQSBC-C5SDQSBC	10,769	-0.21	0.013	0.007	1.827	3.338
Approaches to learning-Teacher	T6LEARN-T5LEARN	8,917	0.04	0.014	0.007	2.044	4.178
Self-control-Teacher	T6CONTRO-T5CONTRO	8,763	0.04	0.014	0.007	2.013	4.054
Interpersonal-Teacher	T6INTERP-T5INTERP	8,646	-0.01	0.015	0.008	1.942	3.772
Externalizing problems-Teacher	T6EXTERN-T5EXTERN	8,845	-0.07	0.011	0.006	1.771	3.136
Internalizing problems-Teacher	T6INTERN-T5INTERN	8,700	-0.01	0.016	0.007	2.310	5.337
Difference between spring-first grade and spring-third grade scores (mean)							
Reading scale score	C6R3RSCL-C4R3RSCL	10,555	66.55	0.381	0.173	2.207	4.871
Mathematics scale score	C6R3MSCL-C4R3MSCL	10,763	54.50	0.394	0.144	2.730	7.455
Approaches to learning-Teacher	T6LEARN-T4LEARN	9,571	0.00	0.015	0.007	2.061	4.247
Self-control-Teacher	T6CONTRO-T4CONTRO	9,415	0.03	0.014	0.007	1.995	3.981
Interpersonal-Teacher	T6INTERP-T4INTERP	9,309	-0.07	0.015	0.008	1.936	3.749
Externalizing problems-Teacher	T6EXTERN-T4EXTERN	9,465	0.03	0.014	0.007	2.136	4.563
Internalizing problems-Teacher	T6INTERN-T4INTERN	9,327	0.08	0.016	0.007	2.336	5.456
Other differences							
Child's Body Mass Index (BMI)	C6BMI-C5BMI	10,345	1.98	0.043	0.021	2.014	4.057
Child's height	C6HEIGHT-C5HEIGHT	10,374	4.57	0.036	0.015	2.332	5.440
Child's weight	C6WEIGHT-C5WEIGHT	10,387	23.14	0.235	0.118	1.987	3.948
Household size	P6HTOTAL-P5HTOTAL	9,568	-0.03	0.014	0.009	1.624	2.636
Number of hours watched TV after dinner	P6TVAFDH-P5TVAFDH	9,403	0.23	0.017	0.010	1.631	2.659
Child and parent characteristics from parent interview (percent)							
Lived in single parent family	P6HFAMIL	9,568	28.30	0.845	0.461	1.834	3.364
Lived in two-parent family	P6HFAMIL	9,568	69.02	0.885	0.473	1.872	3.505
Mom worked 35 hours+/week	P6HMEMP	7,190	67.47	0.982	0.553	1.777	3.159
Primary care is center-based	P6PRIMNW	3,045	29.81	1.354	0.829	1.633	2.667
Primary care is home-based	P6PRIMNW	3,045	70.19	1.354	0.829	1.633	2.667
Parents had high school or less	W5PARED	9,568	31.06	0.835	0.473	1.765	3.116
Household income category below median	W5INCCAT	9,568	47.79	1.045	0.511	2.046	4.187
Parent attended PTA	P6ATTENP	9,555	40.71	1.245	0.503	2.476	6.132
Visited library	P6LIBRAR	9,546	49.17	1.105	0.511	2.161	4.668
Used computer 1-2 times per week	P6HOMECM P6COMPWK	8,255	35.12	1.021	0.525	1.944	3.780
Had Internet access	P6HOMECM P6INTACC	8,073	88.56	0.614	0.354	1.734	3.007

See notes at end of table.

Table 9-4. ECLS-K, spring-first grade/spring-third grade/spring-fifth grade panel: standard errors and design effects using C456CW0-C456CW90 and C456PW0-C456PW90, by selected child and parent variables: School years 1999–2000, 2001–02, and 2003–04—Continued

Survey item	Variable name	Number of cases	Estimate	Design SE[1]	SRS SE[2]	DEFT[3]	DEFF[4]
Child and parent characteristics from parent interview (percent)—Continued							
Used computer 1-2 times/week for homework	P6HOMECM						
	P6CMPEDU	8,063	55.64	0.967	0.554	1.747	3.052
Had family rule for TV	P6TVHOME						
	P6TVRULE	9,508	89.21	0.603	0.318	1.894	3.589
Have someone help with reading homework	P6HELPR	9,437	97.70	0.390	0.154	2.528	6.389
Talked to child about day at school daily	P6OFTTLK	9,538	82.69	0.626	0.387	1.616	2.613
Talked to child about smoking 3+times/year	P6TLKSMK	9,539	73.80	0.754	0.450	1.675	2.806
Talked to child about alcohol 3+times/year	P6TLKALC	9,538	65.70	0.874	0.486	1.797	3.231
Took away privilege when child angry	P6HITPRV	9,451	69.04	1.306	0.475	2.747	7.544
Self-reported in very good health	P6HEALTH	9,355	87.75	0.744	0.339	2.196	4.822
HH received food stamps in last 12 months	P6FSTAMP	9,505	15.12	0.867	0.368	2.358	5.561
Child characteristics from teacher questionnaire (percent)							
Child was in fifth grade	T6GLVL	10,852	85.65	1.041	0.337	3.092	9.560
Participated fully in grade-level assessment	G6ASSMT	9,972	86.63	0.962	0.341	2.821	7.960
Parents attended regular conferences	G6REGCON	9,862	85.00	0.720	0.359	2.003	4.011
Child usually worked best ability in reading	G6ABIL	10,322	56.59	0.985	0.488	2.020	4.079
Child was average in language skills	G6RTLANG	10,310	72.16	0.853	0.441	1.933	3.737
Child was in reading class entire school year	G6LNGTM	10,325	82.64	0.845	0.373	2.268	5.145
Child characteristics (mean)							
Age of child in months	R6AGE	10,801	134.81	0.109	0.046	2.385	5.690
Child’s BMI	C6BMI	10,609	20.67	0.074	0.046	1.622	2.631
Child’s household size	P6HTOTAL	9,568	4.51	0.025	0.014	1.797	3.231
Number of children <18 in child’s HH	P6LESS18	9,568	2.47	0.025	0.012	2.098	4.403
Number of siblings in HH	P6NUMSIB	9,568	1.55	0.022	0.011	1.935	3.743
Number of hours watched TV after dinner	P6TVAFDH	9,496	1.11	0.016	0.009	1.771	3.136
Median						1.991	3.965
Mean						2.045	4.285
Standard deviation						0.325	1.446
Coefficient of variation						0.159	0.337
Minimum						1.616	2.613
Maximum						3.092	9.560

[1] Design SE is the standard error under the ECLS-K sample design. For an explanation of this statistic, see section 4.9.
[2] SRS SE is the standard error assuming simple random sample. For an explanation of this statistic, see section 4.9.
[3] DEFT is the root design effect. For an explanation of DEFT, see section 4.9.
[4] DEFF is the design effect. For an explanation of DEFF, see section 4.9.
SOURCE: U.S. Department of Education, National Center for Education Statistics, Early Childhood Longitudinal Study, Kindergarten Class of 1998–99 (ECLS-K), fall 1999, spring 2000, spring 2002, and spring 2004.

Table 9-5. ECLS-K, spring-kindergarten/spring-first grade/spring-third grade/spring-fifth grade panel: standard errors and design effects using C2_6FC0-C2_6FC90 and C2_6FP0-C2_6FP90, by selected child and parent variables: School years 1998–99, 1999–2000, 2001–02 and 2003–04

Survey item	Variable name	Number of cases	Estimate	Design SE[1]	SRS SE[2]	DEFT[3]	DEFF[4]
Difference between spring-third grade and spring-fifth grade scores (mean)							
Reading scale score	C6R3RSCL-C5R3RSCL	10,530	20.83	0.239	0.125	1.918	3.678
Mathematics scale score	C6R3MSCL-C5R3MSCL	10,588	20.84	0.246	0.106	2.317	5.370
Science scale score	C6R1SSCL-C5R1SSCL	10,578	12.68	0.166	0.076	2.190	4.795
Self-described : Externalizing problems	C6SDQEXT-C5SDQEXT	10,593	-0.13	0.013	0.007	1.937	3.752
Self-described : Internalizing problems	C6SDQINT-C5SDQINT	10,593	-0.15	0.015	0.007	2.123	4.509
Self-described : Competence in math	C6SDQMTC-C5SDQMTC	10,593	-0.22	0.015	0.008	1.817	3.301
Self-described : Competence in peer relation	C6SDQPRC-C5SDQPRC	10,592	-0.04	0.014	0.007	2.092	4.375
Self-described : Competence in reading	C6SDQRDC-C5SDQRDC	10,593	-0.27	0.016	0.007	2.143	4.594
Self-described : Competence in all subjects	C6SDQSBC-C5SDQSBC	10,593	-0.21	0.013	0.007	1.896	3.596
Approaches to learning-Teacher	T6LEARN-T5LEARN	8,792	0.03	0.015	0.007	2.136	4.564
Self-control-Teacher	T6CONTRO-T5CONTRO	8,641	0.04	0.014	0.007	2.032	4.131
Interpersonal-Teacher	T6INTERP-T5INTERP	8,524	-0.01	0.015	0.008	1.953	3.815
Externalizing problems-Teacher	T6EXTERN-T5EXTERN	8,718	-0.07	0.011	0.006	1.739	3.024
Internalizing problems-Teacher	T6INTERN-T5INTERN	8,577	-0.01	0.015	0.007	2.202	4.847
Difference between spring-first grade and spring-third grade scores (mean)							
Reading scale score	C6R3RSCL-C4R3RSCL	10,394	66.49	0.395	0.174	2.275	5.177
Mathematics scale score	C6R3MSCL-C4R3MSCL	10,589	54.51	0.394	0.144	2.733	7.472
Approaches to learning-Teacher	T6LEARN-T4LEARN	9,440	0.00	0.016	0.008	2.110	4.454
Self-control-Teacher	T6CONTRO-T4CONTRO	9,288	0.03	0.014	0.007	1.979	3.915
Interpersonal-Teacher	T6INTERP-T4INTERP	9,182	-0.07	0.015	0.008	1.987	3.948
Externalizing problems-Teacher	T6EXTERN-T4EXTERN	9,333	0.03	0.014	0.007	2.151	4.628
Internalizing problems-Teacher	T6INTERN-T4INTERN	9,200	0.08	0.016	0.007	2.301	5.296
Other differences							
Child’s Body Mass Index (BMI)	C6BMI-C5BMI	10,179	1.99	0.040	0.021	1.890	3.572
Child’s height	C6HEIGHT-C5HEIGHT	10,208	4.59	0.030	0.015	2.004	4.017
Child’s weight	C6WEIGHT-C5WEIGHT	10,220	23.21	0.223	0.120	1.865	3.480
Household size	P6HTOTAL-P5HTOTAL	9,267	-0.03	0.014	0.009	1.636	2.676
Number of hours watched TV after dinner	P6TVAFDH-P5TVAFDH	9,120	0.22	0.017	0.010	1.681	2.825
Child and parent characteristics from parent interview (percent)							
Lived in single parent family	P6HFAMIL	9,267	27.70	0.850	0.465	1.829	3.347
Lived in two-parent family	P6HFAMIL	9,267	69.66	0.892	0.478	1.868	3.489
Mom worked 35 hours+/week	P6HMEMP	6,993	66.95	1.060	0.562	1.885	3.553
Primary care is center-based	P6PRIMNW	2,936	30.42	1.460	0.849	1.719	2.955
Primary care is home-based	P6PRIMNW	2,936	69.58	1.460	0.849	1.719	2.955
Parents had high school or less	W5PARED	9,267	30.65	0.810	0.479	1.692	2.863
Household income category below median	W5INCCAT	9,267	47.19	1.062	0.519	2.048	4.193
Parent attended PTA	P6ATTENP	9,255	40.69	1.265	0.511	2.476	6.133
Visited library	P6LIBRAR	9,246	49.02	0.979	0.520	1.883	3.544
Used computer 1-2 times per week	P6HOMECM P6COMPWK	8,032	34.89	0.926	0.532	1.741	3.030

See notes at end of table.

Table 9-5. ECLS-K, spring-kindergarten/spring-first grade/spring-third grade/spring-fifth grade panel: standard errors and design effects using C2_6FC0-C2_6FC90 and C2_6FP0-C2_6FP90, by selected child and parent variables: School years 1998–99, 1999–2000, 2001–02 and 2003–04—Continued

Survey item	Variable name	Number of cases	Estimate	Design SE[1]	SRS SE[2]	DEFT[3]	DEFF[4]
Child and parent characteristics from parent interview (percent)—Continued							
Had Internet access	P6HOMECM						
	P6INTACC	7,858	88.54	0.614	0.359	1.710	2.923
Used computer 1-2 times/week for homework	P6HOMECM						
	P6CMPEDU	7,849	55.60	0.940	0.561	1.677	2.812
Had family rule for TV	P6TVHOME						
	P6TVRULE	9,210	89.55	0.601	0.319	1.885	3.553
Have someone help with reading homework	P6HELPR	9,142	97.86	0.363	0.151	2.398	5.750
Talked to child about day at school daily	P6OFTTLK	9,237	82.75	0.636	0.393	1.618	2.618
Talked to child about smoking 3+times/year	P6TLKSMK	9,239	73.84	0.824	0.457	1.802	3.249
Talked to child about alcohol 3+times/year	P6TLKALC	9,238	65.62	0.940	0.494	1.903	3.621
Took away privilege when child angry	P6HITPRV	9,156	69.66	1.333	0.480	2.775	7.699
Self-reported in very good health	P6HEALTH	9,074	88.16	0.650	0.339	1.918	3.677
HH received food stamps in last 12 months	P6FSTAMP	9,209	14.79	0.842	0.370	2.276	5.182
Child characteristics from teacher questionnaire (percent)							
Child was in fifth grade	T6GLVL	10,673	85.80	1.061	0.338	3.142	9.871
Participated fully in grade-level assessment	G6ASSMT	9,804	86.14	0.993	0.349	2.844	8.090
Parents attended regular conferences	G6REGCON	9,699	85.32	0.693	0.359	1.929	3.721
Child usually worked best ability in reading	G6ABIL	10,151	56.73	0.991	0.492	2.015	4.060
Child was average in language skills	G6RTLANG	10,138	72.92	0.831	0.441	1.883	3.546
Child was in reading class entire school year	G6LNGTM	10,154	82.38	0.858	0.378	2.270	5.155
Child characteristics (mean)							
Age of child in months	R6AGE	10,625	134.63	0.084	0.043	1.955	3.823
Child's BMI	C6BMI	10,438	20.66	0.076	0.046	1.651	2.725
Child's household size	P6HTOTAL	9,267	4.50	0.024	0.014	1.730	2.992
Number of children <18 in child's HH	P6LESS18	9,267	2.47	0.025	0.012	2.047	4.189
Number of siblings in HH	P6NUMSIB	9,267	1.54	0.021	0.012	1.804	3.255
Number of hours watched TV after dinner	P6TVAFDH	9,199	1.09	0.016	0.009	1.834	3.362
Median						1.933	3.737
Mean						2.018	4.168
Standard deviation						0.313	1.414
Coefficient of variation						0.155	0.339
Minimum						1.618	2.618
Maximum						3.142	9.871

[1] Design SE is the standard error under the ECLS-K sample design. For an explanation of this statistic, see section 4.9.
[2] SRS SE is the standard error assuming simple random sample. For an explanation of this statistic, see section 4.9.
[3] DEFT is the root design effect. For an explanation of DEFT, see section 4.9.
[4] DEFF is the design effect. For an explanation of DEFF, see section 4.9.
SOURCE: U.S. Department of Education, National Center for Education Statistics, Early Childhood Longitudinal Study, Kindergarten Class of 1998–99 (ECLS-K), fall 1998, spring 1999, fall 1999, spring 2000, spring 2002, and spring 2004.

Table 9-6. ECLS-K, fall-kindergarten/spring-kindergarten/spring-first grade/spring-third grade/spring-fifth grade panel: standard errors and design effects using C1_6FC0-C1_6FC90 and C1_6FP0-C1_6FP90, by selected child and parent variables: School years 1998–99, 1999–2000, 2001–02, and 2003–04

Survey item	Variable name	Number of cases	Estimate	Design SE[1]	SRS SE[2]	DEFT[3]	DEFF[4]
Difference between spring-third grade and spring-fifth grade scores (mean)							
Reading scale score	C6R3RSCL-C5R3RSCL	9,666	20.88	0.261	0.131	1.999	3.998
Mathematics scale score	C6R3MSCL-C5R3MSCL	9,720	20.86	0.274	0.111	2.459	6.046
Science scale score	C6R1SSCL-C5R1SSCL	9,711	12.72	0.172	0.079	2.185	4.776
Self-described : Externalizing problems	C6SDQEXT-C5SDQEXT	9,724	-0.13	0.014	0.007	1.945	3.784
Self-described : Internalizing problems	C6SDQINT-C5SDQINT	9,724	-0.15	0.015	0.007	2.080	4.325
Self-described : Competence in math	C6SDQMTC-C5SDQMTC	9,724	-0.23	0.015	0.009	1.735	3.010
Self-described : Competence in peer relation	C6SDQPRC-C5SDQPRC	9,723	-0.03	0.015	0.007	2.040	4.162
Self-described : Competence in reading	C6SDQRDC-C5SDQRDC	9,724	-0.27	0.017	0.008	2.169	4.705
Self-described : Competence in all subjects	C6SDQSBC-C5SDQSBC	9,724	-0.21	0.014	0.007	1.956	3.826
Approaches to learning-Teacher	T6LEARN-T5LEARN	8,021	0.03	0.015	0.007	2.127	4.525
Self-control-Teacher	T6CONTRO-T5CONTRO	7,880	0.04	0.015	0.007	2.060	4.245
Interpersonal-Teacher	T6INTERP-T5INTERP	7,773	-0.01	0.016	0.008	1.951	3.808
Externalizing problems-Teacher	T6EXTERN-T5EXTERN	7,949	-0.06	0.011	0.006	1.712	2.931
Internalizing problems-Teacher	T6INTERN-T5INTERN	7,816	-0.01	0.016	0.007	2.236	5.000
Difference between spring-first grade and spring-third grade scores (mean)							
Reading scale score	C6R3RSCL-C4R3RSCL	9,546	66.71	0.410	0.180	2.281	5.201
Mathematics scale score	C6R3MSCL-C4R3MSCL	9,717	54.69	0.361	0.150	2.411	5.812
Approaches to learning-Teacher	T6LEARN-T4LEARN	8,663	-0.01	0.016	0.008	2.075	4.307
Self-control-Teacher	T6CONTRO-T4CONTRO	8,530	0.03	0.015	0.007	2.001	4.004
Interpersonal-Teacher	T6INTERP-T4INTERP	8,426	-0.07	0.015	0.008	1.928	3.719
Externalizing problems-Teacher	T6EXTERN-T4EXTERN	8,563	0.03	0.015	0.007	2.150	4.624
Internalizing problems-Teacher	T6INTERN-T4INTERN	8,436	0.07	0.016	0.007	2.268	5.142
Other differences							
Child's Body Mass Index (BMI)	C6BMI-C5BMI	9,343	1.99	0.042	0.022	1.928	3.718
Child's height	C6HEIGHT-C5HEIGHT	9,370	4.61	0.032	0.016	1.999	3.995
Child's weight	C6WEIGHT-C5WEIGHT	9,382	23.23	0.228	0.125	1.829	3.346
Household size	P6HTOTAL-P5HTOTAL	8,370	-0.04	0.015	0.009	1.611	2.596
Number of hours watched TV after dinner	P6TVAFDH-P5TVAFDH	8,252	0.22	0.018	0.011	1.682	2.828
Child and parent characteristics from parent interview (percent)							
Lived in single parent family	P6HFAMIL	8,370	27.80	0.813	0.489	1.661	2.758
Lived in two-parent family	P6HFAMIL	8,370	69.53	0.875	0.503	1.739	3.024
Mom worked 35 hours+/week	P6HMEMP	6,323	66.99	1.109	0.591	1.875	3.517
Primary care is center-based	P6PRIMNW	2,640	31.32	1.519	0.903	1.683	2.833
Primary care is home-based	P6PRIMNW	2,640	68.68	1.519	0.903	1.683	2.833
Parents had high school or less	W5PARED	8,370	30.33	0.870	0.503	1.731	2.997
Household income category below median	W5INCCAT	8,370	47.29	1.089	0.546	1.996	3.985
Parent attended PTA	P6ATTENP	8,358	40.54	1.326	0.537	2.468	6.092
Visited library	P6LIBRAR	8,357	49.13	0.999	0.547	1.827	3.339
Used computer 1-2 times per week	P6HOMECM P6COMPWK	7,287	34.85	0.962	0.558	1.723	2.970

See notes at end of table.

Table 9-6. ECLS-K, fall-kindergarten/spring-kindergarten/spring-first grade/spring-third grade/spring-fifth grade panel: standard errors and design effects using C1_6FC0-C1_6FC90 and C1_6FP0-C1_6FP90, by selected child and parent variables: School years 1998–99, 1999–2000, 2001–02, and 2003–04—Continued

Survey item	Variable name	Number of cases	Estimate	Design SE[1]	SRS SE[2]	DEFT[3]	DEFF[4]
Child and parent characteristics from parent interview (percent)—Continued							
Had Internet access	P6HOMECM						
	P6INTACC	7,137	88.74	0.657	0.374	1.755	3.080
Used computer 1-2 times/week for homework	P6HOMECM						
	P6CMPEDU	7,129	55.67	1.046	0.588	1.778	3.160
Had family rule for TV	P6TVHOME						
	P6TVRULE	8,330	89.72	0.647	0.333	1.944	3.778
Have someone help with reading homework	P6HELPR	8,258	97.90	0.430	0.158	2.725	7.423
Talked to child about day at school daily	P6OFTTLK	8,347	83.07	0.744	0.411	1.812	3.283
Talked to child about smoking 3+times/year	P6TLKSMK	8,349	73.87	0.882	0.481	1.834	3.363
Talked to child about alcohol 3+times/year	P6TLKALC	8,348	65.68	0.985	0.520	1.895	3.592
Took away privilege when child angry	P6HITPRV	8,280	69.44	1.416	0.506	2.796	7.819
Self-reported in very good health	P6HEALTH	8,212	88.55	0.711	0.351	2.023	4.091
HH received food stamps in last 12 months	P6FSTAMP	8,325	14.60	0.897	0.387	2.317	5.368
Child characteristics from teacher questionnaire (percent)							
Child was in fifth grade	T6GLVL	9,796	85.52	1.103	0.356	3.102	9.622
Participated fully in grade-level assessment	G6ASSMT	8,982	86.52	0.989	0.360	2.745	7.537
Parents attended regular conferences	G6REGCON	8,888	85.48	0.740	0.374	1.981	3.923
Child usually worked best ability in reading	G6ABIL	9,301	56.59	1.015	0.514	1.976	3.904
Child was average in language skills	G6RTLANG	9,290	72.97	0.882	0.461	1.914	3.665
Child was in reading class entire school year	G6LNGTM	9,305	82.89	0.926	0.391	2.371	5.624
Child characteristics (mean)							
Age of child in months	R6AGE	9,751	134.63	0.081	0.045	1.815	3.294
Child's BMI	C6BMI	9,585	20.61	0.082	0.048	1.709	2.919
Child's household size	P6HTOTAL	8,370	4.49	0.029	0.015	1.930	3.724
Number of children <18 in child's HH	P6LESS18	8,370	2.47	0.027	0.013	2.141	4.582
Number of siblings in HH	P6NUMSIB	8,370	1.54	0.023	0.012	1.883	3.546
Number of hours watched TV after dinner	P6TVAFDH	8,322	1.09	0.018	0.009	1.946	3.787
Median						1.954	3.817
Mean						2.028	4.205
Standard deviation						0.309	1.386
Coefficient of variation						0.152	0.330
Minimum						1.611	2.596
Maximum						3.102	9.622

[1] Design SE is the standard error under the ECLS-K sample design. For an explanation of this statistic, see section 4.9.
[2] SRS SE is the standard error assuming simple random sample. For an explanation of this statistic, see section 4.9.
[3] DEFT is the root design effect. For an explanation of DEFT, see section 4.9.
[4] DEFF is the design effect. For an explanation of DEFF, see section 4.9.
SOURCE: U.S. Department of Education, National Center for Education Statistics, Early Childhood Longitudinal Study, Kindergarten Class of 1998–99 (ECLS-K), fall 1998, spring 1999, fall 1999, spring 2000, spring 2002, and spring 2004.

Table 9-7. ECLS-K, panel of all six rounds: standard errors and design effects for the full sample using C1_6SC0-C1_6SC40 and C1_6SP0-C1_6SP40, by selected child and parent variables: School years 1998–99, 1999–2000, 2001–02, and 2003–04

Survey item	Variable name	Number of cases	Estimate	Design SE[1]	SRS SE[2]	DEFT[3]	DEFF[4]
Difference between spring-third grade and spring-fifth grade scores (mean)							
Reading scale score	C6R3RSCL-C5R3RSCL	2,968	21.03	0.471	0.239	1.972	3.890
Mathematics scale score	C6R3MSCL-C5R3MSCL	2,986	21.03	0.319	0.196	1.624	2.638
Science scale score	C6R1SSCL-C5R1SSCL	2,982	12.87	0.311	0.145	2.143	4.592
Self-described : Externalizing problems	C6SDQEXT-C5SDQEXT	2,986	-0.10	0.026	0.012	2.092	4.378
Self-described : Internalizing problems	C6SDQINT-C5SDQINT	2,986	-0.12	0.022	0.012	1.761	3.100
Self-described : Competence in math	C6SDQMTC-C5SDQMTC	2,986	-0.24	0.024	0.016	1.504	2.261
Self-described : Competence in peer relation	C6SDQPRC-C5SDQPRC	2,986	-0.08	0.027	0.014	1.982	3.930
Self-described : Competence in reading	C6SDQRDC-C5SDQRDC	2,986	-0.27	0.027	0.014	1.971	3.885
Self-described : Competence in all subjects	C6SDQSBC-C5SDQSBC	2,986	-0.21	0.026	0.013	1.939	3.759
Approaches to learning-Teacher	T6LEARN-T5LEARN	2,307	0.02	0.026	0.014	1.922	3.694
Self-control-Teacher	T6CONTRO-T5CONTRO	2,255	0.04	0.026	0.014	1.876	3.521
Interpersonal-Teacher	T6INTERP-T5INTERP	2,228	-0.01	0.028	0.015	1.909	3.646
Externalizing problems-Teacher	T6EXTERN-T5EXTERN	2,292	-0.06	0.020	0.012	1.677	2.812
Internalizing problems-Teacher	T6INTERN-T5INTERN	2,254	0.01	0.024	0.014	1.724	2.972
Difference between spring-first grade and spring-third grade scores (mean)							
Reading scale score	C6R3RSCL-C4R3RSCL	2,908	67.59	0.634	0.319	1.987	3.950
Mathematics scale score	C6R3MSCL-C4R3MSCL	2,974	54.91	0.673	0.264	2.550	6.503
Approaches to learning-Teacher	T6LEARN-T4LEARN	2,585	-0.03	0.026	0.013	1.944	3.778
Self-control-Teacher	T6CONTRO-T4CONTRO	2,544	0.04	0.024	0.014	1.756	3.082
Interpersonal-Teacher	T6INTERP-T4INTERP	2,528	-0.07	0.030	0.015	2.058	4.234
Externalizing problems-Teacher	T6EXTERN-T4EXTERN	2,559	0.03	0.020	0.012	1.665	2.772
Internalizing problems-Teacher	T6INTERN-T4INTERN	2,512	0.11	0.018	0.014	1.333	1.777
Other differences							
Child's Body Mass Index (BMI)	C6BMI-C5BMI	2,855	1.94	0.071	0.040	1.759	3.095
Child's height	C6HEIGHT-C5HEIGHT	2,863	4.61	0.051	0.029	1.735	3.011
Child's weight	C6WEIGHT-C5WEIGHT	2,865	22.96	0.439	0.230	1.908	3.641
Household size	P6HTOTAL-P5HTOTAL	2,566	-0.04	0.028	0.017	1.671	2.792
Number of hours watched TV after dinner	P6TVAFDH-P5TVAFDH	2,520	0.21	0.037	0.021	1.796	3.227
Child and parent characteristics from parent interview (percent)							
Lived in single parent family	P6HFAMIL	2,566	26.12	1.230	0.867	1.418	2.010
Lived in two-parent family	P6HFAMIL	2,566	70.72	1.285	0.899	1.430	2.045
Mom worked 35 hours+/week	P6HMEMP	1,919	64.43	1.619	1.092	1.482	2.195
Primary care is center-based	P6PRIMNW	788	28.66	2.785	1.611	1.729	2.989
Primary care is home-based	P6PRIMNW	788	71.34	2.785	1.611	1.729	2.989
Parents had high school or less	W5PARED	2,566	29.04	1.117	0.896	1.246	1.553
Household income category below median	W5INCCAT	2,566	48.05	1.560	0.986	1.582	2.502
Parent attended PTA	P6ATTENP	2,558	39.70	2.332	0.967	2.411	5.811
Visited library	P6LIBRAR	2,563	48.13	2.035	0.987	2.062	4.252

See notes at end of table.

Table 9-7. ECLS-K, panel of all six rounds: standard errors and design effects for the full sample using C1_6SC0-C1_6SC40 and C1_6SP0-C1_6SP40, by selected child and parent variables: School years 1998–99, 1999–2000, 2001–02, and 2003–04—Continued

Survey item	Variable name	Number of cases	Estimate	Design SE[1]	SRS SE[2]	DEFT[3]	DEFF[4]
Child and parent characteristics from parent interview (percent)—Continued							
Used computer 1-2 times per week	P6HOMECM						
	P6COMPWK	2,208	36.02	1.603	1.022	1.569	2.461
Had Internet access	P6HOMECM						
	P6INTACC	2,174	88.12	1.022	0.694	1.473	2.170
Used computer 1-2 times/week for homework	P6HOMECM						
	P6CMPEDU	2,175	55.16	1.383	1.066	1.297	1.682
Had family rule for TV	P6TVHOME						
	P6TVRULE	2,549	90.57	0.836	0.579	1.445	2.088
Have someone help with reading homework	P6HELPR	2,534	97.25	0.632	0.325	1.943	3.774
Talked to child about day at school daily	P6OFTTLK	2,560	84.18	1.032	0.721	1.432	2.050
Talked to child about smoking 3+times/year	P6TLKSMK	2,561	75.66	1.248	0.848	1.472	2.167
Talked to child about alcohol 3+times/year	P6TLKALC	2,561	65.92	1.831	0.937	1.955	3.822
Took away privilege when child angry	P6HITPRV	2,532	68.94	2.075	0.919	2.257	5.092
Self-reported in very good health	P6HEALTH	2,511	87.01	1.337	0.671	1.993	3.971
HH received food stamps in last 12 months	P6FSTAMP	2,555	15.97	1.297	0.725	1.790	3.203
Child characteristics from teacher questionnaire (percent)							
Child was in fifth grade	T6GLVL	3,000	86.23	1.402	0.629	2.229	4.969
Participated fully in grade-level assessment	G6ASSMT	2,693	86.99	1.484	0.648	2.290	5.242
Parents attended regular conferences	G6REGCON	2,679	83.71	1.317	0.713	1.846	3.408
Child usually worked best ability in reading	G6ABIL	2,791	59.31	1.590	0.930	1.709	2.922
Child was average in language skills	G6RTLANG	2,785	72.12	1.611	0.850	1.896	3.593
Child was in reading class entire school year	G6LNGTM	2,792	81.96	1.161	0.727	1.596	2.546
Child characteristics (mean)							
Age of child in months	R6AGE	2,990	134.72	0.139	0.084	1.661	2.758
Child's BMI	C6BMI	2,932	20.46	0.142	0.089	1.598	2.553
Child's household size	P6HTOTAL	2,566	4.55	0.045	0.026	1.760	3.098
Number of children <18 in child's HH	P6LESS18	2,566	2.50	0.037	0.022	1.713	2.934
Number of siblings in HH	P6NUMSIB	2,566	1.54	0.035	0.022	1.624	2.637
Number of hours watched TV after dinner	P6TVAFDH	2,549	1.12	0.031	0.018	1.748	3.056
Median						1.758	3.089
Mean						1.787	3.267
Standard deviation						0.274	1.015
Coefficient of variation						0.153	0.311
Minimum						1.246	1.553
Maximum						2.550	6.503

[1] Design SE is the standard error under the ECLS-K sample design. For an explanation of this statistic, see section 4.9.
[2] SRS SE is the standard error assuming simple random sample. For an explanation of this statistic, see section 4.9.
[3] DEFT is the root design effect. For an explanation of DEFT, see section 4.9.
[4] DEFF is the design effect. For an explanation of DEFF, see section 4.9.
SOURCE: U.S. Department of Education, National Center for Education Statistics, Early Childhood Longitudinal Study, Kindergarten Class of 1998–99 (ECLS-K), fall 1998, spring 1999, fall 1999, spring 2000, spring 2002, and spring 2004.

Table 9-8 presents the median design effects for subgroups based on school affiliation, child's sex and race/ethnicity, geographic region, type of locale, and the socioeconomic status scales (SES quintiles) of the parents. At the overall level, median design effects are lowest for the panel that includes all six rounds of data collection. Since this panel of children has a much reduced sample size as it includes the fall-first grade subsample from the full base year sample, the clustering effect is smaller resulting in smaller design effects. Within this smallest panel, median design effects range from 2.4 for children whose race is not one of the main groups to 7.1 for American Indian. This last group has a very small sample size and is highly clustered.

For the other four panels, all involving the full sample of children, median design effects have about the same magnitude at the overall level, between 3.7 and 4.0, compared with 3.1 for the reduced panel. By subgroups, the median design effect is smallest for Pacific Islanders in the panels that include kindergarten and first grade and smallest for American Indians in the panel that includes only third and fifth grade. They are highest for children in Catholic schools in all four panels.

Standard errors and design effects were not computed for items from the teacher and school administrator questionnaires since there are no teacher or school weights computed for spring-third grade year. Although standard errors and design effects may also be calculated for the teacher and school administrator questionnaires at the child level, they are quite large compared to those typically found for the ECLS-K data. Design effects for teacher and school items are large because the intraclass correlation is 100 percent for children in the same school and very high for children in the same class; children attending the same school have the same school data, and children in the same class have the same teacher data. The correlation is not 100 percent for children in the same class because teacher data include not only items about the teacher and the class but also items about the individual students as completed by their teachers.

Table 9-8. ECLS-K panel: median design effects for subgroups, kindergarten through fifth grade: School years 1998–99, 1999–2000, 2001–02, and 2003–04

	Spring-third grade/ spring-fifth grade		Spring-first grade/ spring-third grade/ spring-fifth grade		Spring-kindergarten/ spring-first grade/ spring-third grade/ spring-fifth grade		Fall-kindergarten/ spring-kindergarten/ spring-first grade/ spring-third grade/ spring-fifth grade		All six rounds of data collection	
	DEFT[1]	DEFF[2]	DEFT[1]	DEFT[1]	DEFF[2]	DEFF[2]	DEFT[1]	DEFF[2]	DEFT[1]	DEFF[2]
All students	1.965	3.862	1.991	3.965	1.933	3.737	1.954	3.817	1.758	3.089
School affiliation[3]										
Public	1.888	3.563	1.925	3.705	1.903	3.620	1.921	3.689	1.733	3.004
Private	2.203	4.854	2.241	5.023	2.363	5.585	2.012	4.050	1.882	3.541
Catholic private	2.269	5.147	2.368	5.607	2.510	6.296	2.153	4.635	2.021	4.084
Other private	1.970	3.882	1.943	3.776	2.053	4.215	1.824	3.325	1.709	2.920
Sex										
Male	1.839	3.381	1.843	3.397	1.867	3.485	1.876	3.520	1.642	2.696
Female	1.966	3.864	1.926	3.710	1.882	3.542	1.924	3.701	1.728	2.987
Race/ethnicity										
White	1.963	3.854	1.991	3.963	1.943	3.774	1.952	3.809	1.709	2.920
Black	1.835	3.369	1.769	3.128	1.801	3.243	1.805	3.258	1.591	2.532
Hispanic	1.558	2.426	1.538	2.364	1.540	2.371	1.556	2.422	1.567	2.456
Asian	1.520	2.311	1.619	2.622	1.642	2.697	1.670	2.788	1.763	3.106
Pacific Islander	1.409	1.985	1.290	1.665	1.345	1.810	1.395	1.945	2.046	4.186
American Indian	1.254	1.573	1.431	2.048	1.481	2.193	1.477	2.181	2.657	7.058
Other	1.711	2.929	1.710	2.924	1.760	3.098	1.727	2.982	1.557	2.423
Region										
Northeast	2.049	4.198	2.135	4.560	2.020	4.079	2.029	4.115	1.852	3.430
Midwest	1.998	3.994	2.073	4.297	2.097	4.395	2.102	4.416	1.824	3.325
South	1.901	3.615	1.864	3.474	1.884	3.548	1.919	3.682	1.656	2.742
West	1.790	3.205	1.813	3.286	1.838	3.380	1.811	3.278	1.706	2.911
Type of locale										
Central city	1.931	3.729	1.961	3.846	2.019	4.076	1.965	3.860	1.930	3.726
Urban fringe and large town	1.990	3.959	2.037	4.149	2.080	4.328	2.094	4.383	1.734	3.007
Small town and rural area	1.936	3.748	2.027	4.107	1.951	3.804	2.012	4.047	1.866	3.482
Socioeconomic status quintiles										
First (lowest)	1.721	2.961	1.732	2.997	1.764	3.112	1.731	2.997	1.589	2.523
Second	1.857	3.450	1.862	3.466	1.807	3.262	1.779	3.163	1.604	2.573
Third	1.774	3.148	1.779	3.163	1.751	3.066	1.764	3.113	1.602	2.565
Fourth	1.898	3.602	1.951	3.807	1.929	3.720	1.948	3.795	1.682	2.830
Fifth (highest)	1.911	3.653	1.982	3.927	2.061	4.247	2.022	4.090	1.666	2.777

[1] DEFT is the root design effect. For an explanation of DEFT, see section 4.9.
[2] DEFF is the design effect. For an explanation of DEFF, see section 4.9.
[3] The categories of school affiliation in this table do not match categories of school affiliation in chapter 4. This is to allow users to compare median DEFT and DEFF in fifth grade with those in previous years.
NOTE: Each median is based on 58 items, except for spring-third grade/spring-fifth grade which is based on 51 items.
SOURCE: U.S. Department of Education, National Center for Education Statistics, Early Childhood Longitudinal Study, Kindergarten Class of 1998–99 (ECLS-K), fall 1998, spring 1999, fall 1999, spring 2000, spring 2002, and spring 2004.

► *Please note that section 9.4 applies only to the fifth-grade restricted-use and public-use files. This section does not apply to the K–5 longitudinal file.*

9.4 Merging Base Year Child-Level Data with the First-Grade, Third-Grade and Fifth-Grade Child-Level Data

To create a K–5 data file, which combines data from the base-year, first-grade, third-grade and fifth-grade data collections, an analyst should use the *ECLS-K Base Year Restricted-Use or Public-Use Electronic Code Book* (NCES 2000–097 or NCES 2001–029r); the *ECLS-K First Grade Restricted-Use or Public-Use Electronic Code Book* (NCES 2002–127 or NCES 2002–134); the *ECLS-K Third Grade Restricted-Use or Public-Use Electronic Code Book* (NCES 2003–002 or NCES 2004–002); and the *ECLS-K Fifth-Grade Restricted-Use or Public-Use Electronic Codebook* (NCES 2006–033 or NCES 2006–034). The following procedures for creating a longitudinal file are the same whether users merge public-use or restricted-use single-year files. To create a longitudinal file, perform the following steps to merge the base-year child-level variables needed for analysis with the first-grade, third-grade, and fifth-grade child-level variables needed:

1. Select the variables to be analyzed from the base-year ECB child catalog and the variable CHILDID. This creates a "working taglist" (see section 8.4 in chapter 8 for more detail on how to create a working taglist).
2. Run the program generated after extraction to create a base-year data set (DATA1).
3. Using the child catalog from the First-Grade ECB, select the variables to be analyzed and the variable CHILDID.
4. Run the program generated after extraction to create a first-grade data set (DATA2).
5. Using the child catalog from the Third-Grade ECB, select the variables to be analyzed and the variable CHILDID.
6. Run the program generated after extraction to create a third-grade data set (DATA3).
7. Using the child catalog from the Fifth-Grade ECB, select the variables to be analyzed and the variable CHILDID.
8. Run the program generated after extraction to create a fifth-grade data set (DATA4).

9. Sort DATA1, DATA2, DATA3, and DATA4 by CHILDID.

10. Merge DATA1 and DATA2 and DATA3 and DATA4 by CHILDID.

This merged file will contain 21,409 cases, some of which will not have K–5 longitudinal weights. For example, base-year respondents who did not participate in either fall or spring of first grade or spring of third grade or spring of fifth grade, and movers who were not included in the first-grade, third-grade, and fifth-grade samples, will not have any K–5 longitudinal weights. To select cases with K–5 longitudinal data, a user can use a K–5 longitudinal weight appropriate to the analysis.

10. LONGITUDINAL KINDERGARTEN–FIFTH GRADE PUBLIC-USE DATA FILE

► *Please note that this entire chapter is for users of the K–5 longitudinal data file that NCES releases. Users who have created their own longitudinal files should refer to chapter 9. This chapter does not apply to users of the fifth-grade restricted-use or the fifth-grade public-use file.*

10.1 Introduction

For the Early Childhood Longitudinal Study, Kindergarten Class of 1998–99 (ECLS-K), the longitudinal kindergarten–fifth grade public-use data file, referred to hereafter as the K–5 longitudinal data file, combines data from the base-year (kindergarten), first-grade, third-grade, and fifth-grade years. It was created so that analysts can easily examine children's growth and development between kindergarten and fifth grade without having to go through the process of merging several different data files. Longitudinal weights developed for each round of the ECLS-K are included on the file to be used as K–5 longitudinal weights. All children included on the longitudinal K–1 data file released in 2002 and the K–3 data released in 2004 are also included on this file to allow users to conduct K–1 and K–3, as well as K–5 analyses. Thus, this file can be used to study such topics as children's learning across school years, the extent of summer learning or loss between kindergarten and the fall of the following school year, and the school or classroom characteristics that are related to growth in reading and mathematics skills between first and third and third and fifth grades.

Users will obtain basic information about the K–5 longitudinal data public-use file in this chapter. The chapter begins with a description of the individuals included on the file. It then provides an overview of the content of the data file and a description of the K–5 longitudinal weights.

10.2 Individuals Included on the K–5 Longitudinal File

The K–5 longitudinal file contains all children included on the K–1 longitudinal file released in 2002 and on the K–3 file released in 2004, as well as most of the children for whom either a parent interview or a child assessment was completed in fifth grade. To be included on the original K–1 longitudinal file, a parent interview or child assessment must have been conducted in the spring of kindergarten (spring 1999) and in at least one point in time during the first-grade year (fall 1999 or spring 2000). Thus, children were included if there was either a parent or child assessment conducted during the spring of kindergarten **and** a parent interview or child assessment was completed in any of the following data collections:

- Fall-first grade *or*
- Fall-kindergarten and fall-first grade *or*
- Spring-first grade *or*
- Fall-kindergarten and spring-first grade *or*
- Fall-first grade and spring-first grade *or*
- Fall-kindergarten and fall-first grade and spring-first grade.

Not all children on the original K–1 longitudinal file completed a parent interview or a child assessment in the third grade.[1]

In addition to the children from the K–1 longitudinal file, the K–3 longitudinal file also includes 189 children who had either a parent or child assessment conducted during the spring of third grade **and** a parent interview or child assessment completed in any of the following data collections:

- Spring-first grade *or*
- Spring-kindergarten and spring-first grade *or*
- Fall-kindergarten, spring-kindergarten, and spring-first grade *or*
- Fall-kindergarten and spring-first grade *or*
- Fall-kindergarten and fall-first grade and spring-first grade *or*
- Fall-kindergarten, spring-kindergarten, fall-first grade, and spring-first grade.

[1] Of the 17,212 children on the K–1 longitudinal file, 2,402 (13.6 %) had neither a parent interview nor a child assessment completed in third grade; 792 (4.6 %) had a child assessment completed in third grade, but no parent interview; and 1,693 (9.6 %) had a parent interview completed in third grade, but no child assessment. As noted, all of these children are included on the K–3 longitudinal file so that analysts can conduct longitudinal analyses involving just the kindergarten and first grade data.

There are 306 children with either a completed parent interview or a completed child assessment in third grade who are not included on the K–3 longitudinal file. These children were interviewed only in fall-first grade and spring-third grade or were missing both fall- and spring-first grade data.

The K–5 longitudinal file includes children who were either in the K–1 longitudinal file or in the K–3 longitudinal file, and 164 children who had a parent or child assessment completed during the spring of fifth grade **and** a parent interview or child assessment completed in any of the following data collections:

- Spring-third grade *or*
- Fall-kindergarten and spring-third grade *or*
- Fall-kindergarten and fall-first grade and spring-third grade *or*
- Fall-kindergarten and fall-first grade and spring-first grade and spring-third grade *or*
- Spring-kindergarten and spring-third grade *or*
- Fall-kindergarten and spring-kindergarten and spring-first grade *or*
- Fall-kindergarten and spring-kindergarten and fall-first grade and spring-first grade.

Seven children with either a completed parent interview or a completed child assessment in fifth grade are not included on the K–5 longitudinal file because they did not have a completed parent interview for more than one round nor did they have a completed child assessment for more than one round.

Analyses using combinations of K–5 data other than those listed above can be conducted. Even though customized weights were not created for the other K–5 data combinations, existing weights can be used for analyses of these combinations. See section 10.4 on K–5 longitudinal weights for details.

The K–5 longitudinal data file is a child-level file. All parent, teacher, and school information collected for any particular child from each round of data collection has been attached to that child's record (a more detailed description of the record layout follows). In all, the K–5 longitudinal data file has 17,565 child records. For detailed information about response rates in each round of data collection, see chapter 5 of the base-year, first-grade, third-grade, and fifth-grade user's manuals.

10.3 Content

With a few exceptions, the K–5 longitudinal data file contains all data collected from parents, children, teachers, or schools in the base year (fall and spring), first grade (fall and spring), spring-third grade, and spring-fifth grade data collections. To streamline the file, however, the data from the household rosters that listed all household members, their relationship to the sampled child, and selected other characteristics are not included on the file. The composite variables describing critical household roster-based information, such as the children's family structure and selected characteristics of the family members, have been retained on the file. See chapter 7 of the base-year, first-grade, third-grade, and fifth-grade user's manuals for a description of these and other composite variables.

In addition, cross-sectional weights (associated with a single wave of data collection) and within-grade longitudinal weights (for within-kindergarten or within-first grade longitudinal analysis) are not included on the K–5 longitudinal data file, which contains only the K–5 longitudinal weights (for analysis of kindergarten, first-grade, third-grade, and fifth-grade data). There are no cross-sectional or within-grade longitudinal weights included on the file because the K–5 longitudinal data file should not be used to examine only the kindergarten data, only the first-grade data, only the third-grade data, or only the fifth-grade data. The reason that the file should not be used in this way is that not all children interviewed in kindergarten were interviewed in first grade. Similarly, not all children interviewed in first grade were interviewed in third grade, and not all children interviewed in third grade were interviewed in fifth grade. And, as noted previously, not all children interviewed in fifth grade are included on the K–5 longitudinal file. Thus, the population of base year respondents contained in the K–5 longitudinal data file is a subset of those who were interviewed during the base year. Similarly, the population of first-grade, third-grade, and fifth-grade children on the K–5 longitudinal data file is a subset of those who were interviewed during the first-, third-, and fifth-grade data collections, respectively.

Similar to the first- and third-grade files, the K–5 longitudinal data file contains a few base-year variables that were not in the base-year files. They fall into three categories: (1) base-year recalibrated assessment scores, (2) base-year recalibrated Academic Rating Scale (ARS) scores, and (3) new and corrected base-year composites. The direct child assessment scores were recalibrated to obtain gain scores that could be compared across six waves of data. The ARS scores were recalibrated because an error was identified in the base year ARS scores. Specifically, the fall and spring base year ARS scores used slightly different metrics. These scores were recalibrated using a combined calibration of fall- and spring- kindergarten ratings. Therefore, the unit for the corrected fall- and spring-kindergarten scores is the same, though comparisons between fall- and spring-kindergarten scores are not recommended. Although the item stems are similar across grades, the actual items include performance criteria that increase from one grade to the next. Moreover, the ARS score metric is different at each point. Therefore, change scores should *not* be used to compare fifth-grade ratings with those from earlier rounds.

The specifics of the ARS and composite problems are described in the first-grade public-use user's manual in the section titled Base Year Errata and Composites. The other errors listed in that section have either been corrected (errata number 1 through 7) or are not pertinent to the K–5 longitudinal data file (erratum number 8). For example, the base-year poverty and locality composites were detected to have errors and were recreated and included with the first-grade data file (appendix D) and in the K–5 longitudinal data file. Specifically, WKPOV_R replaces WKPOVRTY and KURBAN_R replaces KURBAN. Similarly, the imputation flag IF_INC_R replaces IF_INC. Errata numbers 3, 6, and 7 were corrected but did not require replacing existing variables.

There are two sets of composite variables that have been revised for the kindergarten, first grade, and third grade years. They are the school lunch composites (percent of children eligible for free lunch and percent of children eligible for reduced-price lunch), and the child's disability status.

In the years preceding the fifth grade year, the school lunch composites were computed at the school level for all schools that completed the school administrator questionnaires, and filled in with values from the Common Core of Data (CCD) for public schools that did not complete the school administrator questionnaire. This still left cases with missing values of the school lunch composites since the CCD also had missing values regarding school lunch. In fifth grade, this was changed so that the school lunch composites no longer have missing values for any public school that had at least one child or parent respondent (i.e., at least one child with nonzero child-level weight or child-level parent weight) in spring-fifth grade. This was accomplished by imputing missing values either by filling in with previous rounds values or by imputing using the hot-deck method. See section 7.5.4.6 for a description of how these composites were created for fifth grade.

The imputation of the school lunch composites was applied to third grade, first grade and kindergarten data. These revised values are now included in the K-5 file. The imputation procedures were as follows:

- Imputation was done separately for each school year, starting with third grade and going backwards.

- For any public school with missing school lunch data, non-imputed data on school lunch from the closest school in time was carried forward. For third grade schools, fifth grade data were used first followed by first grade then kindergarten. For first grade schools, third grade data were used first followed by kindergarten then fifth grade. For kindergarten schools, first grade were used first followed by third grade then fifth grade. The rationale for this approach was that the best source of data for a school was the data from the closest or most recent year.

- Data still missing were imputed by the hot deck method using donors (with non-imputed values) from the same school year. Imputation cells were defined by Title I eligibility and participation and school latitude and longitude (soft-boundary cells).

For each round of data collection preceding fifth grade, the resolution of cases having missing data is shown for each school lunch composite in tables 10-1 to 10-8. In each table, the numbers of cases are for the grade level, not the total number of cases in the K-5 file. For example, out of the 17,565 records in the K-5 file, 13,738 records have nonmissing values of the kindergarten school lunch composites. These are records for children in the K-5 file who are in public schools in kindergarten grade.

Table 10-1. Imputation of school lunch composites at the school level, spring-kindergarten: School year 1998–99

School lunch composite	Number of public schools	Number missing	Percent missing	Values from previous round		Imputed by Hot deck	
				n	Percent	n	Percent
Free lunch	752	284	37.8	252	88.7	32	11.3
Reduced-price lunch	752	298	39.6	265	88.9	33	11.1

SOURCE: U.S. Department of Education, National Center for Education Statistics, Early Childhood Longitudinal Study, Kindergarten Class of 1998–99 (ECLS-K), spring 2004.

Table 10-2. Results of imputation of school lunch composites at the child level, spring-kindergarten: School year 1998–99

School lunch composite	Number of children	Number missing	Percent missing	Values from previous round		Imputed by Hot deck	
				n	Percent	n	Percent
Free lunch	13,738	5,077	37.0	4,599	90.6	478	9.4
Reduced-price lunch	13,738	5,353	39.0	4,864	90.9	489	9.1

SOURCE: U.S. Department of Education, National Center for Education Statistics, Early Childhood Longitudinal Study, Kindergarten Class of 1998–99 (ECLS-K), spring 2004.

Table 10-3. School lunch composite

	Number of public schools	Number missing	Percent missing	Values from previous round		Imputed by Hot deck	
				n	Percent	n	Percent
Free lunch	1,650	914	55.4	530	58.0	384	42.0
Reduced-price lunch	1,650	948	57.5	554	58.4	394	41.6

SOURCE: U.S. Department of Education, National Center for Education Statistics, Early Childhood Longitudinal Study, Kindergarten Class of 1998–99 (ECLS-K), spring 2004.

Table 10-4. Results of imputation of school lunch composites at the child level, spring-first grade: School year 1999–2000

School lunch composite	Number of children	Number missing	Percent missing	Values from previous round		Imputed by Hot deck	
				n	Percent	n	Percent
Free lunch	13,538	5,359	39.6	4,554	85.0	805	15.0
Reduced-price lunch	13,538	5,590	41.3	4,745	84.9	845	15.1

SOURCE: U.S. Department of Education, National Center for Education Statistics, Early Childhood Longitudinal Study, Kindergarten Class of 1998–99 (ECLS-K), spring 2004.

Table 10-5. Imputation of school lunch composites at the school level, spring-third grade: School year 2001–02

School lunch composite	Number of public schools	Number missing	Percent missing	Values from previous round		Imputed by Hot deck	
				n	Percent	n	Percent
Free lunch	2,530	1,026	40.6	314	30.6	712	69.4
Reduced-price lunch	2,530	1,034	40.9	312	30.2	722	69.8

SOURCE: U.S. Department of Education, National Center for Education Statistics, Early Childhood Longitudinal Study, Kindergarten Class of 1998–99 (ECLS-K), spring 2004.

Table 10-6. Results of imputation of school lunch composites at the child level, spring-third grade: School year 2001–02

School lunch composite	Number of children	Number missing	Percent missing	Values from previous round		Imputed by Hot deck	
				n	Percent	n	Percent
Free lunch	11,859	2,534	21.4	1,514	59.7	1,020	40.3
Reduced-price lunch	11,859	2,554	21.5	1,503	58.8	1,051	41.2

SOURCE: U.S. Department of Education, National Center for Education Statistics, Early Childhood Longitudinal Study, Kindergarten Class of 1998–99 (ECLS-K), spring 2004.

Table 10-7. Imputation of school lunch composites at the school level, spring-fifth grade: School year 2003–04

School lunch composite	Number of public schools	Number missing	Percent missing	Values from previous round		Imputed by Hot deck	
				n	Percent	n	Percent
Free lunch	2,008	691	34.4	256	37.0	435	63.0
Reduced-price lunch	2,008	712	35.5	265	37.2	447	62.8

SOURCE: U.S. Department of Education, National Center for Education Statistics, Early Childhood Longitudinal Study, Kindergarten Class of 1998–99 (ECLS-K), spring 2004.

Table 10-8. Results of imputation of school lunch composites at the child level, spring-fifth grade: School year 2003–04

School lunch composite	Number of children	Number missing	Percent missing	Values from previous round		Imputed by Hot deck	
				n	Percent	n	Percent
Free lunch	9,323	2,542	27.3	1,775	69.8	767	30.2
Reduced-price lunch	9,323	2,599	27.9	1,818	69.9	781	30.1

SOURCE: U.S. Department of Education, National Center for Education Statistics, Early Childhood Longitudinal Study, Kindergarten Class of 1998–99 (ECLS-K), spring 2004.

As in fifth grade, there were schools in previous years with the two school lunch composites summing up to more than 100 percent. These values came from two sources: (1) from values reported by the school in another year, or (2) from the hot-deck imputation. The reporting error has been present in all rounds of the ECLS-K, and the decision was to keep the reported values in the data file. If the erroneous values came from the hot-deck imputation, then they were corrected so that the two school lunch composites do not add to more than 100 percent. Correction was made by capping the hot-decked values of the two lunch composite variables. This was done by multiplying each value by 100 divided by the sum of the two variables before capping. This way, both values were reduced by the same amount so they sum to 100 percent.

As described in section 7.5.1.8, the disability composite in spring-fifth grade (P6DISABL) is different from the disability composites in previous years of the study. Prior to fifth grade, any child for whom there was a professional diagnosis of a learning, activity, communication, hearing, or vision problem – as well as all children reported to be using services for the disabled – were considered disabled for the ECLS-K composite variables P1DISABL, P4DISABL, and P5DISABL. For variable P6DISABL on the fifth grade file, however, the following modifications were made to that logic:

- Children whose only learning, activity, behavioral, and emotional diagnoses were coded as “no problem” were no longer considered disabled
- Children whose only reported disability was vision, and whose vision was correctable with glasses OR who were able to read large print books were no longer considered disabled.

The disability composites from the first grade and third grade files have been revised in accordance with the changes listed above to create revised disability composites RP4DISAB and RP5DISAB. An additional condition was added to the composite from the kindergarten file: children whose only report of disability stemmed from receipt of services must have been receiving services other than social work, home visits, parent support, or home tutoring to be coded as disabled in variable RPDISAB1. Table 10-9 presents the effects of the change in disability status coding for the kindergarten, first grade, and third grade data.

Table 10-9. Reclassification of disability status resulting from revisions to the disability composites for spring-kindergarten, spring-first grade, and spring-third grade: School years 1998-99, 1999-2000, 2001-02

Grade and disability status	Original classification	Revised classification
Kindergarten	P1DISABL	RPDISAB1
Not Ascertained	29	29
Disabled	2568	2135
Not Disabled	15500	15933
First Grade	P4DISABL	R4DISAB
Not Ascertained	48	50
Disabled	2560	1619
Not Disabled	13016	13957
First Grade	P5DISABL	R5DISAB
Not Ascertained	72	72
Disabled	3619	1631
Not Disabled	9798	11786

SOURCE: U.S. Department of Education, National Center for Education Statistics, Early Childhood Longitudinal Study, Kindergarten Class of 1998–99 (ECLS-K), spring 2004.

10.4 K–5 Longitudinal Weights

There are several sets of K–5 longitudinal weights computed for children with complete data from different combinations of rounds. All K–5 longitudinal weights are child-level weights. There are no K–5 longitudinal weights at the school or teacher level since school- and teacher-level weights were not computed for the first-grade, third-grade, or fifth-grade years due to lack of representativeness.

The K–5 longitudinal weights, available on the K–5 longitudinal data file Electronic Codebook (ECB), are described in exhibit 10-1. The use of the K–5 longitudinal weights is described in exhibit 10-2, which is designed to help users choose appropriate weights for their analyses.

First, decide which two or more points in time are the focus of the analysis. The analysis could pertain to two points in time (e.g., spring-kindergarten and fall-first grade, or spring-kindergarten and spring-first grade, or spring-first grade and spring-third grade); three points in time (e.g., spring-first grade, spring-third grade, and spring-fifth grade); four points in time (any four of fall-kindergarten, spring-kindergarten, fall-first grade, spring-first grade, spring-third grade, and spring-fifth grade); five points in time (any five of fall-kindergarten, spring-kindergarten, fall-first grade, spring-first grade, spring-third grade, and spring-fifth grade); or six points in time (all six rounds of data). For example, if the analysis uses spring-kindergarten and fall-first grade data, then the appropriate weight would be one that begins with C23 (denoting child-level data from round 2 AND round 3). If the analysis uses data from spring-kindergarten, spring-first, and spring-third grade, then the appropriate weight begins with C245 (denoting data from rounds 2, 4, AND 5). If the analysis uses data from spring-kindergarten, spring-first, spring-third, and spring-fifth grade, then the appropriate weight begins with C2_6F.

Second, consider the source of the data, which also affects the choice of the weight. In exhibit 10-2, details under "to be used in the analysis of …" column provide guidance based on whether the data were collected through the child assessments, parent interviews, or teacher questionnaires. If parent data from spring-kindergarten and fall-first grade are needed for the analysis, then C23PW0 should be used, otherwise C23CW0 can be used. Similarly, if an analyst wishes to examine the influence of parent characteristics on gains in assessment scores between kindergarten and third grade, the appropriate weight would be C245PW0, indicating that parent interview data was included. However, if only child or teacher data were used in the analysis, then the appropriate weight to use is C245CW0.

Exhibit 10-1. ECLS-K: K–5 longitudinal weights: School years 1998–99, 1999–2000, 2001–02, and 2003-04

K–5 longitudinal (panel) weight	is nonzero if …
C23CW0	assessment data are present for both spring-kindergarten and fall-first grade, or if the child was excluded from direct assessment in both of these rounds of data collection due to a disability.
C23PW0	parent interview data are present for both spring-kindergarten and fall-first grade.
C123CW0	assessment data are present for fall- and spring-kindergarten and fall-first grade, or if the child was excluded from direct assessment in all three of these rounds of data collection due to a disability.
C123PW0	parent interview data are present for fall- and spring-kindergarten and fall-first grade.
C24CW0	assessment data are present for both spring-kindergarten and spring-first grade, or if the child was excluded from direct assessment in both of these rounds of data collection due to a disability.
C24PW0	parent interview data are present for both spring-kindergarten and spring-first grade.
C124CW0	assessment data are present for fall-kindergarten and spring-kindergarten and spring-first grade, or if the child was excluded from direct assessment in all three of these rounds of data collection due to a disability.
C124PW0	parent interview data are present for fall-kindergarten and spring-kindergarten and spring-first grade.
C1_4CW0	assessment data are present for four rounds of data collection (fall-kindergarten, spring-kindergarten, fall-first grade, and spring-first grade), or if the child was excluded from direct assessment in all of these four rounds of data collection due to a disability.
C1_4PW0	parent interview data are present for four rounds of data collection (fall-kindergarten, spring-kindergarten, fall-first grade, and spring-first grade).
Y2COMW0	assessment data are present for fall-kindergarten and spring-kindergarten and spring-first grade, or if the child was excluded from direct assessment in all three of these rounds of data collection and parent and/or teacher data are present for one or more base year rounds and parent and/or teacher data are present for spring-first grade.
C45CW0	assessment data are present for both spring-first grade and spring-third grade, or if the child was excluded from direct assessment in both of these rounds of data collection due to a disability.
C45PW0	parent interview data are present for both spring-first grade and spring-third grade.
C245CW0	assessment data are present for spring-kindergarten and spring-first grade and spring-third grade, or if the child was excluded from direct assessment in all of these three rounds of data collection due to a disability.
C245PW0	parent interview data are present for spring-kindergarten and spring-first grade and spring-third grade.
C1_5FC0	assessment data are present for four rounds of data collection involving the full sample of children (fall-kindergarten, spring-kindergarten, spring-first grade, and spring-third grade), or if the child was excluded from direct assessment in all four of these rounds of data collection due to a disability.
C1_5FP0	parent interview data are present for four rounds of data collection involving the full sample of children (fall-kindergarten, spring-kindergarten, spring-first grade, and spring-third grade).
C1_5SC0	assessment data are present for all five rounds of data collection (fall-kindergarten, spring-kindergarten, fall-first grade, spring-first grade, and spring-third grade), or if the child was excluded from direct assessment in all five rounds of data collection due to a disability.
C1_5SP0	parent interview data are present for all five rounds of data collection (fall-kindergarten, spring-kindergarten, fall-first grade, spring-first grade, and spring-third grade).

See notes at end of table.

Exhibit 10-1. ECLS-K: K–5 longitudinal weights: School years 1998–99, 1999–2000, 2001–02, and 2003-04—Continued

K–5 longitudinal (panel) weight	is nonzero if …
C56CW0	assessment data are present for both spring-third grade and spring-fifth grade, or if the child was excluded from direct assessment in both of these rounds of data collection due to a disability.
C56PW0	parent interview data are present for both spring-third grade and spring-fifth grade.
C456CW0	assessment data are present for spring-first grade and spring-third grade and spring-fifth grade, or if the child was excluded from direct assessment in all of these three rounds of data collection due to a disability.
C456PW0	parent interview data are present for spring-first grade and spring-third grade and spring-fifth grade.
C2_6FC0	assessment data are present for four rounds of data collection involving the full sample of children (spring-kindergarten, spring-first grade, spring-third grade, and spring-fifth grade), or if the child was excluded from direct assessment in all four of these rounds of data collection due to a disability.
C2_6FP0	parent interview data are present for four rounds of data collection involving the full sample of children (spring-kindergarten, spring-first grade, spring-third grade, and spring-fifth grade).
C1_6FC0	assessment data are present for five rounds of data collection involving the full sample of children (fall-kindergarten, spring-kindergarten, spring-first grade, spring-third grade, and spring-fifth grade), or if the child was excluded from direct assessment in all five of these rounds of data collection due to a disability.
C1_6FP0	parent interview data are present for five rounds of data collection involving the full sample of children (fall-kindergarten, spring-kindergarten, spring-first grade, spring-third grade, and spring-fifth grade).
C1_6SC0	assessment data are present for all six rounds of data collection (fall-kindergarten, spring-kindergarten, fall-first grade, spring-first grade, spring-third grade, and spring-fifth grade), or if the child was excluded from direct assessment in all six rounds of data collection due to a disability.
C1_6SP0	parent interview data are present for all six rounds of data collection (fall-kindergarten, spring-kindergarten, fall-first grade, spring-first grade, spring-third grade, and spring-fifth grade).

SOURCE: U.S. Department of Education, National Center for Education Statistics, Early Childhood Longitudinal Study, Kindergarten Class of 1998–99 (ECLS-K), fall 1998, spring 1999, fall 1999, spring 2000, spring 2002, and spring 2004.

Exhibit 10-2. Use of the K–5 longitudinal weights: School years 1998–99, 1999–2000, 2001–02, and 2003-04

K–5 longitudinal (panel) weights	to be used for analysis of …
C23CW0	child direct assessment data from BOTH spring-kindergarten and fall-first grade, alone or in conjunction with any of the school, teacher, or classroom data, or a limited set of child characteristics (e.g., age, sex, and race/ethnicity).
C23PW0	parent interview data from BOTH spring-kindergarten and fall-first grade, alone or in conjunction with any of the child assessment, school, teacher, or classroom data.
C123CW0	child direct assessment data from fall- AND spring-kindergarten AND fall-first grade, alone or in conjunction with any of the school, teacher, or classroom data, or a limited set of child characteristics (e.g., age, sex, and race/ethnicity).
C123PW0	parent interview data from fall- AND spring-kindergarten AND fall-first grade, alone or in conjunction with any of the child assessment, school, teacher, or classroom data.
C24CW0	child direct assessment data from BOTH spring-kindergarten and spring-first grade, alone or in conjunction with any of the school, teacher, or classroom data, or a limited set of child characteristics (e.g., age, sex, and race/ethnicity).
C24PW0	parent interview data from BOTH spring-kindergarten and spring-first grade, alone or in conjunction with any of the child assessment, school, teacher, or classroom data.
C124CW0	child direct assessment data from fall-kindergarten AND spring-kindergarten AND spring-first grade, alone or in conjunction with any of the school, teacher, or classroom data, or a limited set of child characteristics (e.g., age, sex, and race/ethnicity).
C124PW0	parent interview data from fall-kindergarten AND spring-kindergarten AND spring-first grade, alone or in conjunction with any of the child assessment, school, teacher, or classroom data.
C1_4CW0	child direct assessment data from FOUR rounds of data collection (fall-kindergarten, spring-kindergarten, fall-first grade, and spring-first grade), alone or in conjunction with any of the school, teacher, or classroom data, or a limited set of child characteristics (e.g., age, sex, and race/ethnicity).
C1_4PW0	parent interview data from FOUR rounds of data collection (fall-kindergarten, spring-kindergarten, fall-first grade, and spring-first grade), alone or in conjunction with any of the child assessment, school, teacher, or classroom data.
Y2COMW0	child direct assessment data from fall-kindergarten AND spring-kindergarten AND spring-first grade, in conjunction with parent and/or teacher data from spring-first grade, AND one or more base year rounds of parent and/or teacher data.
C45CW0	child direct assessment data from BOTH spring-first grade and spring-third grade, alone or in conjunction with any of the school, teacher, or classroom data, or a limited set of child characteristics (e.g., age, sex, and race/ethnicity).
C45PW0	parent interview data from BOTH spring-first grade and spring-third grade, alone or in conjunction with any of the child assessment, school, teacher, or classroom data.
C245CW0	child direct assessment data from spring-kindergarten AND spring-first grade AND spring-third grade, alone or in conjunction with any of the school, teacher, or classroom data, or a limited set of child characteristics (e.g., age, sex, and race/ethnicity).
C245PW0	parent interview data from spring-kindergarten AND spring-first grade AND spring-third grade, alone or in conjunction with any of the child assessment, school, teacher, or classroom data.
C1_5FC0	child direct assessment data from FOUR rounds of data collections involving the FULL sample of children (fall-kindergarten, spring-kindergarten, spring-first grade, and spring-third grade), alone or in conjunction with any of the school, teacher, or classroom data, or a limited set of child characteristics (e.g., age, sex, and race/ethnicity).

See notes at end of table.

Exhibit 10-2. Use of the K–5 longitudinal weights: School years 1998–99, 1999–2000, 2001–02, and 2003-04—Continued

K–5 longitudinal (panel) weights	to be used for analysis of …
C1_5FP0	parent interview data from FOUR rounds of data collections involving the FULL sample of children (fall-kindergarten, spring-kindergarten, spring-first grade, and spring-third grade), alone or in conjunction with any of the child assessment, school, teacher, or classroom data.
C1_5SC0	child direct assessment data from FIVE rounds of data collection (fall-kindergarten, spring-kindergarten, fall-first grade, spring-first grade, and spring-third grade), alone or in conjunction with any of the school, teacher, or classroom data, or a limited set of child characteristics (e.g., age, sex, and race/ethnicity).
C1_5SP0	parent interview data from ALL FIVE rounds of data collection (fall-kindergarten, spring-kindergarten, fall-first grade, spring-first grade, and spring-third grade), alone or in conjunction with any of the child assessment, school, teacher, or classroom data.
C56CW0	child direct assessment data from BOTH spring-third grade and spring-fifth grade, alone or in conjunction with any of the school, teacher, or classroom data, or a limited set of child characteristics (e.g., age, sex, and race/ethnicity).
C56PW0	parent interview data from BOTH spring-third grade AND spring-fifth grade, alone or in conjunction with any of the child assessment, school, teacher, or classroom data.
C456CW0	child direct assessment data from spring-first grade AND spring-third grade AND spring-fifth grade, alone or in conjunction with any of the school, teacher, or classroom data, or a limited set of child characteristics (e.g., age, sex, and race/ethnicity).
C456PW0	parent interview data from spring-first grade AND spring-third grade AND spring-fifth grade, alone or in conjunction with any of the child assessment, school, teacher, or classroom data.
C2_6FC0	child direct assessment data from FOUR rounds of data collection involving the FULL sample of children (spring-kindergarten, spring-first grade, spring-third grade, and spring-fifth grade), alone or in conjunction with any of the school, teacher, or classroom data, or a limited set of child characteristics (e.g., age, sex, and race/ethnicity).
C2_6FP0	parent interview data from FOUR rounds of data collection involving the FULL sample of children (spring-kindergarten, spring-first grade, spring-third grade, and spring-fifth grade, alone or in conjunction with any of the child assessment, school, teacher, or classroom data.
C1_6FC0	child direct assessment data from FIVE rounds of data collections involving the FULL sample of children (fall-kindergarten, spring-kindergarten, spring-first grade, spring-third grade, and spring-fifth grade), alone or in conjunction with any of the school, teacher, or classroom data, or a limited set of child characteristics (e.g., age, sex, and race/ethnicity).
C1_6FP0	parent interview data from FIVE rounds of data collections involving the FULL sample of children (fall-kindergarten, spring-kindergarten, spring-first grade, spring-third grade, and spring-fifth grade), alone or in conjunction with any of the child assessment, school, teacher, or classroom data.
C1_6SC0	child direct assessment data from ALL SIX rounds of data collection (fall-kindergarten, spring-kindergarten, fall-first grade, spring-first grade, spring-third grade, and spring-fifth grade), alone or in conjunction with any of the school, teacher, or classroom data, or a limited set of child characteristics (e.g., age, sex, and race/ethnicity).
C1_6SP0	parent interview data from ALL SIX rounds of data collection (fall-kindergarten, spring-kindergarten, fall-first grade, spring-first grade, spring-third grade, and spring-fifth grade), alone or in conjunction with any of the child assessment, school, teacher, or classroom data.

SOURCE: U.S. Department of Education, National Center for Education Statistics, Early Childhood Longitudinal Study, Kindergarten Class of 1998–99 (ECLS-K), fall 1998, spring 1999, fall 1999, spring 2000, spring 2002, and spring 2004.

K–5 longitudinal weights are used to produce estimates of differences between two or more rounds of data collection spanning kindergarten, first grade, third grade, and fifth grade. Simple examples involving two rounds of data collection are as follows: (1) estimating the differences in children's mean assessment scores between spring-third grade and spring-fifth grade using C56CW0 and (2) estimating the difference in Social Rating Scale scores as reported by parents in spring-kindergarten and spring-first grade using C24PW0 (Social Rating Scale scores as reported by parents are not available for fall-first grade, spring-third grade or spring-fifth grade). K–5 longitudinal weights are also used to study the characteristics of children who were assessed in two or more rounds of data collection. For example, one can study the characteristics of kindergarten children that are associated with the greatest gains in learning in third and fifth grades. If the analysis includes data collected from the parents in spring-third grade and spring-fifth grade, then C56PW0 can be used in the analysis. However, if the analysis involves only the key characteristics (e.g., race) available for most children and the child assessment data from spring-third grade and spring-fifth grade, then C56CW0 can be used to estimate changes in assessment scores between spring-third grade and spring-fifth grade. An example in which data from more than two rounds are used is as follows: to examine whether the gains children have made in their reading knowledge and skills during the kindergarten year and from the end of kindergarten to the end of first grade are related to parents' and teachers' beliefs about kindergarten readiness and parental educational expectations, the weight Y2COMW0 would be appropriate. As noted in the first-grade, third-grade, and fifth-grade user's manuals, any longitudinal analysis that uses data from fall-first grade will be limited to a 27 percent subsample of children.[2]

There may be combinations of data for which no weights were developed. For example, there is no specific weight to study changes in children's classroom environments as they move from kindergarten to fifth grade if child assessment or parent data are not used in the analysis. In this example, the data come from the teacher-level teacher's questionnaire (TQA in kindergarten, first grade, and third grade, and teacher-level teacher questionnaire in fifth grade). The preferred weight for this analysis would be C2_6FC0, which is the weight for child direct assessment data from spring-kindergarten, spring-first grade, spring-third grade, and spring-fifth grade. Of children on the longitudinal K–5 file with teacher-level data in all four of these rounds (e.g., spring-kindergarten, spring-first grade, spring-third grade, and spring-fifth grade), 98 percent (8,081) have nonzero C2_6FC0, compared with 90 percent (7,382) with nonzero C1_6FC0 and 26 percent (2,107) with nonzero C1_6SC0, the other two longitudinal weights available for analyses of child data. The preferred weight is the one that will yield the largest number of

[2] As described in the first grade user's manual, fall-first grade was a design enhancement to enable researchers to study the extent of summer learning losses and gains and the factors associated with them. The fall data collection was limited to students in a 30 percent subsample of schools.

records for analysis, which in this case is C2_6FC0. Analytically, it can be argued that since the direct assessments are conducted in schools, this weight comes closest to capturing the children in participating schools and thus to capturing the children with relevant school environment data. Similarly, if data from the school administrator's questionnaire are used in the analysis of the K–5 longitudinal data, then the same arguments can be used to select the weight. In this case, 44 percent of children in the K–5 have school administrator questionnaire data from kindergarten, first grade, third grade, and fifth grade; of these, 97 percent have nonzero C2_6FC0 compared with 88 percent with nonzero C1_6FC0 and 25 percent with nonzero C1_6SC0. Therefore, the preferred weight is also C2_6FC0. For further advice on which weights to use when analyzing a complex combination of data, contact NCES at *ECLS@ed.gov*.

10.5 Characteristics of Longitudinal Weights

The statistical characteristics of the longitudinal weights are presented in table 10-10. For each weight, the number of cases with nonzero values is presented together with the mean weight, the standard deviation, the coefficient of variation (i.e., the standard deviation as a percentage of the mean weight), the minimum value of the weight, the maximum value of the weight, the skewness, the kurtosis, and the sum of weights.

The difference in the estimate of the population of students (sum of weights) between the different panels of students and types of weights results from a combination of factors, among them: (1) the number of base-year respondents who became ineligible (due to death, leaving the country, or being a nonsampled mover) after the base year; (2) the adjustment of the weights for the children of unknown eligibility; and (3) the difference in the number of records used to construct sample-based control totals. Of the longitudinal weights computed in third grade and fifth grade, six weights (C45CW0, C45PW0, C56CW0, C56PW0, C456CW0 and C456PW0) involve children sampled in first grade. For these weights, the child records included in the file used for computing the control totals are records of base-year respondents and records of eligible children sampled in first grade. For all other longitudinal weights, records of children sampled in first grade were not included in the file, causing the sum of weights to be smaller.

For information about the development of the longitudinal weights, see chapter 9 of the first-grade, third-grade, and fifth-grade user's manuals.

Table 10-10. Characteristics of child-level K–5 longitudinal weights: School years 1998–99, 1999–2000, 2001–02, and 2003-04

Variable name	Number of cases	Mean	Standard deviation	CV (× 100)	Minimum	Maximum	Skewness	Kurtosis	Sum
C23CW0	5,216	739.84	587.55	79.42	68.23	7,182.37	3.98	21.56	3,858,997
C23PW0	4,861	793.83	515.75	64.97	84.26	5,853.21	2.97	13.04	3,858,805
C123CW0	4,729	815.99	646.25	79.20	76.08	7,696.79	3.89	21.55	3,858,824
C123PW0	4,295	898.37	597.89	66.55	95.35	6,421.30	3.05	14.20	3,858,492
C24CW0	16,371	234.81	200.69	85.47	1.78	3,272.40	4.22	31.65	3,844,009
C24PW0	14,938	257.25	198.94	77.34	1.93	2,580.41	3.30	19.64	3,842,784
C124CW0	15,001	256.28	228.52	89.17	1.54	3,877.43	3.71	24.60	3,844,472
C124PW0	13,413	286.40	214.80	75.00	2.06	3,275.79	3.84	26.53	3,841,463
C1_4CW0	4,542	847.78	639.83	75.47	77.56	7,528.68	3.49	18.68	3,850,619
C1_4PW0	4,012	959.07	617.93	64.43	108.75	6,780.92	2.86	13.48	3,847,785
Y2COMW0	13,983	274.83	241.55	87.89	2.03	3,803.82	4.26	29.97	3,842,961
C45CW0	13,964	281.86	273.52	97.04	1.68	3,897.42	3.37	19.90	3,935,960
C45PW0	12,652	310.98	266.89	85.82	1.68	3,718.34	3.11	17.32	3,934,550
C245CW0	13,694	280.68	277.47	98.86	1.65	4,119.55	3.55	22.53	3,843,642
C245PW0	12,204	314.92	267.05	84.80	1.78	3,121.66	2.87	14.51	3,843,272
C1_5FC0	12,558	306.07	303.52	99.17	1.68	4,264.25	3.59	22.83	3,843,607
C1_5FP0	10,998	349.42	299.17	85.62	1.92	3,754.91	3.18	17.88	3,842,954
C1_5SC0	4,032	952.67	875.12	91.86	64.97	7,174.65	3.28	13.78	3,841,183
C1_5SP0	3,522	1,090.37	816.79	74.91	104.68	6,801.61	2.56	9.19	3,840,278
C56CW0	11,136	353.53	546.33	154.54	1.85	6,088.46	4.23	22.14	3,936,880
C56PW0	10,079	390.45	552.94	141.62	1.87	6,635.16	3.81	19.01	3,935,347
C456CW0	10,852	362.33	588.43	162.40	1.78	6,681.37	4.13	20.98	3,932,020
C456PW0	9,568	410.86	582.33	141.73	2.18	5,941.85	3.68	16.93	3,931,097
C2_6FC0	10,673	359.60	596.79	165.96	1.75	6,360.58	4.25	22.07	3,838,004
C2_6FP0	9,267	414.05	585.96	141.52	2.19	5,945.74	3.59	15.69	3,836,967
C1_6FC0	9,796	391.72	651.89	166.41	1.62	6,867.64	4.21	21.76	3,837,337
C1_6FP0	8,370	458.36	646.59	141.06	2.16	6,801.76	3.62	16.27	3,836,496
C1_6SC0	3,000	1,274.18	1,841.67	144.54	58.68	11,913.28	3.28	11.10	3,822,526
C1_6SP0	2,566	1,490.10	1,835.53	123.18	86.76	10,279.37	2.71	7.31	3,823,589

SOURCE: U.S. Department of Education, National Center for Education Statistics, Early Childhood Longitudinal Study, Kindergarten Class of 1998–99 (ECLS-K), fall 1998, spring 1999, fall 1999, spring 2000, spring 2002, and spring 2004.

REFERENCES

Durbin, J. (1967). Design of Multi-stage Surveys for the Estimation of Sampling Errors. *Journal of the Royal Statistical Society*, Series C (Applied Statistics), Vol. 16, 152–164.

Gresham, F.M. and Elliott, S.N. (1990). *Social Skills Rating Scale: Elementary Scale A ("How Often?").* Circle Pines, MN: American Guidance Service, Inc..

Kish, L. (1965). *Survey Sampling.* New York: John L. Wiley & Sons.

Linacre, J., and Wright, B. (2000). WINSTEPS. Chicago: MESA Press.

Marsh, H.W. (1992). *Self-Description Questionnaire — I.* Campbelltown, NSW, Australia: University of Western Sydney, Macarthur.

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* (p. 31–49). Washington, DC: Economic Research Service, U.S. Department of Agriculture.

Pollack, J.M., Rock, D.A., Weiss, M.J., and Atkins-Burnett, S. (2005). *Early Childhood Longitudinal Study, Kindergarten Class of 1998–99 (ECLS-K), Psychometric Report for the Third Grade* (NCES 2005–062). U.S. Department of Education. Washington, DC: National Center for Education Statistics.

Pollack, J.M., Najarian, M., Rock, D.A., and Atkins-Burnett, S. (2005). *Early Childhood Longitudinal Study, Kindergarten Class of 1998–99 (ECLS-K), Psychometric Report for the Fifth Grade* (NCES 2006–036). U.S. Department of Education. Washington, DC: National Center for Education Statistics.

Research Triangle Institute. (2004). *SUDAAN Language Manual, Release 9.0.* Research Triangle Park, NC: Research Triangle Institute.

Rock, D. A., and Pollack, J. M. (2002a). *A Model-based Approach to Measuring Cognitive Growth in Pre-reading and Reading Skills During the Kindergarten Year* (ETS Research Report RR-02-18). Princeton, NJ: Educational Testing Service.

Rock, D. A., and Pollack, J. M. (2002b). *Early Childhood Longitudinal Study, Kindergarten Class of 1998–99 (ECLS-K), Psychometric Report for Kindergarten Through First Grade* (NCES 2002–05). U.S. Department of Education. Washington, DC: National Center for Education Statistics.

Tourangeau, K., Brick, M., Byrne, L., Lê, T., Nord, C., West, J., and Germino Hausken, E. (2004). *Early Childhood Longitudinal Study, Kindergarten Class of 1998–99 (ECLS-K), Third Grade Methodology Report* (NCES 2005–018). U.S. Department of Education. Washington, DC: National Center for Education Statistics.

Tourangeau, K., Brick, M., Lê, T., Wan, S., Weant, M., Nord, C., Vaden-Kiernan, N., Hagedorn, M., Bissett, E., Dulaney, R., Fowler, J., Pollack, J., Rock, D., Weiss, M.J., Atkins-Burnett, S., Germino Hausken, E., West, J., Rathbun, A., and Walston, J. (2004). *Early Childhood Longitudinal Study, Kindergarten Class of 1998–99 (ECLS-K), User's Manual for the ECLS-K Third Grade Public-Use Data File and Electronic Code Book* (NCES 2004–001). U.S. Department of Education. Washington, DC: National Center for Education Statistics.

Tourangeau, K., Burke, J., Lê, T., Wan, S., Weant, M., Brown, E., Vaden-Kiernan, N., Rinker, E., Dulaney, R., Ellingsen, K., Barrett, B., Flores-Cervantes, I., Zill, N., Pollack, J., Rock, D., Atkins-Burnett, S., Meisels, S, Bose, J., West, J., Denton, K., Rathbun, A., and Walston, J. (2001a). *Early Childhood Longitudinal Study, Kindergarten Class of 1998–99 (ECLS-K), Base Year Restricted-Head Start File* (NCES 2001–025). U.S. Department of Education. Washington, DC: National Center for Education Statistics.

Tourangeau, K., Burke, J., Lê, T., Wan, S., Weant, M., Brown, E., Vaden-Kiernan, N., Rinker, E., Dulaney, R., Ellingsen, K., Barrett, B., Flores-Cervantes, I., Zill, N., Pollack, J., Rock, D., Atkins-Burnett, S., Meisels, S, Bose, J., West, J., Denton, K., Rathbun, A., and Walston, J. (2001b). *Early Childhood Longitudinal Study, Kindergarten Class of 1998–99 (ECLS-K), Base Year Restricted-Use Salary and Benefits File* (NCES 2001–014). U.S. Department of Education. Washington, DC: National Center for Education Statistics.

Tourangeau, K., Burke, J., Lê, T., Wan, S., Weant, M., Brown, E., Vaden-Kiernan, N., Rinker, E., Dulaney, R., Ellingsen, K., Barrett, B., Flores-Cervantes, I., Zill, N., Pollack, J., Rock, D., Atkins-Burnett, S., Meisels, S, Bose, J., West, J., Denton, K., Rathbun, A., and Walston, J. (2001c). *Early Childhood Longitudinal Study, Kindergarten Class of 1998–99 (ECLS-K), Base Year Restricted-Use Special Education Child File* (NCES 2001–015). U.S. Department of Education. Washington, DC: National Center for Education Statistics.

Tourangeau, K., Burke, J., Lê, T., Wan, S., Weant, M., Brown, E., Vaden-Kiernan, N., Rinker, E., Dulaney, R., Ellingsen, K., Barrett, B., Flores-Cervantes, I., Zill, N., Pollack, J., Rock, D., Atkins-Burnett, S., Meisels, S, Bose, J., West, J., Denton, K., Rathbun, A., and Walston, J. (2001d). *Early Childhood Longitudinal Study, Kindergarten Class of 1998–99 (ECLS-K), Base Year Restricted-Use Student Record Abstract File* (NCES 2001–016). U.S. Department of Education. Washington, DC: National Center for Education Statistics.

Tourangeau, K., Burke, J., Lê, T., Wan, S., Weant, M., Brown, E., Vaden-Kiernan, N., Rinker, E., Dulaney, R., Ellingsen, K., Barrett, B., Flores-Cervantes, I., Zill, N., Pollack, J., Rock, D., Atkins-Burnett, S., Meisels, S, Bose, J., West, J., Denton, K., Rathbun, A., and Walston, J. (2004). *Early Childhood Longitudinal Study, Kindergarten Class of 1998–99 (ECLS-K), Base Year Public-Use Data Files and Electronic Codebook: User's Manual* (NCES 2001–029r). U.S. Department of Education. Washington, DC: National Center for Educational Statistics.

Tourangeau, K., Burke, J., Lê, T., Wan, S., Weant, M., Nord, C., Vaden-Kiernan, N., Bissett, E., Dulaney, R., Fields, A., Byrne, L., Flores-Cervantes, I., Fowler, J., Pollack, J., Rock, D., Atkins-Burnett, S., Meisels, S., Bose, J., West, J. Denton, K., Rathbun, A., and Walston, J. (2002). *Early Childhood Longitudinal Study, Kindergarten Class of 1998–99 (ECLS-K), User's Manual for the ECLS-K First Grade Public-Use Data Files and Electronic Codebook* (NCES 2002–135). U.S. Department of Education. Washington, DC: National Center for Education Statistics.

Tourangeau, K., Nord, C., Lê, T., Wan, S., Bose, J., and West, J. (2002). *Early Childhood Longitudinal Study, Kindergarten Class of 1998–99 (ECLS-K), User's Manual for the ECLS-K Longitudinal Kindergarten–First Grade Public-Use Data File* (NCES 2002–149). U.S. Department of Education, Washington, DC: National Center for Education Statistics.

Tourangeau, K., Nord, C., Lê, T., Wan, S., and West, J. (2004). *Early Childhood Longitudinal Study, Kindergarten Class of 1998–99 (ECLS-K), User's Manual for the ECLS-K Longitudinal Kindergarten–Third Grade Public-Use Data File and Electronic Code Book* (NCES 2004–088). U.S. Department of Education. Washington, DC: National Center for Education Statistics.

Tourangeau, K., Lê, T., and Nord, C. (2005). *Early Childhood Longitudinal Study, Kindergarten Class of 1998–99 (ECLS-K), Fifth-Grade Methodology Report* (NCES 2006–037). U.S. Department of Education. Washington, DC: National Center for Education Statistics.

U.S. Department of Agriculture, Food and Nutrition Service. (2000). *Guide to Measuring Household Food Security, Revised 2000*. Alexandria, VA: Author.

U.S. Department of Agriculture, Economic Research Service. (2002). *Measuring Children's Food Security in U.S. Households, 1995-99.* Washington, DC: Author.

U.S. Department of Commerce, Bureau of the Census. (1982). *1980 Census of Population—Index of Industries and Occupations, Final Edition.* Washington, DC: U.S. Government Printing Office.

U.S. Department of Commerce, Office of Federal Statistical Policy and Planning. (1980). *Standard Occupational Classification Manual.* Washington, DC: U.S. Government Printing Office.

U.S. Department of Education, National Center for Educational Statistics (1995–96). *Common Core of Data, Public School Universe Survey*. Washington, DC: Author.

U.S. Department of Education, National Center for Education Statistics. (1998). *Private School Universe Survey, 1995–96*. Washington, DC: Author.

U.S. Department of Education, National Center for Education Statistics. (1999). *Manual for Coding Industries and Occupations.* Appendix to Data File User's Manual, Volume IV, of the National Household Education Survey of 1999 (NCES 2000–077). Washington, DC: Author.

U.S. Department of the Interior, Bureau of Indian Affairs. (1995–96). *Office of Indian Education Programs Education Directory*. Washington, DC: Author.

Walston, J., and West, J. (2004). *Full-day and Half-day Kindergarten in the United States: Findings from the Early Childhood Longitudinal Study, Kindergarten Class of 1998-99* (NCES 2004–078). U.S. Department of Education. Washington, DC: National Center for Education Statistics.

Westat (2001). *WesVar 4.0 User's Guide.* Rockville, MD: Author.

Wolter, K. (1985). *Introduction to Variance Estimation.* New York: Springer-Verlag.

Woodcock, R.W., McGrew, K.S., and Werder, J.K. (1994). *Woodcock-McGrew-Werder Mini-Battery of Achievement.* Itasca, IL: Riverside Publishing.

This page is intentionally left blank.