ies NATIONAL CENTER FOR EDUCATION STATISTICS
Institute of Education Sciences

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

Combined User's Manual for the ECLS-K Eighth-Grade and K–8 Full Sample Data Files and Electronic Codebooks

NCES 2009–004

U.S. DEPARTMENT OF EDUCATION

ies NATIONAL CENTER FOR EDUCATION STATISTICS
Institute of Education Sciences

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

Combined User's Manual for the ECLS-K Eighth-Grade and K–8 Full Sample Data Files and Electronic Codebooks

February 2009

Karen Tourangeau
Christine Nord
Thanh Lê
Alberto G. Sorongon
Westat

Michelle Najarian
Educational Testing Service

Elvira Germino Hausken
Project Officer
National Center for Education Statistics

NCES 2009–004

U.S. DEPARTMENT OF EDUCATION

U.S. Department of Education
Arne Duncan
Secretary

Institute of Education Sciences
Sue Betka
Acting Director

National Center for Education Statistics
Stuart Kerachsky
Acting 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 2009

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.

This report is available only on the Web at http://nces.ed.gov/pubsearch.

Suggested Citation

Tourangeau, K., Nord, C., Lê, T., Sorongon, A. G., and Najarian, M. (2009). *Early Childhood Longitudinal Study, Kindergarten Class of 1998–99 (ECLS-K), Combined User's Manual for the ECLS-K Eighth-Grade and K–8 Full Sample Data Files and Electronic Codebooks* (NCES 2009–004). National Center for Education Statistics, Institute of Education Sciences, U.S. Department of Education. Washington, DC.

Content Contact
Gail Mulligan
(202) 502-7491
gail.mulligan@ed.gov

CONTENTS

CONTENTS (continued)

CONTENTS (continued)

CONTENTS (continued)

CONTENTS (continued)

CONTENTS (continued)

CONTENTS (continued)

CONTENTS (continued)

CONTENTS (continued)

CONTENTS (continued)

List of Tables

CONTENTS (continued)

List of Tables (continued)

CONTENTS (continued)

List of Tables (continued)

CONTENTS (continued)

List of Tables (continued)

CONTENTS (continued)

List of Tables (continued)

CONTENTS (continued)

List of Tables (continued)

CONTENTS (continued)

List of Tables (continued)

CONTENTS (continued)

List of Tables (continued)

List of Exhibits

CONTENTS (continued)

List of Exhibits (continued)

CONTENTS (continued)

List of Exhibits (continued)

List of Exhibits (continued)

CONTENTS (continued)

List of Exhibits (continued)

This page is intentionally left blank.

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 directs 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, hereinafter referred to as the User's Manual. In this chapter, major differences between the eighth-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, two files are available for analyzing eighth-grade data: (1) a restricted-use data file containing information collected during the eighth-grade round and recalibrated assessment scores for all rounds and (2) a kindergarten–eighth grade (K–8) full sample public-use data file that has been produced in the place of both an eighth-grade public-use file and a K–8 longitudinal file. As described in chapter 10, the full sample data file can be used for within-year analyses of any round of data collection from kindergarten through eighth grade, and it also can be used for any combination of cross-year analyses.

This manual serves as a guide for users of both of these files. Most of the User's Manual chapters apply to both the public-use and the restricted-use data files, but a few sections apply to only one of the two. Exhibit A summarizes the User's Manual sections that do not apply to both files and indicates the data file 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 both data files.

In preparing public-use data files, the National Center for Education Statistics (NCES) takes steps to minimize the likelihood that an individual school, teacher, parent, or child[1] 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 eighth-grade restricted-use file have been made to the K–8 full sample public-use data 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 data file. Chapter 1, section 1.4.1, provides a general description of the differences between public-use and

[1] To be consistent with documentation from earlier rounds of the ECLS-K, this manual refers to student respondents in the eighth-grade round as "children."

restricted-use files. Table 7-16 in chapter 7 contains a list of eighth-grade variables that have been modified. Section 7.10 contains additional information about the "masking" process.

Exhibit A. Sections of User's Manual that do not apply to both data files

Section	Description	Data file to which section applies
► *Please note …*		
7.9: 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 eighth-grade data in the K–8 full sample public-use data file is contained in the last column of table 7-15.
7.10	Masked variables	Eighth-grade data in the public-use K–8 full sample file
9.4	Merging base-year, first-, third-, fifth-, and eighth-grade data	Eighth-grade restricted-use file
10	Using the kindergarten–eighth-grade full sample file	This chapter applies to users of the K–8 full sample public-use data file

Major Differences in the Eighth-Grade Data Collection and Release

Although the eighth-grade data collection shares many similarities with earlier rounds, some modifications were made to capture important information relevant to children in eighth grade. The major differences between the eighth-grade data collection and the earlier rounds are summarized below:

- **Parent data were collected in the fall rather than in the spring, as was the method in previous rounds.** Because the data were collected at the beginning of the school year, items tapping parent involvement in various school functions were followed by items asking whether parents had yet had an opportunity to be involved in those functions.

- **New construct areas were added to the parent interview for eighth grade.** These new construct areas included the following:
  - expectations of how far child will go in school;
  - family activities (e.g., working on homework together, going shopping, attending concerts, plays, or movies);
  - family rules (e.g., rules new to round 7 are about the child maintaining a certain grade point average, doing homework, and hours spent on the computer or playing video games);
  - parent monitoring (e.g., checking homework, having and enforcing a curfew);
  - days per week that child has adult supervision after school;
  - parent reading habits;
  - child's use of tutors in science or English/Language Arts;
  - parent discussions with child (e.g., about courses at school, events);
  - characteristics of parent's relationship with child;
  - child performance in school;
  - whether school is in the assigned district;
  - school suspension;
  - parent perceptions of and satisfaction with the school;
  - characteristics of parent's relationship with spouse;
  - parent religious practices;
  - parent political views;
  - nonresident parent contribution to medical and other expenses;
  - child health questions regarding depression, weight and eating disorders, diabetes, and various treatments (e.g., medicine, individual therapy);
  - child internalizing and externalizing problems;
  - parent depression (the same questions were used in round 2 of the study);
  - stressful life events;

- home ownership, value, and mortgage debt; and
- savings for post-high school education.

- **The sample of children included on the K-8 longitudinal public-use data file differs from the sample included in prior ECLS-K longitudinal files.** In each of the previous ECLS-K longitudinal files, children were included if they had at least one nonzero weight among the weights computed for the rounds included in the longitudinal file. However, the K–8 longitudinal public-use data file included any child who was ever sampled in the base year who had base-year data, and any child sampled in the first-grade year who had at least one round of data in first grade and beyond.

- **In eighth grade, children were assessed in proctored group settings rather than one on one.** In earlier rounds, the mathematics, reading, and science assessments were conducted via one-on-one direct assessment. In the eighth grade, however, children were expected to be familiar with proctored testing in school. Thus, groups of ECLS-K sampled children who attended the same school were assessed in a single, proctored group administration. The content changes of the assessment are described in section 2.1.2.

- **Two-level (high versus low) second-stage assessment forms were used, rather than three-level forms used in previous rounds.** In the eighth-grade timed assessment session, all children were given separate routing tests in each subject area to determine the level (high versus low) of their second-stage reading, mathematics, and science assessments. Routing children into two, rather than three, second-stage forms facilitated accurate and efficient distribution of the second-stage forms. Results of the spring 2006 field test showed that there was no loss of data by using a two-level second-stage form. Information on the results of the spring 2006 field test can be found in the *ECLS-K Methodology Report for the Eighth Grade* (NCES 2009–003) (Tourangeau et al. forthcoming). Information on the quality of the eighth-grade assessment data can be found in the *ECLS-K Psychometric Report for the Eighth Grade* (NCES 2009–002) (Najarian, Pollack, and Sorongon forthcoming).

- **Age-appropriate changes were made to the rating items used to tap children's perceptions of their social skills, interest in school subjects, self-concept, and control they had over their own lives.** In the kindergarten and first-grade rounds of the ECLS-K, parents and teachers reported on children's social skills. In the third and fifth grade of the ECLS-K, the children provided information about themselves by completing a short self-description questionnaire that included items from a published instrument appropriate for third- and fifth-graders (Self Description Questionnaire I) (Marsh 1992a). In eighth grade, a new version of the self-description questionnaire was developed using items from a published instrument designed to be used with adolescents (Self Description Questionnaire II) (Marsh 1992b). See sections 2.1.1, 3.3, and 3.4 for additional information on the eighth-grade self-description questionnaire. In addition, two scales from the student questionnaire adapted from the National Education Longitudinal Study (NELS:88) tapped children's self-concept and their perceptions of how much control they had over their own lives. See sections 3.3

and 3.4 for more information on these scales and the scores that are available for analysis.

- **The procedures for collecting height and weight data were modified.** In the previous rounds of the ECLS-K, height and weight data were collected during the one-on-one direct assessment sessions. In the eighth grade, height and weight data were collected during the group assessment sessions. In most cases the groups were small (in many cases there was a single child). However, in some cases, the assessment sessions had several children participating. In the group assessment sessions, children were measured one at a time at a single height and weight station. The average size of the assessment group was three children and ranged from one to nine children per group. See section 5.5.2 or the *ECLS-K Eighth-Grade Methodology Report* (NCES 2009–003) (Tourangeau et al. forthcoming) for additional information on the height and weight data collection.

- **In eighth grade, children completed self-administered paper and pencil questionnaires about their school experiences, their activities, their perceptions of themselves, and their weight, diet, and level of exercise.** This questionnaire was completed during the group assessment session.

- **The Academic Rating Scale (ARS) was replaced with other items tapping children's classroom behavior and performance.** English, mathematics, and science teachers were asked to rate children on their respective domain-relevant skills. Teachers also rated children on their effort (e.g., "Does this student usually work hard for good grades in your class?"), behavior (e.g., "Does this student seem to relate well to other students in your class?"), and attendance (e.g., "How often is this student absent from your class?"). Teachers also were asked to report if they had either spoken to a guidance counselor regarding a child's poor performance or if they had recommended children for academic honors or advanced placement. Information on the scaling of these items can be found in section 3.2.

- **Information about children's food consumption was collected through a self-administered questionnaire**. In previous rounds, the assessor read the questionnaire items for the children and recorded their responses. In the eighth-grade round, the food consumption items were included in the self-administered questionnaire completed during the group assessment session.

- **Collection of school record abstracts and school facilities checklists was discontinued.** These instruments were discontinued due to cost constraints and low response rates in prior rounds. Items associated with Individualized Education Programs (IEPs) that were collected from school record abstracts in previous rounds were collected in the special education teacher questionnaire (B).

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 website for commonly asked questions and responses: http://nces.ed.gov/ecls.

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

- **The sample is not representative of children in eighth grade, 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 children who had not attended kindergarten in the United States in the previous year, the first-grade sample is representative of children attending first grade 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 eighth-grade sample is not representative of children in eighth grade, eighth-grade teachers, or schools with eighth grades. Children who started their schooling in the U.S. after first grade are not represented in the sample. The data should not be used to make statements about eighth-graders, schools with eighth grades, or eighth-grade teachers.

- **Not all sample children were in eighth grade.** The eighth-grade data file includes children who were in eighth grade in spring 2007, and others who were either held back (e.g., seventh-graders) or promoted ahead an extra year or more (e.g., ninth-graders). Users should be aware of this fact when using the data and interpreting the findings. Most children in the sample had been in school for at least 9 years (K–8) and some more than 9 years (those who were repeating kindergarten in the base year). A very small number may have been in school less than 9 years (some were part of the freshened sample added in first grade).

- **Child mobility and its consequences.** A random subsample of children 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, 4.5, and 4.6 describe the subsampling of movers. A number of variables on the file can be used to determine if a child moved to a different school between rounds. Section 7.8 describes these variables.

- **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. Five different possible missing data codes are used

on the file. See section 7.4 for a discussion of the different missing values codes and the circumstances in which 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-eighth 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 eighth-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 Eighth Grade* (NCES 2009–002) (Najarian, Pollack, and Sorongon forthcoming) for more information.

- **Use of weights.** The eighth-grade restricted-use data file contains 5 sets of cross-sectional weights and 12 longitudinal (panel) weights. Although a variety of weights exist 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 4.8, 9.3.1, and 10.4.

- **Defining special populations.** The ECLS-K includes a number of analytic groups of interest that can be identified and studied separately. For example, the eighth-grade data file contains variables that identify children who have a disability diagnosed by a professional (P7DISABL) and those who live in households with incomes below the federal poverty threshold (W8POVRTY). 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. Users who wish to study a specific subpopulation should consult the ECLS-K composite variables (table 7-15) or the data collection instruments to identify variables that might help them identify their population of interest.

- **Examining school and classroom effects.** Examination of classroom effects is possible with kindergarten and first-grade data because child assessment data were collected at the start and end of each of these grades. When studying the effects of schools and classrooms, it is important to group the subject children in the same classroom and/or same school. 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 website 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 eighth grade. There are several different approaches to measuring gains. See section 3.1.5 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). (Given the variations of Windows Vista, it is uncertain what issues may be encountered when attempting to run the ECB on this operating system). 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 dataset. 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 ECLS-K CD-ROM. This text data file is referred to as the "child catalog." The **restricted-use eighth grade** child catalog is named child8r.dat in the restricted-use CD-ROM root directory. The **K–8 full sample public-use** child catalog is named childk8p.dat in the public-use DVD root directory. For more information about the data file, see appendix E on the CD-ROM or DVD.

- **Data files.** The eighth-grade **restricted-use** child catalog contains one record for each of 9,725 responding children in spring-eighth grade. The **K–8 full sample public-use** child catalog contains one record for each of the 21,409 children responding in any round from fall-kindergarten to spring-eighth grade data collections. Data collected from teachers and schools are stored in the child catalog. Appendix B on the eighth-grade CD-ROM and DVD contains the data file record layout for the child catalog. It is strongly recommended that users access the data for both data files by using the ECB software available on the CD-ROM and DVD rather than access the ASCII file directly.

- **Identification variables.** The eighth-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 J71T_ID (reading teacher ID) and J72T_ID (mathematics or science teacher ID); schools are identified by the ID variable S7_ID. See section 7.1 in the User's Manual for further information on these identification variables.

- **Instruments.** For the ECLS-K eighth-grade data collection, data were collected using computer-assisted interviewing for parent interviews. Eighth-graders completed cognitive assessments in paper-and-pencil format in timed group administrations. They also completed self-administered paper-and-pencil questionnaires about their school experiences, their activities, their perceptions of themselves, and their weight, diet, and level of exercise. Self-administered questionnaires in paper-and-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 and DVD contains, with some exceptions, the eighth-grade instruments. The exceptions are measures that contain copyright-protected items.

- **Composite variables.** Numerous composite variables have been constructed for the ECLS-K data to make it easier for users to use the dataset. Most composite variables were created using two or more variables that are on the data file or using information from other sources. Other composites 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 the percentage of minority children in class and children's grade level. Composites based on the school data include the percentage of minority children and school type. Composites based on the parent data include parent education, poverty status, and socioeconomic status. See section 7.6 and table 7-15 of the User's Manual for details on all the composites contained on the eighth-grade restricted-use and the K–8 full sample public-use data files. 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 cognitive assessments and measures of children's social development. Chapter 2 provides a general description of the survey instruments, including the cognitive assessments. The eighth-grade cognitive assessment contained items in reading, mathematics, and science. See section 3.1 of the User's Manual for details on the cognitive assessment and the scores that are available for analysis. Section 3.1.4 of the User's Manual discusses choosing the appropriate score for analysis. Section 3.1.5 discusses approaches to measuring gains in child achievement.

  The measures of children's social development consisted of a self-description questionnaire in which the children rated their own perceptions of competence and

interest in reading and mathematics and also reported problem behaviors. In addition, two scales from the student questionnaire adapted from the National Education Longitudinal Study of 1988 (NELS:88) tapped children's self-concept and their perceptions of how much control they had over their own lives. See sections 3.3 and 3.4 for more information on these scales 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 4.8 describes the eighth-grade 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 eighth-grade longitudinal weights is provided in chapter 9. Section 9.3.1 describes the K–8 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–8 longitudinal (panel) weights. Section 10.4 describes the eighth-grade cross-sectional weights and the K–8 longitudinal (panel) weights available on the K–8 full sample public-use data file and provides advice for which weight to use for a given type of analysis.

- **Creating a longitudinal file**. It is possible to merge the eighth-grade restricted-use data with data from earlier rounds. Instructions on how to create such a file are provided in chapter 9, section 9.4.

1. INTRODUCTION

► *Please note that this manual will refer to student respondents in the eighth-grade round as "children" to be consistent with the terminology used in documentation from earlier rounds of the ECLS-K.*

This manual provides guidance and documentation for users of the eighth-grade data[2] 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 and ending with eighth grade. 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. In the eighth-grade data collection, a student paper-and-pencil questionnaire was added. The ECLS-K was developed under the sponsorship of the U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics (NCES). Westat conducted this study with assistance provided by Educational Testing Service (ETS) in Princeton, New Jersey.

The ECLS-K followed a nationally representative cohort of children from kindergarten into middle 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.[3] The fall-first grade data collection was limited to a 30 percent subsample of schools[4] (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

[2] The term "eighth grade" is used throughout this document to refer to the data collections that took place in the 2006–07 school year, at which time most of the sampled children—but not all of them—were in eighth grade.
[3] 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.
[4] Approximately 27 percent of the base-year children who were eligible to participate in year 2 attended the 30 percent subsample of schools.

collection, which included 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 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.[5] 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.[6] 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 academically and socially. 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, 2003–04, and 2006–07

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
Spring-eighth grade	Spring 2007	Full sample

[1] Fall data collection consisted of a 30 percent sample of schools containing approximately 27 percent of the base-year children 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 2007.

[5] Their addition is referred to as "freshening" the sample. See chapter 4, section 4.3.2 for more detail on the freshening process.
[6] 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.

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.[7] In addition to the data collection components used in third grade, children also were asked about their food consumption at school and other places (e.g., home, restaurants) 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.

A seventh wave of data was collected in the spring of the 2006–07 school year when most, but not all, of the sampled children were in eighth grade.[8] In addition to the data collection components used in fifth grade, children were asked to complete a paper-and-pencil questionnaire about their school experiences, their activities, their perceptions of themselves, and their weight, diet, and level of exercise. The spring-eighth grade data collection can be used to measure school progress and to describe the eighth-grade learning environment of children in the study.

The sample of children in the eighth-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 children 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 eighth-grade data are representative of the population cohort rather than all eighth-graders in 2006–07. Comparisons of the weighted population of ECLS-K children enrolled in the eighth grade with the weighted population of eighth-graders reported in the 2006 Current Population Survey[9] suggest that the ECLS-K represents about 80 percent of all U.S. eighth-graders in the 2006–07 school year.[10] Some examples of subpopulations of eighth-graders who are not represented in the ECLS-K in 2006–07 include children who started kindergarten before fall of 1998 and were retained in a later grade, children who immigrated to the United States after first grade, and children who were home-schooled until after first grade. Data were collected from teachers and schools to provide important contextual information about the school environment for the sampled children, but the teachers and schools are not representative of eighth-grade teachers and schools in the country in 2006–07. For this reason, the only weights produced from the study for eighth-grade estimates are for making statements about children, including statements about the teachers and schools of those children.

[7] 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 some other grade (e.g., second, third, or sixth grade).
[8] Approximately 89 percent of the children interviewed were in eighth grade during the 2006–07 school year, 9 percent were in seventh grade, and less than 2 percent were in some other grade (e.g., such as fifth, sixth, or ninth grade).
[9] The Current Population Survey is the monthly survey of households conducted by the Bureau of the Census for the Bureau of Labor Statistics of the U.S. Department of Labor (see http://www.bls.gov/cps/).
[10] The estimate of the percent of eighth-graders captured by the ECLS-K was calculated by dividing the sum of the child weight (C7CW0) by the number of eighth-graders according to the 2006 Current Population Survey.

The ECLS-K has several major objectives and numerous potential applications. The ECLS-K combines (1) a study of achievement in the elementary and middle school 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 elementary and middle school; (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 kindergarten, elementary school, and middle school.

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 middle school. The ECLS-K also provides a rich dataset 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 middle school performance; and to record children's academic growth as they move into middle school.

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.

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.

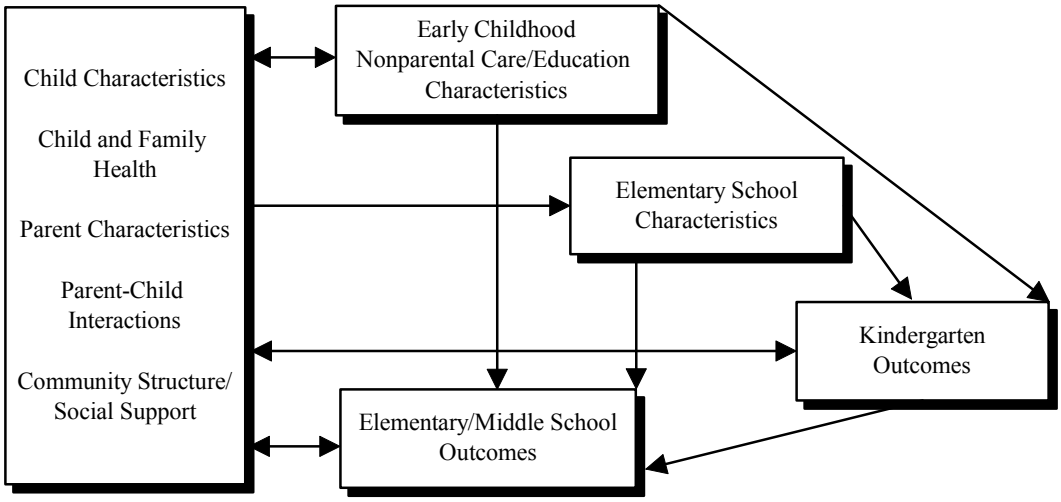
As a study of early achievement, the ECLS-K allows researchers to examine how children's progress is associated with such factors as placement in high or low ability groups, receipt of special services or remedial instruction, grade retention, and frequent changes in schools attended because of family moves. Data on these early school experiences were collected as they occurred, with the exception of their experiences before kindergarten, which were collected retrospectively. Collecting this information as the experiences occurred 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 to later grades.

The spring-eighth grade data collection can be used to describe the diversity of the children in the study and the classrooms and schools they attended. 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 was 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 paid particular attention to the role that parents and families played in helping children adjust to formal school and in supporting their education through the elementary and middle school grades. It 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 included the collection of data from the child, the child's parents/guardians, teachers, and schools.

- **Children** participated in various activities to measure the extent to which they exhibited those abilities and skills deemed important to success in school. They were 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. Children were assessed in each round of data collection. During kindergarten and elementary school, most measures of a child's cognitive skills were obtained through an untimed one-on-one assessment of the child. In the eighth grade, children were assessed in a formal group setting. Beginning with the third-grade data collection, children also reported on their own perceptions of their abilities and achievement, their interest in and enjoyment of reading, mathematics, and other school subjects, their peer relationships, and their own problem behaviors. Children in eighth grade completed a self-administered paper-and-pencil questionnaire about their school experiences, their activities, their perceptions of themselves, and their weight, diet, and level of exercise.

- **Parents/guardians** were an important source of information about the families of the children selected for the study and about themselves. Parents provided information

about children's development at school entry and their experiences both with family members and with others. Information was collected from parents/guardians in each round of data collection.

- **Teachers**, like parents, represented a valuable source of information on themselves, the children in their classrooms, and the children's learning environment (i.e., the classroom). Teachers were not only asked to provide information about their own backgrounds, teaching practices, and experience; they were also called on to provide information on the classroom setting for the sampled children they taught 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 were 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 completed self-administered questionnaires each time children were assessed.

- **School administrators**, or their designees, were asked to provide information on the physical, organizational, and fiscal characteristics of their schools, and on the schools' learning environment and programs. Special attention was paid to the instructional philosophy of the school and its expectations for children. School administrators or their designees were 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- and eighth-grade data collections. Information was collected from school administrators via self-administered questionnaires 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 data 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 public-use data 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 public-use and restricted-use data files:

- Outlier values are top- or bottom-coded;[11]
- 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

[11] 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, 1, 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

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 data files. While very few of the variables are suppressed, some users might require the restricted-use data files. 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 find that the restricted-use data 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 data 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 data files to ascertain whether their specific analytic objectives can be met using those data files.

1.4.2 Overview of Available Data Files

Several ECLS-K data files are available for use by analysts. These are described below beginning with the eighth-grade data files.

- **ECLS-K Eighth-Grade Restricted-Use Data File.** The eighth-grade data are available only as a child-level data file. The file includes all data collected from or about the children and their schools including data from the child assessments and the student, parent, teacher, and school administrator questionnaires. No eighth-grade teacher or school files are released because the sample of teachers and schools is not nationally representative of eighth-grade teachers or schools with eighth grades. Analysts who wish to examine children's experiences in eighth grade and the influence of their classroom or school characteristics on their eighth-grade experiences should use the eighth-grade restricted-use file or the K–8 full sample public-use file described below.

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

- **Kindergarten–Eighth Grade Full Sample Public-Use Data File.** This public-use data file combines data from the base, first-, third-, fifth-, and eighth-grade years. It contains both within-year and cross-year weights so that analysts can examine children's growth and development between kindergarten and eighth grade. Unlike the public-use longitudinal files released in previous rounds, this file contains all data for all ECLS-K sample cases that have been publicly released in any of the rounds. Thus, it can be used for within-year (cross-sectional) analyses of any round of data collection and cross-year (longitudinal) analyses of combinations of rounds. See chapter 10 for details on how to use the K–8 full sample public-use file.

- **ECLS-K Fifth-Grade Restricted- and Public-Use Data Files.** The fifth-grade data are available only as child-level data files. The files include all data collected from or about the children and their schools including data from the child assessments and from their parents, teachers, or schools. No fifth-grade teacher or school files were 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 data file or the K–8 full sample public-use data file.

  The fifth-grade data file can be used not only to analyze data collected in the fifth grade but also to provide 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 were included in 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. However, most analysts will find it more convenient to use the already created K–8 full sample public-use data file described above. For more information on these files, refer to the *ECLS-K Combined User's Manual for the ECLS-K Fifth-Grade Data Files and Electronic Codebooks* (NCES 2006–032) (Tourangeau et al. 2006).

- **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. This file is now superseded by the K–8 full sample public-use data file.

- **ECLS-K Third-Grade Restricted- and Public-Use Data Files.** The third-grade data are available only as child-level data files. The files include 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 were released because the sample of teachers and schools is not nationally representative of third-grade teachers or 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 data file or the K–8 full sample public-use data 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 were included in 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. However, most analysts will find it more convenient to use the already created K–8 full sample public-use data file described above. For more information on these files, refer to the *ECLS-K 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).

- **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. This file is now superseded by the K–8 full sample public-use data file

- **ECLS-K First-Grade Restricted- and Public-Use Data Files.** The first-grade data (fall and spring) are available only as child-level data files. The files include 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 or 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 data file or the K–8 full sample public-use data 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. However, most analysts will find it more convenient to use the already-created K–8 full sample public-use data file described above. For more information on these files, refer to the *ECLS-K User's Manual for the ECLS-K First Grade Public-Use Data Files and Electronic Codebook* (NCES 2002–135) (Tourangeau 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. This file has now been superseded by the K–8 full sample public-use data file.

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

  The child file data 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 children 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

data file or the K–8 full sample public-use data 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 data 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 data file. These child-, teacher-, and school-level data 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 2001–029rev) (Tourangeau, Burke et al. 2004).

- The **Salary and Benefits File** is 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 child 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 et al. 2001b).

- The **Special Education File** is a child-based data 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 et al. 2001c).

- The **Student Records Abstract File** contains information from school records about children's school enrollment and attendance; Individualized Education Program (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 data 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 data file. For more information about this file, see the *ECLS-K Base Year Restricted-Use Student Record Abstract File* (NCES 2001–016) (Tourangeau 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 data file. For more information about this file, see the *ECLS-K Base Year Restricted-Use Head Start File* (NCES 2001–025) (Tourangeau 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.

- The **Census Data and Geocoded Location File** contains census tract and ZIP Code tabulation area (ZCTA) codes for ECLS-K children's homes and schools for each round of the ECLS-K up to third grade. It also has about 600 census variables (or census-derived variables) for each census tract and ZCTA including income, race/ethnicity, and many other sociodemographic characteristics. Supporting documentation included on the CD consists of a user's manual, data file record layouts describing the variables on each of the ASCII data files, and SAS code for converting the data files. This file is a restricted-use data file available upon request from the Institute of Education Sciences Security Data Officer.

1.5 Contents of Manual

This manual provides documentation for users of the eighth-grade data files (the eighth-grade restricted-use data file and the K–8 full sample public-use data file) of the ECLS-K. Prior to fifth grade, separate manuals were issued for each data file. Please refer to the previous chapter, *Getting Started*, for a summary of which sections of the manual do not apply to both files and for an overview of the major differences between the eighth-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 eighth-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, fifth-grade, and eighth-grade files (chapter 9). Finally, chapter 10 presents information on the kindergarten–eighth grade full sample public-use data file. The Electronic Codebook contains unweighted frequencies for all variables. Because this manual focuses on the eighth-grade data collection, minimal information is provided about the base-year, first-grade, third-grade, or fifth-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 2001–029rev) (Tourangeau, Burke et al. 2004); the *User's Manual for the ECLS-K First Grade Public-Use Data Files and Electronic Codebook* (NCES 2002–135) (Tourangeau et al. 2002), 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), or the *Combined User's Manual for the ECLS-K Fifth-Grade Data Files and Electronic Codebooks* (NCES 2006–032) (Tourangeau et al. 2006). 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 Early Childhood Longitudinal Study, Kindergarten Class of 1998–99 (ECLS-K) eighth-grade data collection instruments. The ECLS-K eighth-grade data collection instruments consisted of eight questionnaires (student, parent, teacher, special education teacher/service, and school administrator), three achievement tests (reading, mathematics, and science), and one physical measurement record form.

The eighth-grade data collection instruments, with the exception of the assessments and the items adapted from the Self Description Questionnaire II (Marsh 1992)[12] in the student questionnaire, are available on the ECLS-K DVD and CD-ROM as appendix A. The assessments and Self Description Questionnaire II items contain copyright-protected materials.

For information on the data collection instruments used in any of the past rounds of the ECLS-K, please refer to chapter 2 of the ECLS-K base-year, first-grade, third-grade, fifth-grade, and eighth-grade user's manuals. These can be found on the Web at http://nces.ed.gov/pubsearch.

2.1 Child Assessments and Questionnaire

The child assessments were paper-and-pencil assessments administered in small group settings timed and proctored by a trained test administrator in the spring of the 2006–07 school year. Children were assessed with the same assessment regardless of whether they were on grade level (i.e., in eighth grade). As in the previous rounds, the eighth-grade assessments included cognitive and physical (i.e., height and weight) components. In addition, a self-administered student questionnaire was completed during the eighth-grade assessment session. This included an adaptation of the Self Description Questionnaire (SDQ) II (Marsh 1992b) and the Self-Concept and Locus of Control scales from the National Education Longitudinal Study of 1988 (NELS:88), with questions about children's socioemotional development. The questionnaire also included questions about children's activities in and out of school and their relationships with their friends and parents. Items about children's food consumption were included in the "Your Diet" section, with questions about the kinds of food they could

[12] The student questionnaire used items adapted with permission from the *Self Description Questionnaire (SDQ) II (*Marsh, 1992).

buy at school and the food that they had eaten in the past week. The entire assessment session was 2 hours in duration.

Chapter 3 contains a detailed description of the assessment scores and information on their use and interpretation.

2.1.1 Cognitive Assessments

The ECLS-K eighth-grade direct cognitive assessment battery was designed to assess children's academic achievement in spring of eighth grade, and to provide a means of measuring academic growth since kindergarten entry. A panel of child development, middle school education, and content area experts recommended that the knowledge and skills assessed by the ECLS-K eighth-grade assessments should represent the typical and important academic goals of middle school curricula in English, mathematics, and science. Reading, mathematics, and science were the three cognitive domains assessed in the eighth grade.

While the direct cognitive assessments were individually administered at all six previous time points, in spring-eighth grade, groups of ECLS-K sampled children who attended the same school were assessed in a single, proctored group administration. All children were given separate routing assessment forms to determine the level (high/low) of their reading, mathematics, and science assessments. The two-stage cognitive assessment approach was used to maximize the accuracy of measurement and reduce administration time by using the child's responses from a brief first-stage routing form to select the appropriate second-stage level form.[13] For the reading, mathematics, and science routing forms, children read items in a booklet and recorded their responses on an answer form. These answer forms were then scored by the test administrator. Based on the score of the respective routing forms, the test administrator then assigned a high or low second-stage level form of the reading, mathematics, and science assessments. For the second-stage level tests, children read items in the assessment booklet and recorded their responses in the same assessment booklet. The routing tests and the second-stage level tests were timed and took 80 minutes to complete.

[13] For additional detail on the eighth-grade cognitive assessments, see the *ECLS-K Psychometric Report for the Eighth Grade* (NCES 2008–069) (Najarian, Pollack, and Sorongon forthcoming).

Accommodations that did not significantly affect the assessment were provided to those children whose Individualized Education Programs (IEPs) required accommodations for assessments. These included allowing for additional time or the presence of a health care aide. Children were excluded from the direct assessment if they had a disability (e.g., blindness or deafness) that could not be accommodated by the ECLS-K direct assessment, if their IEP prevented their participation in assessments, or they required an accommodation not offered by the ECLS-K assessments. Chapter 5, section 5.5.2 has more information on accommodations and exclusions in the ECLS-K.

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. Assessment items in each of the content domains were drawn from assessments used in other large-scale studies of similar-aged youth, such as the National Assessment of Educational Progress (NAEP), the National Education Longitudinal Study of 1988 (NELS:88), and the Education Longitudinal Study of 2002 (ELS:2002), the Texas Assessment of Knowledge and Skills (TAKS), as well as previous rounds of the ECLS-K. Items were chosen to extend the longitudinal scales initiated in kindergarten, first grade, third grade, and fifth grade, but were grade-appropriate in terms of content and format. Items were reviewed by content area specialists for appropriateness of content and difficulty, and for relevance to the assessment framework. In addition, items were reviewed for issues related to sensitivity to minorities. Items that passed these content, construct, and sensitivity screenings were field tested in the spring of 2006. For additional detail on the selection of items for the eighth-grade cognitive assessments, see the *ECLS-K Psychometric Report for the Eighth Grade* (NCES 2009–002) (Najarian, Pollack, and Sorongon forthcoming).

Reading. The eighth-grade reading assessment focused on four aspects of reading comprehension skills including forming a general understanding of the text, developing a more complete understanding of what was read, making connections from the text with personal background knowledge, and critically evaluating, comparing and contrasting, and understanding the effect of literary devices or the author's intentions.

The kindergarten through eighth-grade proficiency levels included (1) Letter Knowledge—identifying upper- and lower-case letters of the alphabet by name; (2) Beginning Sounds—associating letters with sounds at the beginning of words; (3) Ending Sounds—associating letters with sounds at the end of words; (4) Sight Words—recognizing common "sight" words; (5) Words in Context—reading words in context; (6) Literal Inference—making inferences using cues that were directly stated with key

words in text; (7) Extrapolation—identifying clues used to make inferences; (8) Evaluation—demonstrating understanding of author's craft and making connections between a problem in the narrative and similar life problems; (9) Evaluating Nonfiction—comprehension of biographical and expository text; and (10) Evaluating Complex Syntax—evaluating complex syntax and understanding high-level vocabulary.

Mathematics. The eighth-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.

The kindergarten through eighth-grade mathematics proficiency levels include (1) Number and Shape—identifying some one-digit numerals, recognizing geometric shapes, and one-to-one counting up to 10 objects; (2) Relative Size—reading all one-digit numerals, counting beyond 10, recognizing a sequence of patterns, and using nonstandard units of length to compare the size of objects; (3) Ordinality and 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; (4) Addition and Subtraction—solving simple addition and subtraction problems; (5) Multiplication and Division—solving simple multiplication and division problems and recognizing more complex number patterns; (6) Place Value—demonstrating understanding of place value in integers to hundreds' place; (7) Rate and Measurement—using knowledge of measurement and rate to solve word problems; (8) Fractions—solving problems using fractions; and (9) Area and Volume—solving word problems involving area and volume. No new mathematics proficiency level was added at the eighth grade because it was not warranted. Previously defined proficiency levels were sufficiently "difficult" to allow for the demonstration of growth in the higher proficiency levels at eighth grade.

Science. In the eighth-grade assessment, equal emphasis was placed on life science, earth and space science, and physical science. Similar to the third- and fifth-grade science assessments, children needed to demonstrate understanding of the physical and natural world, draw inferences, and comprehend relationships. In addition, they needed to interpret scientific data, formulate hypotheses, and identify the best plan to investigate a given question. As with the third- and fifth-grade science assessments, no set of proficiency levels was developed. The subject matter content of the science assessment domain was too

diverse and the items insufficiently ranked or graded to permit the formation of a set of proficiency levels. Instead, a single score was calculated to represent each child's breadth and depth of understanding and knowledge of the world.

For additional detail on the development of the eighth-grade cognitive assessments, see the *ECLS-K Psychometric Report for the Eighth Grade* (NCES 2009–002) (Najarian, Pollack, and Sorongon. forthcoming).

2.1.2 Student Questionnaire

Children completed the student questionnaire after completing the routing test. The student questionnaire was timed, and children had 20 minutes to complete the questionnaire. They entered their responses to each item into the student questionnaire booklet. Topics covered by the student questionnaire included the following:

- school experiences—school safety, importance of grades, time spent on homework, peer relationships;
- activities—participation in school-sponsored and out-of-school activities;
- social-emotional development—how children thought and felt about themselves both academically and socially;
- weight and exercise—level of exercise per week, participation in physical education classes; and
- diet—what kinds of food they could buy at school and the food they had eaten in the past week.

The student questionnaire included two scales to measure their socioemotional development. The first was the self-description questionnaire (SDQ), which was used to determine how children thought and felt about themselves both academically and socially. Children rated their perceived competence and interest in reading and mathematics. They also reported on internalizing problem behaviors with which they might struggle. The Internalizing Problems scale included items on sadness, loneliness, and anxiety.

The SDQ consists of 16 statements. Children rated whether each item was "not at all true," "a little bit true," "mostly true," or "very true." Three subscales were produced from the SDQ items. The scale scores on all SDQ scales represent the mean rating of the items included in the scale.

- **The SDQ Perceived Interest/Competence–Reading** subscale includes four items on grades in English and the child's interest in and enjoyment of reading.
- **The SDQ Perceived Interest/Competence–Math** subscale includes four items on mathematics grades and the child's interest in and enjoyment of mathematics.
- **The SDQ Internalizing Behavior** subscale 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 two subscales of the ECLS-K SDQ were adapted with permission from the Self Description Questionnaire (SDQ) II (Marsh 1992b). The items in the internalizing problem behavior subscale were developed specifically for the ECLS-K and used in the third- and fifth-grade rounds.

The second set of scales consisted of the Self-Concept and Locus of Control scales adapted from the National Education Longitudinal Study of 1988 (NELS:88). The Self-Concept scale comes from the Rosenberg Self-Esteem Scale (RSE) (Rosenberg 1965). These scales asked children about their perceptions about themselves and the amount of control they had of their own lives. Items were drawn from the NELS:88 student questionnaire and asked children to indicate the degree to which they agreed with 13 statements about themselves. They chose from the following responses: "strongly agree," "agree," "disagree," or "strongly disagree" for each item.

As noted earlier, to measure children's food consumption, the student questionnaire included 19 items that asked them about the kinds of food they could buy at school and the food they had eaten in the past week. The first set of questions was about foods that are high in fat, sodium, and/or added sugars (e.g., candy, salty snacks, soft drinks). 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.

The eighth-grade food consumption items were the same as those used at the fifth-grade round. Items tapping food consumption were taken mainly from existing surveys, although some were developed for the ECLS-K. Two main sources for questions were two surveys by the Centers for Disease Control and Prevention (CDC)/Division of Adolescent and School Health Surveys: the Youth Risk Behavior Surveillance Survey (YRBSS) and the School Health Programs and Policies Survey (SHPPS).[14] 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.

2.1.3 Physical Components

Anthropometric data were collected in all seven rounds of the ECLS-K. The anthropometric data consisted of recording the children's height (in inches to the nearest quarter-inch) and weight (in pounds to the nearest half-pound) to measure their physical growth and development. The Shorr Board vertical stadiometer and a Seca digital scale were used to obtain standing height and weight measurements, which were recorded on a height and weight recording form. Height and weight were measured twice for each child and took approximately 2 minutes to complete. For additional detail on the procedures used to collect height and weight, see the *ECLS-K Eighth-Grade Methodology Report* (NCES 2009–003) (Tourangeau et al. forthcoming).

2.2 Parent Interview

The eighth-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 (2.2 percent) were conducted in person.

Data collection for the eighth-grade parent interview started in fall 2006. The parent interview lasted on average 46 minutes and contained approximately 300 questions concerning eighth-

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

grade school experiences, parent characteristics, and child health. Topics covered in the eighth-grade parent interview included the following:

- parent involvement in school activities;
- family structure—demographics, household roster, marital status;
- home environment and cognitive stimulation—frequency of literacy activities, computer use, television viewing, homework, family routines;
- child's schooling;
- critical family processes, such as marital satisfaction and religiosity;
- parent/child interaction—parent discipline;
- nonresident parent—contact with child, school involvement, and child support;
- primary language spoken in home;
- child's health and well-being—physical functioning, parent ratings of child's strengths and difficulties,[15] services for children with special needs, receipt of prescription for attention and/or hyperactivity disorders, family therapy;
- parent health and emotional well-being;
- parental educational expectations for the child;
- parent education;
- parent employment;
- welfare and other public assistance use;
- food security; and
- parent income and assets.

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 in the eighth-grade data collection (94 percent), the eighth-grade parent respondent was the same as the

[15] These parent interview items (CHQ.900) are from the Strengths and Difficulties Questionnaire (ages 11 –17) copyrighted by Dr. Robert Goodman, Ph. D., of the Psychiatric Institute of London, England. Agencies may use these questions without charge or permission providing the wording is not modified, all questions are retained, and copyright is acknowledged. More information can be found at http://www.sdqinfo.com/ or Appendix V in http://www.cdc.gov/nchs/data/nhis/srvydesc.pdf.

respondent from the previous round. The child's mother was the respondent in 88 percent of the cases and the child's father in 9 percent. Other adults completed the parent interview in 3 percent of the cases (typically grandparents of the sampled child).

2.3 General Education Teacher Questionnaires

During the spring-eighth grade data collection, one teacher-level background and three child-level subject matter (i.e., English, mathematics, and science) questionnaires were used to collect data from the sampled children's teachers. The self-administered teacher-level background questionnaire covered a variety of topics, including views on teaching and the school, teacher demographic information, teaching experiences, and education and certification information.

The English, mathematics, and science teacher questionnaires were each organized in the same manner. Each questionnaire was divided into three sections. The first section included questions that collected data on the child's social skills, class performance, and his or her skills in relevant areas. The English teacher questionnaire asked about the child's skills in written and oral expression. The mathematics teacher questionnaire asked about the child's skills in mathematics, such as problem solving and demonstrating mathematical reasoning. The science teacher questionnaire asked about the child's skills in science, such as designing an experiment to solve a scientific question and writing up and preparing a presentation of scientific data.

The second section included questions about characteristics of the children in the classroom. The third section included questions about the instructional practices in the classroom, such as specific instructional activities and curricular focus, and assigned books and textbooks. In this last section, the items specified activities and practices that were relevant to the subject domain (i.e., English, mathematics, or science).

Two subject-matter questionnaires were completed for each sampled child. Therefore, data were gathered on each sampled child's skills in the areas of English and mathematics, or in the areas of English and science.

Topics covered in the spring-eighth grade teacher questionnaires included the following:

- race/ethnicity of children in the classroom;
- materials and resources available, such as computers;
- instructional time on different topics;
- behavior of children in classroom;
- instructional information;
- teachers' evaluation and grading practices;
- perceptions of school climate;
- teacher demographic information;
- teacher experience and education;
- job satisfaction;
- children's domain-relevant skills (i.e., written and oral expression, science, and mathematics skills); and
- children's behavior and performance in class.

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 a mathematics or science teacher completed questionnaires. This latter approach was also used in spring-eighth grade, in which each sampled child's English teacher and either a 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.

Each child's selected teacher(s) received a self-administered teacher-level background questionnaire. In addition to the teacher-level questionnaire, each teacher received at least one of the three child-level questionnaires (English, mathematics, or science, based on the subject(s) they taught) specifically about the focal child. All children were assigned to have an English teacher complete questionnaires. In fifth grade, half of the children were randomly assigned to have a mathematics teacher complete questionnaires, and the other half of the children were assigned to have a science teacher complete questionnaires. This assignment made for the mathematics or science teacher questionnaire in

fifth grade was carried forward in eighth grade so that the same children who had a mathematics teacher questionnaire in fifth grade would have a mathematics teacher questionnaire in eighth grade, and those with a science teacher questionnaire in fifth grade would have a science teacher questionnaire in eighth grade. In cases where the same eighth-grade teacher taught the sampled child English, mathematics, and science, the teacher was asked to complete an English questionnaire and either a mathematics or science questionnaire, depending upon the domain for which the child was sampled.

2.4 Special Education Teacher Questionnaires

In the spring-eighth-grade data collection, field supervisors asked the school coordinators to identify the ECLS-K children receiving special education services and the names of their special education service providers. The supervisor then listed special education staff working with each child (e.g., speech pathologists, reading instructors, and audiologists). Field supervisors determined the primary service provider of children receiving special education services from multiple service providers. The primary special education teacher/service provider was defined as follows:

- the teacher who managed the child's IEP;
- the teacher who spent the greatest 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.

Special education teachers of children in the ECLS-K were asked to complete two questionnaires. The questionnaires addressed topics such as the child's disability, IEP goals, the amount and type of services used by sampled children, and communication with parents and general education teachers. Part A of the special education teacher questionnaire was designed to collect information about the special education teacher's professional background and experience, including the following:

- teacher's sex;
- teacher's race/ethnicity;
- teaching experience;
- educational background;

- special education teacher background;
- location of service provision;
- student load per week; and
- teacher's main assignment.

The special education teacher was asked to complete part B. Part B asked about the special education services provided to the child and the nature of the child's special education curriculum. Items covered such topics as the following:

- disability category;
- IEP goals for the school year;
- extent of services;
- types of services provided for the year;
- primary placement;
- teaching practices, methods, and materials;
- assistive technologies used by the child;
- general education goals, expectations, and participation in school-wide assessments;
- collaboration/communication with the child's general education teacher;
- frequency of communicating with the child's parents;
- child receipt of formal evaluations in the past year;
- when the child first had the IEP;
- likelihood that the child would have an IEP next school year;
- percentage of IEP goals that were met during this school year; and
- receipt of special education or related services because of attention deficit/ hyperactivity disorder.

The special education teacher was asked to complete part B for each sampled child for whom he or she was the primary service provider.

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 2007. This self-administered questionnaire was intended to gather information about the school, student body, teachers, school policies, and characteristics of the administrator. 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.

The school administrator was also asked questions regarding 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 children 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 children a la carte lunch or breakfast items 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 machine); and how full the cafeteria was at peak meal times. Questions on the availability of foods that were not part of USDA meal programs and on cafeteria crowding were taken from the School Health Policies and Programs Study (SHPPS).

The content areas addressed in this questionnaire in spring-eighth grade included the following:

- school characteristics—type of school, length of school year and start and end dates, school size, average daily attendance, highest and lowest grades;
- academic course offerings for eighth-graders;
- child population characteristics—race/ethnicity of children, participation in special services, percent Limited English Proficient (LEP);
- school facilities and resources;

- community characteristics and school safety;
- average starting salary of full-time first year teachers;
- school policies and programs—assessments and testing, free and reduced-price breakfast and lunch;
- programs for special populations—English as a Second Language (ESL) and bilingual education, special education, gifted and talented;
- principal characteristics—sex, race/ethnicity, age of principal, experience and education;
- school governance and climate—goals and objectives for teachers, school functioning and decisionmaking; and
- availability of different types of foods during school hours.

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 eighth grade. These scores were for the direct cognitive assessment, the teacher ratings of English, mathematics, and science skills, 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.[16] Guidelines for when and how to use each cognitive assessment score are also provided in this chapter.

3.1 Direct Cognitive Assessment

The eighth-grade direct cognitive assessment contained items in reading, mathematics, and science. In each subject area, children received a 10-item routing test. Performance on the routing items guided the selection and administration of one of two second-stage (high and low) forms in each subject area. The second-stage forms contained items of appropriate difficulty for the level of ability indicated by the routing items.[17]

The eighth-grade direct cognitive assessment was built from the framework established in the previous kindergarten through fifth-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 study. 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 administering 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.

[16] This manual is applicable to the data gathered during the 2006–07 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), 2001–02 school year (third grade), and 2003–04 (fifth grade) is provided primarily for background and comparison purposes.

[17] See chapter 2, section 2.1 of the *ECLS-K Psychometric Report for the Eighth Grade* (NCES 2008–069) (Najarian et al. forthcoming) for additional information on the two-stage process for the direct cognitive assessments.

- **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, fifth grade, and again for the eighth 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 kindergarten and first grade (K-1) general knowledge assessment included basic natural science concepts as well as concepts in social studies. For third, fifth, and eighth grades, a science assessment replaced the general knowledge assessment. Thus, the longitudinal scale for measuring gains in science spans only the third- through eighth-grade rounds.

- **Assessment format:** The format of the eighth-grade assessment was modified from that of prior rounds to accommodate administration differences for the older sample. In all previous rounds, an assessor presented the questions to the child and entered responses into a computer for each individually administered assessment. In third grade and fifth grade, the 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. The individually administered easel assessments were less appropriate for the older sample in eighth grade. Therefore, the eighth-grade assessments were paper-based and were administered in groups (where possible). The passages and items were in booklet form; an answer sheet was provided for the routing test responses, while responses for the second-stage forms were entered into the assessment booklets. There were two second-stage forms for each domain.

- **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, were reported for the third- and fifth-grade assessments. There were no cluster scores for the eighth-grade round.

- **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. Similarly, a bridge sample was not done to bridge the gap between fifth- and eighth-graders.

The scores used to describe a child's performance on the direct cognitive assessment include broad-based measures that report 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 are simple counts of correct answers, while others are 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 are 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 measure children's performance on a set of questions with a broad range of difficulty. Standardized scores (T-scores) report children's performance relative to their peers. Criterion-referenced proficiency scores evaluate children's performance with respect to subsets of items that mark specific skills.

Tables 3-1 through 3-9 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: 1 (fall-kindergarten), 2 (spring-kindergarten), 3 (fall-first grade), 4 (spring-first grade), 5 (spring-third grade), 6 (spring-fifth grade), or 7 (spring-eighth grade). Weighted means in tables containing only eighth-grade scores use the round 7 cross-sectional weight, C7CW0, to represent population estimates for eighth grade. Weighted estimates in tables containing scores for all earlier rounds are based on C1_7SC0, the round 1-2-3-4-5-6-7 panel weight, while tables containing only scores for science, assessed only in third, fifth, and eighth grades, use C57CW0, the round 5-7 panel weight. Kindergarten through fifth-grade scores in this database differ somewhat from the corresponding scores in the previously released data files because they were re-estimated along with the eighth-grade scores (see section 3.1.2). In addition, all kindergarten through fifth-grade score statistics presented here differed from previous estimates because the panel weight used restricted estimates to children who participated in all seven rounds of data collection (for reading and mathematics scores), or rounds 5, 6, and 7 (science scores).

3.1.1 Number-Right Scores

Routing test number-right scores are counts of the raw number of items a child answered correctly on the routing test. Number-right scores are not comparable to one another when the assessment differs in difficulty (i.e., high vs. low second-stage form). For example, a child who took the high-difficulty mathematics second-stage form would probably have answered more questions correctly if the

easier low form had been administered. For this reason, raw number-right scores are reported in the database only for the first-stage (routing) tests, which are 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 K-1 reading routing test emphasized pre-reading skills, while the routing tests in fifth and eighth grade contained questions based on reading passages with progressively more difficult content. 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, fifth, and eighth 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*, the fifth-grade scores only *within fifth grade*, and the eighth-grade scores only *within eighth grade*, 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, table 3-3 for the fifth-grade routing tests, and table 3-4 for the eighth-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
C1R4RNOR	C1 RC4 Reading Routing #Right - K-1 Assmt	0–20	5.96	3.91
C2R4RNOR	C2 RC4 Reading Routing #Right - K-1 Assmt	0–20	10.02	4.05
C3R4RNOR	C3 RC4 Reading Routing #Right - K-1 Assmt	0–20	11.83	4.14
C4R4RNOR	C4 RC4 Reading Routing #Right - K-1 Assmt	0–20	16.41	3.70
C1R4MNOR	C1 RC4 Math Routing #Right - K-1 Assmt	0–16	4.65	3.01
C2R4MNOR	C2 RC4 Math Routing #Right - K-1 Assmt	0–16	7.30	3.36
C3R4MNOR	C3 RC4 Math Routing #Right - K-1 Assmt	0–16	9.01	3.30
C4R4MNOR	C4 RC4 Math Routing #Right - K-1 Assmt	0–16	11.87	2.84

NOTE: Table estimates based on C1_7SC0 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 7.5 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
C5R4RNR3	C5 RC4 Reading Routing #Right - Gr3 Assmt	0–15	10.05	2.82
C5R4MNR3	C5 RC4 Math Routing #Right - Gr3 Assmt	0–17	9.13	4.34
C5SROUNR	C5 Science Routing #Right - Gr3 Assmt	0–15	8.41	3.39

NOTE: Table estimates based on C1_7SC0 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, Rock et al. 2005) because of sample attrition. See chapter 7, section 7.5 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
C6R4RNR5	C6 RC4 Reading Routing #Right - Gr5 Assmt	0–25	11.59	5.10
C6R4MNR5	C6 RC4 Math Routing #Right - Gr5 Assmt	0–18	10.02	4.80
C6R1SNR5	C6 RC1 Science Routing #Right - Gr5 Assmt	0–21	13.43	4.09

NOTE Table estimates based on C1_7SC0 panel weight. Table estimates may differ from those reported in earlier user's manuals and the *ECLS-K Psychometric Report for the Fifth Grade* (NCES 2006–036rev (Pollack, Atkins-Burnett et al. 2005) because of sample attrition. See chapter 7, section 7.5 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 3-4. Direct cognitive assessment: routing test number-right, eighth-grade assessment: School year 2006–07

Variable	Description	Range of values	Weighted mean	Standard deviation
C7R4RNR8	C7 RC4 Reading Routing #Right - Gr8 Assmt	0–10	5.95	2.53
C7R4MNR8	C7 RC4 Math Routing #Right - Gr8 Assmt	0–10	5.88	2.61
C7R2SNR8	C7 RC2 Science Routing #Right - Gr8 Assmt	0–10	6.46	2.46

NOTE: Table estimates based on C1_7SC0 panel weight. Table estimates may differ from those reported in the *ECLS-K Psychometric Report for the Eighth Grade* (NCES 2009–002) (Najarian, Pollack, and Sorongon, forthcoming) because of sample attrition. See chapter 7, section 7.5 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 2007.

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 child guessing several difficult 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-, fifth-, and eighth-grade data to stabilize the longitudinal estimates. In addition, the maximum values of the scale scores were extended to include the more difficult items administered in the eighth-grade assessments. The scale scores for each round of user files are defined on the basis of performance on all tasks administered *up to and including* the current round. The re-estimated kindergarten/first-grade, third-grade, fifth-grade, and eighth-grade IRT scores in this database differ from the IRT scores in the kindergarten/first-grade, third-grade, and fifth-grade files previously released. For example, the reading scale score in the fifth-grade file is based on test items used in kindergarten through fifth grade, while the current reading score is an estimate based on an expanded set of items, all of those used in kindergarten through eighth 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 (or recalibrated) so that comparisons can be made.

The IRT scale scores in the database represent estimates of the number of items children would have answered correctly at each point in time if they had taken all of the 212 questions in all of the first- and second-stage reading forms administered in all rounds, the 174 questions in all of the mathematics forms, and the 111 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, spring-fifth grade from spring-eighth 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 to eighth 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 3-5 for variable names, descriptions, ranges, weighted means, and standard deviations for the IRT scale scores.

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 provide only 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.

Table 3-5. Direct cognitive assessment: Item Response Theory (IRT) scale scores: School year 2006–07

Variable	Description	Range of values	Weighted mean	Standard deviation
C1R4RSCL	C1 RC4 Reading IRT Scale Score	0–212	35.47	9.86
C2R4RSCL	C2 RC4 Reading IRT Scale Score	0–212	46.52	13.88
C3R4RSCL	C3 RC4 Reading IRT Scale Score	0–212	52.73	16.93
C4R4RSCL	C4 RC4 Reading IRT Scale Score	0–212	77.07	23.70
C5R4RSCL	C5 RC4 Reading IRT Scale Score	0–212	125.70	28.57
C6R4RSCL	C6 RC4 Reading IRT Scale Score	0–212	148.67	26.85
C7R4RSCL	C7 RC4 Reading IRT Scale Score	0–212	167.24	28.03
C1R4MSCL	C1 RC4 Math IRT Scale Score	0–174	26.13	9.09
C2R4MSCL	C2 RC4 Math IRT Scale Score	0–174	36.17	12.00
C3R4MSCL	C3 RC4 Math IRT Scale Score	0–174	43.57	14.22
C4R4MSCL	C4 RC4 Math IRT Scale Score	0–174	61.50	17.66
C5R4MSCL	C5 RC4 Math IRT Scale Score	0–174	98.77	24.96
C6R4MSCL	C6 RC4 Math IRT Scale Score	0–174	122.94	25.18
C7R4MSCL	C7 RC4 Math IRT Scale Score	0–174	139.28	23.10
C5SR2SSCL	C5 RC4 Science IRT Scale Score	0–111	49.91	15.29
C6SR2SSCL	C6 RC4 Science IRT Scale Score	0–111	63.87	15.73
C7SR2SSCL	C7 RC2 Science IRT Scale Score	0–111	82.72	17.07

NOTE: Table estimates based on C1_7SC0 panel weight. Table estimates may differ from those reported in earlier user's manuals and psychometric reports because of re-estimation of scores on a longitudinal scale that includes eighth grade, and because of sample attrition. See chapter 7, section 7.5 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 2007.

The standardized 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 fifth-grade reading T-score of 45 (C6R4RTSC) represents a reading achievement level that is one-half of a standard deviation lower than the mean for the population represented by the assessed sample of ECLS-K round 6 participants. If the same child had a reading T-score of 50 in eighth grade (C7R4RTSC), this would indicate that the child had made up his or her initial deficit and was 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 3-6 for variable names, descriptions, and ranges for the standardized T-scores across all rounds.

Table 3-6. Direct cognitive assessment: standardized scores: School year 2006–07

Variable	Description	Range of values	Weighted mean	Standard deviation
C1R4RTSC	C1 RC4 Reading T-Score	0–96	50.82	9.90
C2R4RTSC	C2 RC4 Reading T-Score	0–96	50.76	9.79
C3R4RTSC	C3 RC4 Reading T-Score	0–96	50.52	9.62
C4R4RTSC	C4 RC4 Reading T-Score	0–96	50.55	9.77
C5R4RTSC	C5 RC4 Reading T-Score	0–96	50.28	10.13
C6R4RTSC	C6 RC4 Reading T-Score	0–96	50.56	9.76
C7R4RTSC	C7 RC4 Reading T-Score	0–96	50.13	9.68
C1R4MTSC	C1 RC4 Math T-Score	0–96	50.65	10.10
C2R4MTSC	C2 RC4 Math T-Score	0–96	50.50	9.95
C3R4MTSC	C3 RC4 Math T-Score	0–96	50.73	9.59
C4R4MTSC	C4 RC4 Math T-Score	0–96	50.83	9.09
C5R4MTSC	C5 RC4 Math T-Score	0–96	50.68	9.95
C6R4MTSC	C6 RC4 Math T-Score	0–96	50.92	9.79
C7R4MTSC	C7 RC4 Math T-Score	0–96	50.25	9.93
C5R2STSC	C5 RC2 Science T-Score	0–96	50.37	10.10
C6R2STSC	C6 RC2 Science T-Score	0–96	50.61	9.63
C7R2STSC	C7 RC2 Science T-Score	0–96	50.23	9.83

NOTE: Table estimates based on C1_7SC0 panel weight. Table estimates may differ from those reported in earlier user's manuals and psychometric reports because of re-estimation of scores on a longitudinal scale that includes eighth grade, and because of sample attrition. See chapter 7, section 7.5 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 2007.

The K-8 Full Sample Public Use data file includes the IRT theta (ability) scores for each data collection round for each domain (reading, mathematics, general knowledge, science) along with the standard error of measurement (SEM) associated with each theta score. The theta scores represent a child's ability measured at each round along a single continuous scale. The theta scores represent underlying ability (which is normally distributed at all rounds) while the IRT scale scores represent predicted performance on the ECLS-K assessments (which is not normally distributed at all rounds). The theta scores are ideally suited for measuring growth from kindergarten through the eighth grade. The theta score distribution range is approximately -3 to 3.

3.1.3 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 10 points along the reading and 9 points along the math score scales for the 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 child 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 eighth grade. No proficiency scores were computed for the science assessment because the questions did not follow a hierarchical pattern.

3.1.3.1 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;
- **Level 9: Evaluating nonfiction:** critically evaluating, comparing and contrasting, and understanding the effect of features of expository and biographical texts; and
- **Level 10: Evaluating complex syntax:** evaluating complex syntax and understanding high-level nuanced vocabulary in biographical text.

3.1.3.2 Mathematics

- **Level 1: Number and shape:** identifying some one-digit numerals, recognizing geometric shapes, and one-to-one counting of up to 10 objects;
- **Level 2: Relative size:** reading all single-digit numerals, counting beyond 10, 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 child 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 children in kindergarten through eighth 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. For the first 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; in round 7 (grade 8) these figures were 3 percent for mathematics and less than 1 percent for reading. 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.

3.1.3.3 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 assessment was a two-stage design (where not all children were administered all items), and since more advanced assessment instruments were administered in third, fifth, and eighth grades, the 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 children'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 and used to assign an imputed right or wrong answer. 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 for the kindergarten through fifth-grade rounds. In eighth grade, the scores determined on the basis of item response data dropped to 19 percent for reading and 47 percent for math, a result of the necessary placement of the proficiency level items on either the low or high second-stage form, based on their estimated difficulty levels. The rest utilized IRT-based probabilities for some or all of the missing items, since the "missingness" is a consequence of the child's ability or grade level and requires special treatment in order to avoid misleading results. (The *ECLS-K Psychometric Report for the Eighth Grade* (NCES 2009–002) (Najarian, Pollack, and Sorongon forthcoming) describes this treatment in more detail.) 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 nonmastery of all higher levels. A zero score indicates nonmastery 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 3-7 for variable names, descriptions, and weighted percentages for the highest proficiency level mastered scores.

Table 3-7. Direct cognitive assessment: highest proficiency level mastered, in percent: School year 2006–07

Variable	Description	Below Level 1	Level 1	Level 2	Level 3	Level 4	Level 5	Level 6	Level 7	Level 8	Level 9	Level 10
C1R4RPF	C1 RC4 Reading Highest Prof Lvl Mastered	30	36	17	15	2	1	0	0	0	0	0
C2R4RPF	C2 RC4 Reading Highest Prof Lvl Mastered	6	17	23	40	10	3	1	0	0	0	0
C3R4RPF	C3 RC4 Reading Highest Prof Lvl Mastered	5	13	18	40	14	8	2	1	0	0	0
C4R4RPF	C4 RC4 Reading Highest Prof Lvl Mastered	0	2	4	12	33	33	12	3	1	0	0
C5R4RPF	C5 RC4 Reading Highest Prof Lvl Mastered	0	0	0	1	4	19	24	27	23	2	0
C6R4RPF	C6 RC4 Reading Highest Prof Lvl Mastered	0	0	0	0	1	7	16	34	34	7	1
C7R4RPF	C7 RC4 Reading Highest Prof Lvl Mastered	0	0	0	0	0	4	8	19	34	27	7
C1R4MPF	C1 RC4 Math Highest Prof Lvl Mastered	6	32	37	20	4	0	0	0	0	0	†
C2R4MPF	C2 RC4 Math Highest Prof Lvl Mastered	1	11	27	41	17	2	0	0	0	0	†
C3R4MPF	C3 RC4 Math Highest Prof Lvl Mastered	1	9	18	42	24	5	1	0	0	0	†
C4R4MPF	C4 RC4 Math Highest Prof Lvl Mastered	0	1	5	20	48	22	5	0	0	0	†
C5R4MPF	C5 RC4 Math Highest Prof Lvl Mastered	0	0	0	4	17	30	31	16	2	0	†
C6R4MPF	C6 RC4 Math Highest Prof Lvl Mastered	0	0	0	1	5	15	33	30	14	2	†
C7R4MPF	C7 RC4 Math Highest Prof Lvl Mastered	0	0	0	1	1	7	22	31	22	17	†

†Not applicable.
NOTE: Table estimates based on C1_7SC0 panel weight. Table estimates may differ from those reported in earlier user's manuals and psychometric reports because of re-estimation of scores on a longitudinal scale that includes eighth grade, and because of sample attrition. See chapter 7, section 7.5 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 2007.

3.1.3.4 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 his or her 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 children's 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 mathematics level 8 probability at round 6 (C6R4MPB8) from the level 8 probability at round 7 (C7R4MPB8) indicates whether a child advanced in mastery of the particular set of level 8 skills (i.e., fractions) during the time interval between the fifth- and eighth-grade assessments. Thus, children's school experiences can be related to improvements in specific skills.

See tables 3-8 and 3-9 for variable names, descriptions, ranges, weighted means, and standard deviations for the proficiency probability scores in reading and mathematics.

Table 3-8. Eighth-grade direct cognitive assessment: proficiency probability scores—reading: School year 2006–07

Variable	Description	Range of values	Weighted mean	Standard deviation
C1R4RPB1	C1 RC4 Prob1 - Letter Recognition	0–1	0.68	0.33
C1R4RPB2	C1 RC4 Prob2 - Beginning Sounds	0–1	0.31	0.34
C1R4RPB3	C1 RC4 Prob3 - Ending Sounds	0–1	0.18	0.26
C1R4RPB4	C1 RC4 Prob4 - Sight Words	0–1	0.03	0.14
C1R4RPB5	C1 RC4 Prob5 - Word in Context	0–1	0.02	0.09
C1R4RPB6	C1 RC4 Prob6 - Literal Inference	0–1	0.00	0.03
C1R4RPB7	C1 RC4 Prob7 - Extrapolation	0–1	0.00	0.01
C1R4RPB8	C1 RC4 Prob8 - Evaluation	0–1	0.00	0.01
C1R4RPB9	C1 RC4 Prob9 - Evaluating Nonfiction	0–1	0.00	0.00
C1R4RPB10	C1 RC4 Prob10 - Evaluating Complex Syntax	0–1	0.00	0.00
C2R4RPB1	C2 RC4 Prob1 - Letter Recognition	0–1	0.93	0.17
C2R4RPB2	C2 RC4 Prob2 - Beginning Sounds	0–1	0.70	0.32
C2R4RPB3	C2 RC4 Prob3 - Ending Sounds	0–1	0.51	0.34
C2R4RPB4	C2 RC4 Prob4 - Sight Words	0–1	0.16	0.26
C2R4RPB5	C2 RC4 Prob5 - Word in Context	0–1	0.07	0.16
C2R4RPB6	C2 RC4 Prob6 - Literal Inference	0–1	0.02	0.08
C2R4RPB7	C2 RC4 Prob7 - Extrapolation	0–1	0.00	0.04
C2R4RPB8	C2 RC4 Prob8 - Evaluation	0–1	0.00	0.02
C2R4RPB9	C2 RC4 Prob9 - Evaluating Nonfiction	0–1	0.00	0.00
C2R4RPB10	C7 RC4 Prob10 - Evaluating Complex Syntax	0–1	0.00	0.00
C3R4RPB1	C3 RC4 Prob1 - Letter Recognition	0–1	0.96	0.13
C3R4RPB2	C3 RC4 Prob2 - Beginning Sounds	0–1	0.82	0.27
C3R4RPB3	C3 RC4 Prob3 - Ending Sounds	0–1	0.67	0.32
C3R4RPB4	C3 RC4 Prob4 - Sight Words	0–1	0.28	0.32
C3R4RPB5	C3 RC4 Prob5 - Word in Context	0–1	0.13	0.22
C3R4RPB6	C3 RC4 Prob6 - Literal Inference	0–1	0.03	0.12
C3R4RPB7	C3 RC4 Prob7 - Extrapolation	0–1	0.01	0.06
C3R4RPB8	C3 RC4 Prob8 - Evaluation	0–1	0.01	0.03
C3R4RPB9	C3 RC4 Prob9 - Evaluating Nonfiction	0–1	0.00	0.00
C3R4RPB10	C3 RC4 Prob10 - Evaluating Complex Syntax	0–1	0.00	0.00
C4R4RPB1	C4 RC4 Prob1 - Letter Recognition	0–1	0.99	0.05
C4R4RPB2	C4 RC4 Prob2 - Beginning Sounds	0–1	0.97	0.12
C4R4RPB3	C4 RC4 Prob3 - Ending Sounds	0–1	0.92	0.19
C4R4RPB4	C4 RC4 Prob4 - Sight Words	0–1	0.75	0.32
C4R4RPB5	C4 RC4 Prob5 - Word in Context	0–1	0.48	0.32
C4R4RPB6	C4 RC4 Prob6 - Literal Inference	0–1	0.18	0.23
C4R4RPB7	C4 RC4 Prob7 - Extrapolation	0–1	0.06	0.13
C4R4RPB8	C4 RC4 Prob8 - Evaluation	0–1	0.03	0.06
C4R4RPB9	C4 RC4 Prob9 - Evaluating Nonfiction	0–1	0.00	0.00
C4R4RPB10	C4 RC4 Prob10 - Evaluating Complex Syntax	0–1	0.00	0.00

See notes at end of table.

Table 3-8. Eighth-grade direct cognitive assessment: proficiency probability scores—reading: School year 2006–07—Continued

Variable	Description	Range of values	Weighted mean	Standard deviation
C5R4RPB1	C5 RC4 Prob1 - Letter Recognition	0–1	1.00	0.00
C5R4RPB2	C5 RC4 Prob2 - Beginning Sounds	0–1	1.00	0.00
C5R4RPB3	C5 RC4 Prob3 - Ending Sounds	0–1	1.00	0.01
C5R4RPB4	C5 RC4 Prob4 - Sight Words	0–1	0.98	0.09
C5R4RPB5	C5 RC4 Prob5 - Word in Context	0–1	0.90	0.17
C5R4RPB6	C5 RC4 Prob6 - Literal Inference	0–1	0.68	0.28
C5R4RPB7	C5 RC4 Prob7 - Extrapolation	0–1	0.44	0.31
C5R4RPB8	C5 RC4 Prob8 - Evaluation	0–1	0.25	0.21
C5R4RPB9	C5 RC4 Prob9 - Evaluating Nonfiction	0–1	0.01	0.03
C5R4RPB10	C5 RC4 Prob10 - Evaluating Complex Syntax	0–1	0.00	0.00
C6R4RPB1	C6 RC4 Prob1 - Letter Recognition	0–1	1.00	0.00
C6R4RPB2	C6 RC4 Prob2 - Beginning Sounds	0–1	1.00	0.00
C6R4RPB3	C6 RC4 Prob3 - Ending Sounds	0–1	1.00	0.00
C6R4RPB4	C6 RC4 Prob4 - Sight Words	0–1	1.00	0.01
C6R4RPB5	C6 RC4 Prob5 - Word in Context	0–1	0.97	0.07
C6R4RPB6	C6 RC4 Prob6 - Literal Inference	0–1	0.85	0.19
C6R4RPB7	C6 RC4 Prob7 - Extrapolation	0–1	0.67	0.29
C6R4RPB8	C6 RC4 Prob8 - Evaluation	0–1	0.44	0.27
C6R4RPB9	C6 RC4 Prob9 - Evaluating Nonfiction	0–1	0.06	0.15
C6R4RPB10	C6 RC4 Prob10 - Evaluating Complex Syntax	0–1	0.01	0.02
C7R4RPB1	C7 RC4 Prob1 - Letter Recognition	0–1	1.00	0.00
C7R4RPB2	C7 RC4 Prob2 - Beginning Sounds	0–1	1.00	0.00
C7R4RPB3	C7 RC4 Prob3 - Ending Sounds	0–1	1.00	0.00
C7R4RPB4	C7 RC4 Prob4 - Sight Words	0–1	1.00	0.00
C7R4RPB5	C7 RC4 Prob5 - Word in Context	0–1	0.98	0.03
C7R4RPB6	C7 RC4 Prob6 - Literal Inference	0–1	0.92	0.14
C7R4RPB7	C7 RC4 Prob7 - Extrapolation	0–1	0.82	0.25
C7R4RPB8	C7 RC4 Prob8 - Evaluation	0–1	0.64	0.30
C7R4RPB9	C7 RC4 Prob9 - Evaluating Nonfiction	0–1	0.26	0.34
C7R4RPB10	C7 RC4 Prob10 - Evaluating Complex Syntax	0–1	0.06	0.13

NOTE: Table estimates based on C1_7SC0 panel weight. Table estimates may differ from those reported in earlier user's manuals and psychometric reports because of re-estimation of scores on a longitudinal scale that includes eighth grade, and because of sample attrition. See chapter 7, section 7.5 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 2007.

Table 3-9. Eighth-grade direct cognitive assessment: proficiency probability scores—mathematics: School year 2006–07

Variable	Description	Range of values	Weighted mean	Standard deviation
C1R4MPB1	C1 RC4 Prob1 - Count, Number, Shape	0–1	0.92	0.17
C1R4MPB2	C1 RC4 Prob2 - Relative Size	0–1	0.55	0.36
C1R4MPB3	C1 RC4 Prob3 - Ordinality, Sequence	0–1	0.22	0.30
C1R4MPB4	C1 RC4 Prob4 - Add/Subtract	0–1	0.04	0.12
C1R4MPB5	C1 RC4 Prob5 - Multiply/Divide	0–1	0.00	0.03
C1R4MPB6	C1 RC4 Prob6 - Place Value	0–1	0.00	0.00
C1R4MPB7	C1 RC4 Prob7 - Rate & Measurement	0–1	0.00	0.00
C1R4MPB8	C1 RC4 Prob8 - Fractions	0–1	0.00	0.00
C1R4MPB9	C1 RC4 Prob9 - Area and Volume	0–1	0.00	0.00
C2R4MPB1	C2 RC4 Prob1 - Count, Number, Shape	0–1	0.99	0.05
C2R4MPB2	C2 RC4 Prob2 - Relative Size	0–1	0.84	0.24
C2R4MPB3	C2 RC4 Prob3 - Ordinality, Sequence	0–1	0.54	0.38
C2R4MPB4	C2 RC4 Prob4 - Add/Subtract	0–1	0.17	0.26
C2R4MPB5	C2 RC4 Prob5 - Multiply/Divide	0–1	0.02	0.08
C2R4MPB6	C2 RC4 Prob6 - Place Value	0–1	0.00	0.01
C2R4MPB7	C2 RC4 Prob7 - Rate & Measurement	0–1	0.00	0.00
C2R4MPB8	C2 RC4 Prob8 - Fractions	0–1	0.00	0.00
C2R4MPB9	C2 RC4 Prob9 - Area and Volume	0–1	0.00	0.00
C3R4MPB1	C3 RC4 Prob1 - Count, Number, Shape	0–1	1.00	0.02
C3R4MPB2	C3 RC4 Prob2 - Relative Size	0–1	0.92	0.17
C3R4MPB3	C3 RC4 Prob3 - Ordinality, Sequence	0–1	0.73	0.33
C3R4MPB4	C3 RC4 Prob4 - Add/Subtract	0–1	0.33	0.33
C3R4MPB5	C3 RC4 Prob5 - Multiply/Divide	0–1	0.05	0.14
C3R4MPB6	C3 RC4 Prob6 - Place Value	0–1	0.00	0.03
C3R4MPB7	C3 RC4 Prob7 - Rate & Measurement	0–1	0.00	0.00
C3R4MPB8	C3 RC4 Prob8 - Fractions	0–1	0.00	0.00
C3R4MPB9	C3 RC4 Prob9 - Area and Volume	0–1	0.00	0.00
C4R4MPB1	C4 RC4 Prob1 - Count, Number, Shape	0–1	1.00	0.00
C4R4MPB2	C4 RC4 Prob2 - Relative Size	0–1	0.99	0.04
C4R4MPB3	C4 RC4 Prob3 - Ordinality, Sequence	0–1	0.95	0.16
C4R4MPB4	C4 RC4 Prob4 - Add/Subtract	0–1	0.71	0.31
C4R4MPB5	C4 RC4 Prob5 - Multiply/Divide	0–1	0.23	0.30
C4R4MPB6	C4 RC4 Prob6 - Place Value	0–1	0.03	0.11
C4R4MPB7	C4 RC4 Prob7 - Rate & Measurement	0–1	0.00	0.02
C4R4MPB8	C4 RC4 Prob8 - Fractions	0–1	0.00	0.00
C4R4MPB9	C4 RC4 Prob9 - Area and Volume	0–1	0.00	0.00
C5R4MPB1	C5 RC4 Prob1 - Count, Number, Shape	0–1	1.00	0.00
C5R4MPB2	C5 RC4 Prob2 - Relative Size	0–1	1.00	0.00
C5R4MPB3	C5 RC4 Prob3 - Ordinality, Sequence	0–1	1.00	0.02
C5R4MPB4	C5 RC4 Prob4 - Add/Subtract	0–1	0.97	0.10
C5R4MPB5	C5 RC4 Prob5 - Multiply/Divide	0–1	0.77	0.32

See notes at end of table.

Table 3-9. Eighth-grade direct cognitive assessment: proficiency probability scores—mathematics: School year 2006–07—Continued

Variable	Description	Range of values	Weighted mean	Standard deviation
C5R4MPB6	C5 RC4 Prob6 - Place Value	0–1	0.43	0.40
C5R4MPB7	C5 RC4 Prob7 - Rate & Measurement	0–1	0.14	0.24
C5R4MPB8	C5 RC4 Prob8 - Fractions	0–1	0.01	0.06
C5R4MPB9	C5 RC4 Prob9 - Area and Volume	0–1	0.00	0.01
C6R4MPB1	C6 RC4 Prob1 - Count, Number, Shape	0–1	1.00	0.00
C6R4MPB2	C6 RC4 Prob2 - Relative Size	0–1	1.00	0.00
C6R4MPB3	C6 RC4 Prob3 - Ordinality, Sequence	0–1	1.00	0.00
C6R4MPB4	C6 RC4 Prob4 - Add/Subtract	0–1	1.00	0.02
C6R4MPB5	C6 RC4 Prob5 - Multiply/Divide	0–1	0.93	0.18
C6R4MPB6	C6 RC4 Prob6 - Place Value	0–1	0.75	0.35
C6R4MPB7	C6 RC4 Prob7 - Rate & Measurement	0–1	0.43	0.38
C6R4MPB8	C6 RC4 Prob8 - Fractions	0–1	0.14	0.27
C6R4MPB9	C6 RC4 Prob9 - Area and Volume	0–1	0.03	0.10
C7R4MPB1	C7 RC4 Prob1 - Count, Number, Shape	0–1	1.00	0.00
C7R4MPB2	C7 RC4 Prob2 - Relative Size	0–1	1.00	0.00
C7R4MPB3	C7 RC4 Prob3 - Ordinality, Sequence	0–1	1.00	0.00
C7R4MPB4	C7 RC4 Prob4 - Add/Subtract	0–1	1.00	0.00
C7R4MPB5	C7 RC4 Prob5 - Multiply/Divide	0–1	0.98	0.07
C7R4MPB6	C7 RC4 Prob6 - Place Value	0–1	0.89	0.25
C7R4MPB7	C7 RC4 Prob7 - Rate & Measurement	0–1	0.67	0.37
C7R4MPB8	C7 RC4 Prob8 - Fractions	0–1	0.36	0.41
C7R4MPB9	C7 RC4 Prob9 - Area and Volume	0–1	0.16	0.30

NOTE: Table estimates based on C1_7SC0 panel weight. Table estimates may differ from those reported in earlier user's manuals and psychometric reports because of re-estimation of scores on a longitudinal scale that includes eighth grade, and because of sample attrition. See chapter 7, section 7.5 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 2007.

The following are some examples of interpretation and use of the proficiency probability scores:

- Children's skills in making inferences based on cues directly stated in text (literal inference) increased dramatically between first and third grade, from 18 percent, or a mean probability = 0.18 (C4R4RPB6), to 68 percent (C5R4RPB6). Nearly all children, 92 percent, had mastered this skill by eighth grade (C7R4RPB6).

- 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. Only 25 percent mastered the evaluation level in third grade (C5R4RPB8), with 44 percent demonstrating mastery in fifth grade (C6R4RPB8) and 64 percent in eighth grade (C7R4RPB8).

- Twenty-six percent of eighth-graders were proficient at critical evaluation of nonfiction (C7R4RPB9), up from only 6 percent in fifth grade (C6R4RPB9).

- Only 6 percent of eighth-graders were able to evaluate complex syntax and understand high-level vocabulary in a biographical passage (C7R4RPB10).

- Fourteen percent of children understood interpretation and manipulation of simple fractions (C6R4MPB8) by the spring of fifth grade, and 36 percent by spring of eighth grade (C7R4MPB8).

- Three percent of fifth-graders could solve word problems involving area and volume (C6R4MPB9), with 16 percent of children demonstrating mastery in eighth grade (C7R4MPB9).

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 3.1.5 for further discussion of measurement of gain.

3.1.4 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; or

- a criterion-referenced versus norm-referenced interpretation.

3.1.4.1 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, if using eighth-grade scores, or data from prior rounds only, or both (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 statuses.

- **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 proficiency scores each target 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 eighth grade may be working on learning to comprehend complex syntax and vocabulary and make evaluative judgments based on reading material, which would show up as large gains in reading levels 8, 9, and 10. Less skilled readers may show their largest gains between fifth and eighth grades at levels 6 or 7, literal inference and extrapolation, catching up with the skill levels achieved by many of their peers in earlier rounds. 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.

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

The routing test number-right 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 eighth-grade reading, mathematics, and science assessments are based on 10 items in each domain. 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. However, because of the limited number of items in the routing tests, it is important to remember that these scores do not represent a comprehensive sample of the relevant domain of knowledge. The primary purpose of the routing tests was selection of appropriate second-stage forms.

- **Highest proficiency level mastered** is based on the same sets of items as the proficiency probability scores but consists 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 in the earlier rounds, and about 80 percent for reading and 50 percent for mathematics in eighth grade, the item data was supplemented with IRT-based estimates so that the "highest level" scores would not have to be reported as missing data. The higher percentages in eighth grade are a result of the necessary placement of proficiency level items on either the low or high second-stage forms, based on their estimated difficulty levels. Therefore, analysis of missing data that is 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. The *ECLS-K Psychometric Report for the Eighth Grade* (NCES 2009–002) (Najarian , Pollack, and Sorongon forthcoming) describes this treatment in more detail.

3.1.5 Measuring Gains

This section outlines approaches to measuring gains that rely on multiple criterion-referenced points to identify different patterns of child 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 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 11 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. Between spring-kindergarten and spring-first grade, mean reading scale scores increased by about 30 points for all subgroups, with additional 49-point gains by third grade, 23 more points by fifth grade, and 19 point gains by eighth grade. Similarly, each of these groups gained about 10 points, on average, on the mathematics scale during kindergarten, again starting from a very different initial status. Gains as of first, third, fifth, and eighth grades averaged approximately 25 points, 37 points, 24 points, and 17 points, respectively, for most subgroups. The differences among groups in *gains* in scale score points are relatively small, while the differences in subgroup scale score means are much larger. The *ECLS-K Psychometric Report for the Eighth Grade* (NCES 2009–002) (Najarian, Pollack, and Sorongon forthcoming) describes this analysis in more detail.

It would be incorrect to conclude that, because different subgroups of children are quantitatively gaining 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 mastering reading comprehension skills.

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 test-takers have an equal chance to gain on the vertical scale. 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 each child's respective 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 10 reading and 9 mathematics proficiency levels described in section 3.1.3, 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 0.23, analogous to 23 percent of the first-grade population demonstrating mastery of this set of items. The mean probability at the end of third grade, 0.77, is equivalent to a population mastery rate of 77 percent, with a mastery rate of 93 percent by the end of fifth grade and 98 percent in eighth 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 14 percent at the end of third grade, 43 percent at fifth grade, and 67 percent at eighth grade. By the end of eighth grade, nearly all children (89 percent) demonstrated mastery of level 6 mathematics skills (understanding place value), while the majority had not yet shown the same level of competence at level 8 (fractions: 36 percent proficient) and level 9 (area and volume: 16 percent proficient). These proportions, and the average gains in the proportions for the various skills, would very likely be quite different for subgroups of children defined by various demographic and school-process categories. Similarly, gains at each level between one assessment round and a subsequent round 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 any two selected time points 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 outcome measures to be explained by background and process variables.

Each different analytical approach provides a different perspective with respect to understanding children's 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-fifth grade to spring-eighth grade, but may be more important within grade, especially fall- to spring-kindergarten. Analysts should also keep in mind that, as the longitudinal data collection progresses, increasing numbers of children are not in the modal grade for the sample. Children's grade levels, and the consequent differences in curriculum exposure at the time of assessment rounds, should be taken into account. In designing an analysis plan, it is important to consider whether and how differences in ages, assessment dates and intervals, and children's grade levels 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.

3.1.6 Reliability

Reliability statistics assess consistency of measurement, in other words, the extent to which test items in a set are 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 routing test number-right, 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 .84 to .92 for the three subjects in eighth grade, 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 .70 to .76 for the eighth-grade assessment forms. These coefficients are relatively low because the routing tests consisted of only 10 items each. Alpha coefficients for the second-stage forms in each subject ranged from .68 to .82. The restriction of range of ability of children taking each second-stage form would tend to depress the alpha coefficients (relative to the routing test), while the greater number of items in the second stage would have the opposite effect. The alpha coefficients for individual sections of the tests are reported here although the test scores that are most useful and informative are those based on the children's complete sets of test responses.

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 child 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 child's highest level mastered could be determined on the basis of actual item response data alone for only about 19 percent of the reading and 47 percent of the mathematics eighth-grade scores, because the clusters of items marking some of the proficiency levels appeared only in some of the test forms. Alternatively, IRT ability estimates and item parameters could be used to generate pass/fail scores, and the composite highest level scores, for these same children. 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 (albeit slightly lower in eighth grade) between the methods indicates that the IRT approach supports the use of the highest level score sufficiently well for use in aggregate statistics.

Tables 3-10 through 3-12 present the reliability statistics for all of the assessment scores in eighth grade.

Table 3-10. 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, 2003–04, and 2006–07

Domain	Fall-kindergarten	Spring-kindergarten	Fall-first grade	Spring-first grade	Spring-third grade	Spring-fifth grade	Spring-eighth grade
Reading	.92	.95	.96	.96	.94	.93	.87
Mathematics	.91	.93	.94	.94	.95	.95	.92
Science	†	†	†	†	.87	.87	.84

† Not applicable.

NOTE: Approximately 89 percent of the children interviewed were in eighth grade during the 2006–07 school year, 9 percent were in seventh grade, and 2 percent were in sixth 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), the *ECLS-K Psychometric Report for the Third Grade* (NCES 2005—062) (Pollack, Rock et al. 2005), and the *ECLS-K Psychometric Report for the Fifth Grade* (NCES 2006–-036rev) (Pollack, Atkins-Burnett et al. 2005) because of re-estimation of scores on a longitudinal scale that includes eighth grade, and because of sample attrition.

SOURCE: U.S. Department of Education, National Center for Education Statistics, Early Childhood Longitudinal Study, Kindergarten Class of 1998–99 (ECLS-K), spring 2007.

Table 3-11. Reliability of routing test number correct (alpha coefficient), by round of data collection and domain: School years 1998–1999, 1999–2000, 2001–02, 2003–04, and 2006–07

Domain	Fall-kindergarten	Spring-kindergarten	Fall-first grade	Spring-first grade	Spring-third grade	Spring-fifth grade	Spring-eighth grade
Reading	.86	.88	.88	.86	.75	.88	.73
Mathematics	.78	.81	.83	.80	.86	.88	.76
Science	†	†	†	†	.75	.79	.70

† Not applicable.

NOTE: Approximately 89 percent of the children interviewed were in eighth grade during the 2006–07 school year, 9 percent were in seventh grade, and 2 percent were in sixth 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), the *ECLS-K Psychometric Report for the Third Grade* (NCES 2005–062) (Pollack, Rock et al. 2005), and the *ECLS-K Psychometric Report for the Fifth Grade* (NCES 2006–036rev) (Pollack, Atkins-Burnett et al. 2005) because of re-estimation of scores on a longitudinal scale that includes eighth grade, and because of sample attrition.

SOURCE: U.S. Department of Education, National Center for Education Statistics, Early Childhood Longitudinal Study, Kindergarten Class of 1998–99 (ECLS-K), spring 2007.

Table 3-12. Percent agreement of highest proficiency level mastered score, by round of data collection: School years 1998–1999, 1999–2000, 2001–02, 2003–04, and 2006–07

Domain	Fall-kindergarten	Spring-kindergarten	Fall-first grade	Spring-first grade	Spring-third grade	Spring-fifth grade	Spring-eighth grade
Reading							
Exact Agreement	63	54	55	55	50	51	44
Exact + Off by 1	96	94	94	95	95	95	89
Mathematics							
Exact Agreement	54	51	52	57	56	55	61
Exact + Off by 1	97	95	96	97	97	97	98

NOTE: Approximately 89 percent of the children interviewed were in eighth grade during the 2006–07 school year, 9 percent were in seventh grade, and 2 percent were in sixth 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), the *ECLS-K Psychometric Report for the Third Grade* (NCES 2005–062) (Pollack, Rock et al. 2005), and the *ECLS-K Psychometric Report for the Fifth Grade* (NCES 2006–-036rev) (Pollack, Atkins-Burnett et al. 2005) because of re-estimation of scores on a longitudinal scale that includes eighth grade, and because of sample attrition.

SOURCE: U.S. Department of Education, National Center for Education Statistics, Early Childhood Longitudinal Study, Kindergarten Class of 1998–99 (ECLS-K), spring 2007.

3.1.7 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, and the judgments of curriculum experts all provided input to test specifications.

The ECLS-K test specifications were derived from a variety of sources. For the third- through eighth-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- and eighth-grade frameworks are similar to the NAEP fourth- and eighth-grade frameworks, with some differences due to ECLS-K formatting and administration constraints. The NAEP fourth-grade frameworks were modified for third and fifth grades (and for the earlier K-1 forms), while the eighth-grade frameworks were used as defined in NAEP. An expert panel of secondary school educators, including curriculum specialists in the subject areas, examined the pool of items. 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 eighth-grade assessment forms. Some compromises in matching target percentages were necessary to satisfy constraints related to other issues, including linking to K-1, third-grade, and fifth-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.)

3.2 Indirect Cognitive Assessment

English, mathematics, and science teachers were asked to rate each sampled child on his or her skills in areas relevant to the subject taught. English teachers were asked about children's skills in written and oral expression. Mathematics teachers were asked about children's skills in mathematics, such as problem solving and demonstrating mathematical reasoning. Science teachers were asked about children's skills in science, such as designing an experiment to solve a scientific question and writing a report and preparing a presentation of scientific data. In earlier grades, teachers also rated children's achievement in a fourth domain: social studies. Teachers rated each child's skills, knowledge, and behaviors as "Outstanding (5)," "Very Good (4)," "Good (3)," "Fair (2)," or "Poor (1)." If a skill, knowledge, or behavior had not been introduced into the classroom yet, or if the teacher otherwise did not have the opportunity to observe the skill, the teacher was able to code that item as "Not Applicable/Not Observed." In eighth grade, many schools are departmentalized so different teachers may be rating the child on science and mathematical thinking. All children were rated on their English skills by their English teacher. Half of the children were rated on their mathematics skills by their mathematics teacher,

and half were rated on their science skills by their science teacher. The differences between the direct and indirect cognitive assessments, and the scores available, are described here.

3.2.1 Comparison to Direct Cognitive Assessment

The teacher ratings overlap and 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 important, the teacher rating scales include 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 how they express their ideas, solve mathematical problems, or investigate scientific phenomena. The language and literacy teacher ratings collect information on children's oral expression and written composition, areas not assessed on the direct measure.

These criterion-referenced indirect measures are targeted to the specific grade level of the child and draw upon the daily observations made by teachers of the children in their class.

3.2.2 Scores Available for the Teacher Ratings

IRT analysis using a generalized partial credit model (Muraki 1992) was used to create measures of the reported performance of children on a hierarchy of skills, knowledge, and behavior. The generalized partial credit model, as implemented in the SSI Parscale computer program, uses the pattern of ratings on items to obtain an estimate of the difficulty of each item and to place each child on an interval scale set with a minimum score of one and a maximum score of five. The analysis showed that the reliability of the estimates of the child's ability was very high for all domains (see table 3-13).

Table 3-13. Teacher rating scale reliability statistics for the IRT-based score, by category: School year 2006–07

Category	Grade 8
Written Expression skill ratings	.96
Oral Expression skill ratings	.93
Mathematics skill ratings	.95
Science skill ratings	.95

NOTE: Approximately 89 percent of the children interviewed were in eighth grade during the 2006–07 school year, 9 percent were in seventh grade, and 2 percent were in sixth or other grades. See chapter 7, section 7.5 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 2007.

As mentioned earlier, the teacher rating 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 child scores are placed on a common scale. Children had a high probability of receiving a high rating on items whose difficulty was below their scale score, and a lower probability of receiving a high rating on items above their scale score. Therefore, the scores received on the subscales should not be interpreted as mean scores, but as the child's relative probability of success with the items. Bayesian estimation techniques allow children who received maximum ratings on all the items or minimum ratings on all the items to receive a rating score.

The variable names, descriptions, value ranges, weighted means, and standard deviations for the eighth-grade (T7) teacher rating scores are shown in table 3-14. 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 eighth-grade teacher ratings are different from the Academic Rating Scale (ARS) ratings in earlier rounds of data collection, so the scores are not directly comparable to those for kindergarten, first, third, or fifth grades. The children's scores are calculated in relation to the item difficulty. With different items used across the grades and separate calibrations performed, the scale metric differs from one grade to another.

Table 3-14. Teacher rating scale range, mean, and standard deviation (weighted): School year 2006-07

Variable name	Description	Weighted mean	Standard deviation
T7ARSMAT	T7 Mathematics skills score	2.48	1.17
T7ARSSCI	T7 Science skills score	2.38	1.28
T7ARSORL	T7 English oral expression score	2.73	1.20
T7ARSWRT	T7 English writing skills score	2.40	1.30

NOTE: Table estimates based on C7CW0 weight. See chapter 7, section 7.5 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 2007.

Tables 3-15 to 3-18 report item difficulty estimates for the eighth-grade teacher questionnaire rating scales. Higher values imply that teachers rated fewer children as proficient on those items. Children would have a greater than 50 percent probability of receiving ratings of "5" on items below their ability level.

Table 3-15. Spring-eighth grade Oral Expression Skills item difficulties (arranged in order of difficulty): School year 2006–07

Item difficulty	Item number and abbreviated content
2.19	Q12a. Uses Spoken English Grammar
2.61	Q12c. Expresses Creative Thinking
2.72	Q12b. Expresses Analytical or Critical Thinking

SOURCE: U.S. Department of Education, National Center for Education Statistics, Early Childhood Longitudinal Study, Kindergarten Class of 1998–99 (ECLS-K), spring 2007.

Table 3-16. Spring-eighth grade Written Expression Skills item difficulties (arranged in order of difficulty): School year 2006–07

Item difficulty	Item number and abbreviated content
2.40	Q11a. Organizes Ideas Logically and Coherently
2.46	Q11c. Gathers Information for Research Purposes
2.46	Q11b. Employs English Grammar and Usage
2.53	Q11d. Writes Various Types of Composition
2.85	Q11e. Uses Style and Rhetoric

SOURCE: U.S. Department of Education, National Center for Education Statistics, Early Childhood Longitudinal Study, Kindergarten Class of 1998–99 (ECLS-K), spring 2007.

Table 3-17. Spring-eighth grade Mathematic Skills item difficulties (arranged in order of difficulty): School year 2006–07

Item difficulty	Item number and abbreviated content
1.48	Q11f. Uses Calculator to Solve Problems
2.23	Q11g. Uses Computer to Complete Mathematics Assignments
2.68	Q11a. Applies Mathematical Concepts to Real World
2.68	Q11c. Talks about Reasoning in Solving a Problem
2.74	Q11e. Uses Representations to Model Mathematical Ideas
2.82	Q11d. Explains Reasoning in Solving a Problem in Writing
2.85	Q11b. Conducts Proofs or Demonstrates Mathematical Reasoning

SOURCE: U.S. Department of Education, National Center for Education Statistics, Early Childhood Longitudinal Study, Kindergarten Class of 1998–99 (ECLS-K), spring 2007.

Table 3-18. Spring-eighth grade Science Skills item difficulties (arranged in order of difficulty): School year 2006–07

Item difficulty	Item number and abbreviated content
2.33	Q11a. Organizes Data in Tables and Charts
2.50	Q11f. Applies Science Concepts to Solve Real World Problems
2.52	Q11c. Talks about Investigations to Solve Problems
2.57	Q11b. Writes Up Results or Presentation for Research Project
2.64	Q11d. Makes Presentation to Class about Science Analysis
2.79	Q11e. Designs Experiment to Solve Scientific Question

SOURCE: U.S. Department of Education, National Center for Education Statistics, Early Childhood Longitudinal Study, Kindergarten Class of 1998–99 (ECLS-K), spring 2007.

The teacher ratings scale was designed to provide information on children's abilities at a given point in time, not necessarily over time. Moreover, these teacher rating scales are placed on a different metric than the ARS scores in previous rounds. Therefore, change scores cannot be calculated between time points.

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 Self-Description Questionnaire

Beginning in the third-grade data collection in the ECLS-K, children were asked to provide self-assessments of their academic and social skills. For the eighth-grade data collection, children rated

their perceived competence and interest in English and mathematics. Children also reported on problem behaviors with which they might struggle. The Internalizing Problems scale included items on sadness, loneliness, and anxiety. Items for the English and mathematics scales were drawn from the Self Description Questionnaire (SDQ) II,[18] which was designed for children in middle and high school. Items for the eighth-grade Internalizing Problems scale were drawn from the fifth-grade Internalizing Problems scale as recommended by the Content Review Panel because these items better reflected the constructs that the study intended to measure and also allowed for comparison with previous rounds of data collection. For further description of the ECLS-K self-description questionnaire (SDQ) see chapter 2, section 2.1.2.

Children rated whether each item was "not at all true," "a little bit true," "mostly true," or "very true." Three scales were produced from the eighth-grade SDQ items. The scale scores on all eighth-grade SDQ scales represent the mean rating of the items included in the scale. Children who responded to the eighth-grade 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).

Table 3-19 presents the internal consistency reliability estimates of the eighth-grade SDQ scales, as measured by Cronbach's coefficient alpha. The Cronbach's coefficient alpha for the Perceived Interest and Competence in Math is similar to that found by the scale's authors (alpha = .89; Ellis, Marsh, and Richards 2002). However, the coefficient for the eighth-grade Perceived Interest and Competence in English scale is lower than that found by the scale's authors (alpha = .88; Ellis, Marsh, and Richards). The coefficient alpha for the eighth-grade Internalizing Problem Behaviors scale is consistent with the findings from the ECLS-K fifth-grade data (alpha = .79) (Pollack, Atkins-Burnett et al. 2005).

[18] The items were adapted with permission from the Self Description Questionnaire (SDQ II), from *Self Description Questionnaire (SDQ) II: A theoretical and empirical basis for the measurement of multiple dimensions of adolescent self-concept. An interim test manual and a research monograph*, by H.W. Marsh (Sydney: University of Western Sydney, SELF Research Centre, 1992). (Original work published in 1990.)

Table 3-19. Self-description questionnaire scale reliabilities (alpha coefficient): School year 2006–07

Variable	Description	Number of items	Alpha coefficient
C7SDQRDC	C7 SDQ Prcvd Interest[1]/Competence - Reading	4	.76
C7SDQMTC	C7 SDQ Prcvd Interest/Competence - Math	4	.89
C7SDQINT	C7 SDQ Internalizing Problems	8	.75

[1] "Prcvd Interest" = Perceived Interest.
NOTE: See chapter 7, section 7.5 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 2007.

Table 3-20 presents the variable names, scale ranges, means, and standard deviations (weighted) for the self-description questionnaire.

Table 3-20. Self-description questionnaire scale range, mean, and standard deviation (weighted): School year 2006-07

Variable	Description	Range of values	Weighted mean	Standard deviation
C7SDQRDC	C7 SDQ Prcvd Interest[1]/Competence - Reading	1–4	2.52	.78
C7SDQMTC	C7 SDQ Prcvd Interest/Competence - Math	1–4	2.62	.91
C7SDQINT	C7 SDQ Internalizing Problems	1–4	2.03	.57

[1] "Prcvd Interest" = Perceived Interest.
NOTE: Table estimates based on C7CW0 weight. See chapter 7, section 7.5 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 2007.

3.4 Self-Concept and Locus of Control Scale Scores

The Self-Concept and Locus of Control scales were adopted from the National Education Longitudinal Study of 1988 (NELS:88). These scales ask children about their self-perceptions and the amount of control they have over their own lives. Items were drawn from the NELS:88 student questionnaire and asked children to indicate the degree to which they agreed with 13 statements about themselves. Statements reflected perceptions children might have about themselves and about how much control they felt they had over their own lives. Children rated whether they "strongly agree," "agree," "disagree," or "strongly disagree" with each item.

In order to be as comparable as possible to NELS:88, scale scores were calculated with the same procedures as NELS:88. Some items were positively worded, and some were negatively worded. As

a result, scoring for some items was reversed to provide an appropriate score. For the Self-Concept scale, three of the seven items in the scale were reverse scored before performing computations, so that higher scores indicate more positive self- concept:

- I certainly feel useless at times.
- At times I think I am no good at all.
- I feel I do not have much to be proud of.

The seven items in the scale were then standardized separately to a mean of zero and a standard deviation of 1. The scale score is an average of the seven standardized scores.

For the Locus of Control scale, five items were reverse scored so that higher scores indicate greater perception of control over one's own life:

- I don't have enough control over the direction my life is taking.
- In my life, good luck is more important than hard work for success.
- Every time I try to get ahead, something or somebody stops me.
- My plans hardly ever work out, so planning only makes me unhappy.
- Chance and luck are very important for what happens in my life.

The six items in the scale were then standardized separately to a mean of zero and a standard deviation of 1. The scale score is an average of the six standardized scores.

Children who responded to the Self-Concept and Locus of Control items answered virtually all of the questions, so treatment of missing data was not an issue.

Table 3-21 presents the internal consistency reliability estimates of the Self-Concept and Locus of Control scales, as measured by Cronbach's coefficient alpha. The coefficient alpha for both scales is consistent with the findings from the NELS:88 data ($alpha_{Self\text{-}Concept} = .79$, $alpha_{Locus\ of\ Control} = .68$) (Ingels et al. 1990).

Table 3-21. Self-Concept and the Locus of Control scale reliabilities (alpha coefficient): School year 2006–07

Variable	Description	Number of items	Alpha coefficient
C7CONCPT	C7 Self concept	7	.81
C7LOCUS	C7 Locus of control	8	.75

NOTE: See chapter 7, section 7.5 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 2007.

The variable names, descriptions, value ranges, weighted means, and standard deviations of the Self-Concept and Locus of Control scales are shown in table 3-22.

Table 3-22. Self-Concept and the Locus of Control scale range, mean, and standard deviation (weighted): School year 2006-07

Variable	Description	Range of values	Weighted mean	Standard deviation
C7CONCPT	C7 Self concept	-1.12 - +3.06	0.00	.70
C7LOCUS	C7 Locus of control	-1.53 - +2.50	0.02	.64

NOTE: Items were standardized to a mean of 0 and a standard deviation of 1. Table estimates based on C7CW0 weight. See chapter 7, section 7.5 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 2007.

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 children 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 child 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 children who had transferred from their kindergarten schools was followed.

The third-grade data collection targeted base-year respondents and children sampled in first grade through the freshening operation in which 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 in which only a subsample of children who had transferred from their kindergarten schools 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 follow-up if they moved in third grade.

The fifth-grade data collection set differential sampling rates for movers in different categories. It also excluded four special groups of children, irrespective of other subsampling procedures that were implemented. The excluded children were those who became ineligible in an earlier round because they died or moved out of the country; who were subsampled out in previous rounds because they were movers; whose parents emphatically refused to cooperate (hard refusals); and who were eligible for the third-grade data collection but had neither first-grade nor third-grade data. Of the remaining children, those who moved from their original schools during fifth grade or earlier were subsampled for follow-up. Children whose home language was 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.

The eighth-grade sample included all children eligible after fifth grade regardless of their fifth-grade response status. The ineligible children were those who moved out of the country, were deceased, or moved to another school and were not subsampled for follow-up in fifth grade. There was no subsampling of movers for follow-up as in previous rounds since the vast majority of children were not in the same school from kindergarten to eighth grade (having moved out of elementary schools into middle schools), and subsampling these movers would result in substantial losses in sample size and precision of the estimates for eighth grade.

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 with 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 NCES Common Core of Data (CCD), "Public Elementary/Secondary School Universe Survey," 1995–96, and the NCES Private School Universe Survey (PSS), 1995–96. 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 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 first grade and twelfth 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 and spring 1999.

The third-stage sampling units were children of kindergarten age, selected within each sampled school. The goal of the child sample design was to obtain an approximately self-weighting sample of children 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 children and the second, all other children. Within each stratum, children were selected using equal probability systematic sampling, using a higher rate for the API stratum.[19] 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 was the same for fall- and spring-kindergarten, the child sample was 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 2001–029rev) (Tourangeau, Burke et al. 2004).

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

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

Characteristic	Total	Sector	
		Public	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
More than one race	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 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
Master's	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 first grade and twelfth 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 and spring 1999.

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.

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 first grade and twelfth grade.
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 they were 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.

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
More than one race	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
Master's	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 first grade and twelfth grade.
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 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 child 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 children still enrolled in their base-year schools were recontacted, only a 50 percent subsample of base-year sampled children who had transferred from their kindergarten school was followed for data collection.

4.3.1 Subsampling Movers

As noted earlier, 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 follow-up, 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 children followed during spring-first grade, and any transfer child 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 children. Region, locale, school affiliation, and school type describe the school in which the child attended kindergarten.

Table 4-6. Number (unweighted) of eligible children in spring-first grade sample excluding freshened children, 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
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
More than one race	434	347	87
Unknown	54	45	9

See notes at end of table.

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

Characteristic	Total	Sector	
		Public	Private
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
Master's	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 first grade and twelfth grade.
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 Child Freshening

The spring-first grade child 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 in which transfer children were flagged for follow-up. Each of these schools was asked to prepare an alphabetized roster of children enrolled in first grade, and the names of ECLS-K kindergarten-sampled children were identified on this list. Beginning with the name of the first kindergarten-sampled child, school records were checked to see whether the child 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 child name, and so on. Note: the child 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). Child 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 child 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. However, 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 children who had not transferred from the school in which they had been sampled during the base year. A more detailed discussion of freshened child 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 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 children entered the sample.

While all children still enrolled in their base-year schools were recontacted, slightly more than 50 percent of the base-year sampled children who had transferred from their kindergarten school were followed for data collection. This subsample of children 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 children). 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 follow-up still targeted the same 50 percent subsample of children in the base-year

schools). In addition, language minority children who moved between first grade 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 follow-up 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 affects only 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 follow-up procedure provides a larger sample of children whose home language is not English. The mover follow-up 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 children. 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 children, 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
More than one race	432	351	81
Unknown	39	32	7

See notes at end of table.

Table 4-7. Number (unweighted) of eligible children in spring-third grade sample excluding freshened children, 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
Master's	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 first grade and twelfth grade.
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 *ECLS-K 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 were 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 2003–04

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
Socioeconomic status 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 follow-up. 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 were 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 were as follows:

- 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 follow-up 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 follow-up activities targeted a 42 percent subsample of movers who 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 children. Region, locale, school affiliation, and school type describe the school at which the child attended kindergarten.

A new feature of the fifth-grade sample was 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.

Table 4-9. Number (unweighted) of eligible children in spring-fifth grade sample excluding freshened children, 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
More than one race	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 children, 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
Master's	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 first grade and twelfth grade.
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.

4.6 Eighth-Grade Sample

The sample design for eighth grade called for including all 12,129 children eligible after fifth grade (regardless of their fifth-grade response status), and following all movers without any subsampling. In the ECLS-K first-grade to fifth-grade data collections, subsampling of movers was used to reduce data collection costs. The initial sample size was developed taking into account the reduction in sample size and increase in the variability of the weights of the respondents resulting from the subsampling. As the design was extended beyond fifth grade (the initial planning of the ECLS-K did not plan for this extension into eighth grade), a change in the methods of handling movers to avoid subsampling them was needed to achieve the major analytic goals. The vast majority of children were not in the same school from kindergarten to eighth grade (having moved out of elementary schools into middle schools), and subsampling these movers would result in substantial losses in sample size and precision of the estimates for the eighth grade.

Table 4-10 shows the characteristics of eligible children in the spring-eighth grade sample, excluding freshened children. Region, locale, school affiliation, and school type describe the school in which the child attended kindergarten.

4.7 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.

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 4-11 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 4-12 and 4-13 show the same children separately by the original sample school affiliation (public/private).

Table 4-10. Number (unweighted) of eligible children in spring-eighth grade sample excluding freshened children, by selected characteristics: School year 2006–07

Characteristic	Total	Sector	
		Public	Private
Total	11,929	9,482	2,447
Region			
Northeast	2,223	1,679	544
Midwest	3,107	2,341	766
South	3,820	3,211	609
West	2,779	2,251	528
Type of locale			
Large city	2,171	1,601	570
Midsize city	2,352	1,684	668
Urban fringe of large city	3,394	2,740	654
Urban fringe of midsize city	831	737	94
Large town	370	293	77
Small town	1,131	877	254
Rural	1,680	1,550	130
School affiliation			
Public	9,482	9,482	†
Catholic	1,467	†	1,467
Non-Catholic, religious	697	†	697
Nonreligious, private	283	†	283
School type			
Regular[1]	11,529	9,334	2,195
Ungraded	11	3	8
No grade beyond kindergarten	202	141	61
Unknown	187	4	183
Child's race/ethnicity			
White	6,815	5,053	1,762
Black	1,354	1,219	135
Hispanic, with race	1,092	924	168
Hispanic, without race	1,144	1,012	132
Asian	846	697	149
Pacific Islander	153	139	14
Native American	224	200	24
More than one race	285	226	59
Unknown	16	12	4

See notes at end of table.

Table 4-10. Number (unweighted) of eligible children in spring-eighth grade sample excluding freshened children, by selected characteristics: School year 2006–07—Continued

Characteristic	Total	Sector	
		Public	Private
Highest parent level of education			
Less than high school	984	964	20
High school graduate	2,379	2,164	215
Vocational/technical	677	602	75
Some college	3,333	2,721	612
College graduate	2,717	1,893	824
Master's	1,116	722	394
Ph.D./professional	679	376	303
Unknown	44	40	4
Home language			
Not English	3,436	2,863	573
English	8,493	6,619	1,874

† Not applicable.
[1] School offers kindergarten and at least another grade between first grade and twelfth grade.
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 2007.

Table 4-11. Number (unweighted) of children in the ECLS-K sample, by response status and data collection round: School years 1998–99, 1999–2000, 2001–02, 2003–04, and 2006–07

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
Spring-eighth grade	12,129[4]	36	67	†	2,301	9,725

† 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, hard-to-field cases such as hard refusals, and children with neither first-grade nor third-grade data.
[4] 12,129 fifth-grade respondents and eligible respondents were eligible for the eighth-grade data collection.
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 follow-up 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 include only 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, spring 2004, and spring 2007.

Table 4-12. 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, 2003–04, and 2006–07

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
Spring-eighth grade	9,655[4]	28	60	†	1,919	7,648

† 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 from the original sample of 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, hard-to-field cases such as hard refusals, and children with neither first-grade nor third-grade data.

[4] 9,655 fifth-grade respondents and eligible respondents from the original sample of public schools were eligible for the eighth-grade data collection.

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 follow-up 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, spring 2004, and spring 2007.

Table 4-13. 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, 2003–04, and 2006–07

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
Spring-eighth grade	2,474[4]	8	7	†	382	2,077

† 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 the original sample of 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, hard-to-field cases such as hard refusals, and children with neither first-grade nor third-grade data.
[4] 2,474 fifth-grade respondents and eligible respondents from the original sample of private schools were eligible for the eighth-grade data collection.
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 follow-up 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 include only 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, spring 2004, and spring 2007.

The number of children who participated in all five years of the ECLS-K data collection (base year, first grade, third grade, fifth grade, and eighth grade) is 8,706 (6,911 in original public schools and 1,795 in original private schools). This represents 41 percent of the base-year respondents or 38 percent of children sampled for the base year.

4.8 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-, fifth- or eighth-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-, fifth-, and eighth-grade data are representative of the population cohort rather than all third-graders in 2001–02 or all fifth-graders in 2003–04 or all eighth-graders in 2006–07. 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 data collection, the estimated number of fifth-graders is approximately 83 percent of all fifth-graders. From the eighth-grade data collection, the estimated number of eighth-graders is approximately 80 percent of all eighth-graders. While the vast majority of children in third grade in the 2001–02 school year, in fifth grade in the 2003–04 school year, and in the eighth grade in the 2006–07 school year are members of the cohort, third-graders who repeated second or third grade, fifth-graders who repeated third or fourth grade, eighth-graders who repeated fifth, sixth, or seventh 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, of fifth-grade teachers and schools in 2003–04, nor of eighth-grade teachers and schools in 2006–07. 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 eighth 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; student questionnaires (third, fifth and eighth grade only); parent instruments; several types of teacher instruments completed by reading or English, 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 eighth-grade cross-sectional sampling weights be computed for use in analyzing the eighth-grade ECLS-K data, as was the case with previous rounds of data collection. Several combinations of kindergarten through eighth-grade longitudinal weights were also computed. Details on these longitudinal weights are available in chapter 9 and in chapter 10 for users of the K-8 full sample public-use data file. This section describes the different types of eighth-grade cross-sectional weights, how they were calculated, how they should be used, and their statistical characteristics.

4.8.1 Types of Cross-Sectional Sample Weights

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

- C7CW0 is nonzero if assessment data or student questionnaire data are present (or the child was excluded from direct assessment due to a disability).
- C7PW0 is nonzero if parent interview data are present.
- C7CPTE0 is nonzero if assessment data or student questionnaire data are present (or the child was excluded from direct assessment due to a disability), and parent interview data, and teacher-level data from the English teacher are present.
- C7CPTM0 is nonzero if the child was sampled to have a child-level questionnaire completed by the mathematics teacher, and assessment data or student questionnaire data are present (or the child was excluded from direct assessment due to a disability), and parent interview data, and teacher-level data (either from the English teacher or the mathematics teacher) are present.
- C7CPTS0 is nonzero if the child was sampled to have a child-level questionnaire completed by the science teacher, and assessment data or student questionnaire data are present (or the child was excluded from direct assessment due to a disability), and parent interview data, and teacher-level data (either from the English teacher or the science teacher) are present.

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

Prior to the fifth-grade data collection, only one child-parent-teacher weight was computed based on the presence of the teacher questionnaire B (teacher-level). With the addition beginning in fifth grade 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) with or without 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 eighth-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 eighth-grade cross-sectional weights: School year 2006–07

Weight	To be used for analysis of ...
C7CW0	child direct assessment or student questionnaire data from spring-eighth grade, alone or in combination with (a) a limited set of child characteristics (e.g., age, sex, and race/ethnicity), (b) data from any spring-eighth grade teacher questionnaire (teacher-level or child-level), or (c) data from the spring-eighth grade school administrator questionnaire.
C7PW0	parent interview data from spring-eighth grade, alone or in combination with (a) spring-eighth grade child assessment or student questionnaire data, (b) data from any spring-eighth grade teacher questionnaire (teacher-level or child-level), or (c) data from the spring-eighth grade school administrator questionnaire. *Exception*: If data from the parent interview AND child assessments AND teacher-level (with or without child-level teacher) questionnaires are used together, then C7CPTE0, C7CPTM0, or C7CPTS0 should be used.
C7CPTE0	child direct assessment or student questionnaire data from spring-eighth grade with spring-eighth grade parent interview data and spring-eighth grade English teacher-level data with or without child-level data from the English teacher, alone or in combination with data from the spring-eighth grade school administrator questionnaire.
C7CPTM0	child direct assessment or student questionnaire data from spring-eighth grade with spring-eighth grade parent interview data and spring-eighth grade English or mathematics teacher-level data with or without child-level data from the mathematics teacher, alone or in combination with data from the spring-eighth grade school administrator questionnaire. This weight is to be used only if the analytic sample is restricted to the subset of children who were sampled to have a mathematics teacher questionnaire.
C7CPTS0	child direct assessment or student questionnaire data from spring-eighth grade with spring-eighth grade parent interview data and spring-eighth grade English or science teacher-level data with or without child-level data from the science teacher, alone or in combination with data from the spring-eighth grade school administrator questionnaire. This weight is to be used only if the analytic sample is restricted to the subset of children who were sampled to have a science teacher questionnaire.

SOURCE: U.S. Department of Education, National Center for Education Statistics, Early Childhood Longitudinal Study, Kindergarten Class of 1998–99 (ECLS-K), spring 2007.

Weight C7CW0 is used to estimate child-level characteristics or assessment scores for eighth grade. Examples of such estimates are the percent of children who are in private schools, the percent of children who are API, the percent of children who are 15 years old at the beginning of the eighth-grade

data collection, and the mean reading score of children in the eighth-grade data collection. These weights exist not only for children who had assessment data but also for children who completed the student questionnaire or who could not be assessed due to a disability.[20] 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, and the school administrator questionnaire. Rating scale scores from teachers on children's skills in language and literacy (oral and written expression), science, and mathematical thinking (see chapter 3) are also available for these children, regardless of whether they completed the direct child assessment.

C7PW0 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 have child care, and the percent of children whose parents were concerned about their child's weight. 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. C7CPTE0 should be used if teacher-level data from the English teacher are analyzed with or without child-level data from the English teacher. Note that the teacher-level questionnaire can be completed by more than one teacher (English and/or mathematics; or, English and/or science). Therefore, C7CPTM0 should be used if teacher-level data from the English or mathematics teacher are analyzed with or without child-level data from the mathematics teacher. Likewise, C7CPTS0 should be used if teacher-level data from the English or science teacher are analyzed with or without child-level data from the science teacher. Weight C7CW0 may be used when analyzing child assessment data in conjunction with English 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. C7CPTE0 is used when child direct assessment or student questionnaire data *and* parent data *and* English teacher-level data *with or without* child-level data from English teachers are combined in an analysis; for example, in the analysis of the relationship between parent education, teacher education, and children's reading

[20] 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.

knowledge and skills. If it is the children's mathematics knowledge and skills as reported by the teacher that are analyzed, then C7CPTM0 should be used. Likewise, C7CPTS0 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.

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 (because the weighting adjustments were not correctly accounted for in the estimates). For example, if C7CPTE0 were used in an analysis of child- and teacher-level data only, then the resulting estimates will be inefficient compared to estimates using C7CW0. The lower parent response causes C7CPTE0 to result in a smaller sample with positive weights. If using C7CPTE0 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.

4.8.2 Weighting Procedures

Two features of the eighth-grade sample design that are different from previous grades are that (1) only fifth-grade respondents and eligible nonrespondents were included in the eighth-grade sample and (2) children who changed schools between fifth and eighth grade were not subsampled out but were all followed into their new school. However, a feature of the fifth-grade sample whereby children were subsampled for the administration of the mathematics or science questionnaires as discussed in section 4.5 was retained for eighth grade. The mathematics and science teacher questionnaires were administered to the same halves of the sample as was done in fifth grade. This is to allow for longitudinal analyses of data from the mathematics and science teacher questionnaires. The same subsamples of children selected for these two instruments in the fifth grade were maintained for eighth grade, i.e., children who had been assigned to have mathematics teacher questionnaires in fifth grade had mathematics teacher questionnaires in eighth grade, and likewise for the science teacher questionnaire.

These features of the design are taken into account in the weighting. The weighting procedures were divided into three main stages.

The starting point for the eighth-grade child weight is the fifth-grade child weight before adjustment for fifth-grade child nonresponse. It includes 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;
- adjustment of the base-year child weight for subsampling of schools for freshening in first grade (for children sampled in first grade only);
- adjustment for fifth-grade mover subsampling; and
- adjustment for fifth-grade unknown eligibility status.

Except for the last two adjustments, this starting weight is the same 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) was eligible for subsequent rounds of data collection. The starting weight was extracted from the first-grade weighting file to be used in eighth grade. The procedures used for computing these weights are described again in section 4.8.3 for completeness.

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

- eighth-grade unknown eligibility status; and
- eighth-grade child-level nonresponse.

For the mathematics and science child-parent-teacher weights, an additional adjustment was necessary (before the second-stage adjustment 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 4.8.4.2.

The computation of the initial child weights is described in section 4.8.3. The subsequent weight adjustments are described in section 4.8.4. Section 4.8.5 describes the different types of weights computed for spring-eighth 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.

4.8.3 Computation of Spring-Eighth 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), (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), (4) the adjustment for fifth-grade mover subsampling, and (5) the adjustment for fifth-grade unknown eligibility status. These weights were already computed for spring-fifth grade. For completeness, they are described below, in section 4.8.3.1 for the school-level weights, and in section 4.8.3.2 for the child-level weights.

4.8.3.1 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. See section 4.1 for a description of how schools were selected as part of the frame freshening procedure.

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 was 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.

4.8.3.2 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 who 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 the *ECLS-K Base Year Public-Use Data Files and Electronic Codebook: User's Manual* (NCES 2001–029rev) (Tourangeau, Burke 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 4.8.3.1) was adjusted for the subsampling of schools for freshening. Child 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 follows:

- 0 if the school was not in the set of schools subsampled for freshening;[21] 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 that 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 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.

[21] 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.

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 follows:

- 0 if the child was a mover with unknown grade (group c);
- 1 if the child was a nonmover, in first grade or in 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 follows:

- 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.

Base-year child weights adjusted for fifth-grade mover subsampling and fifth-grade unknown eligibility. First, the base-year child weights were adjusted to reflect the subsampling of movers in fifth grade. In the ECLS-K, a child could move more than once and at different times. For example, a child could move out of his or her original sample school because the school did not have grades higher than kindergarten. Then he or she 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 or she 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 or she 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 follow-up 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 follow-up		
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 base-year child weight was adjusted to reflect the subsampling of fifth-grade 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. Note that for the computation of the fifth-grade final weights, large mover adjusted weights for 12 children were trimmed by 40 percent, and the excess weight was not redistributed at this step since the total sum of weights was re-established later at the raking step of the fifth-grade final weights. For eighth grade, it was the untrimmed mover-adjusted weight that was used, so that the excess weight was not discarded at this point.

After the adjustment for subsampling movers, the child weights were adjusted for fifth-grade children whose eligibility was unknown (since subsampled out movers and children of unknown eligibility in fifth grade were not included in the eighth-grade sample). In fifth grade, a portion of children of unknown eligibility was assumed to be ineligible, equal to the proportion of children of known eligibility who were ineligible. To carry out this adjustment, each fifth-grade 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 adjustment factor for children of unknown eligibility) was computed as follows:

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

4.8.4 Computation of Spring-Eighth Grade Child Weights

4.8.4.1 Adjustment for Unknown Eligibility and Nonresponse

The initial child weights described in section 4.8.3 were adjusted for nonresponse in eighth grade, and raked to sampled-based control totals to obtain the final spring-eighth grade child weights.

The eighth-grade initial child weights described in section 4.8.3 were adjusted for eighth-grade nonresponse. As in previous years, 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 were 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 (children who were out of the country or deceased), or (d) of

unknown eligibility (children who could not be located for assessment). The first adjustment factor (for children of unknown eligibility) was computed as follows:

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

- 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 using fifth-grade moving status (all cross-sectional weights); response status of the child assessment and parent interview in the previous rounds (C7CW0 and C7PW0); the race/ethnicity of the child (C7CW0 and C7PW0); whether the child belonged to the language minority group (all cross-sectional weights); the type of household collected from the parent interviews (all cross-sectional weights except C7CW0); and the school affiliation including whether the child was homeschooled (C7CPTE0, C7CPTM0 and C7CPTS0 only). After nonresponse adjustment and prior to raking, very large weights were trimmed but not redistributed because the sum of weights was re-established after raking, described in section 4.8.4.2 below.

4.8.4.2 Raking to Sample-Based Control Totals

To reduce the variability due to the subsampling of schools and movers in fifth grade, the child weights were then raked to sample-based control totals computed using the initial child weights computed as described in section 4.8.3. The child records included in the file used for computing the control totals are records of fifth-grade respondents, eligible nonrespondents, and ineligible children. Records of fifth-grade ineligibles were part of raking in fifth grade, and needed to be included in the file for computing control totals for eighth grade (even though they were not eligible for eighth grade) in order for the sum of weights to be the estimated number of children who were in kindergarten in 1998–99 or in first grade in 1999–2000.

In the nonresponse adjustment step, the weights of the eighth-grade nonresponding children were distributed to the eighth-grade responding children while the weights of the eighth-grade ineligible children were not affected. At the end of raking, the weights of the ineligible children are nonzero, but will be set to zero because these children are not included in the analysis of the spring-eighth grade data. The reason for including the ineligible children in the raking step is that these children were included in the 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, census region, urbanicity, sex, age, race/ethnicity, socioeconomic status (SES), language minority status, whether sampled in kindergarten or first grade, and, if sampled in kindergarten, mover status.

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

In all three child-parent-teacher weights described in section 4.8.1, 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 he or she was assigned (i.e., English and mathematics, or English and science). A child could have one teacher who taught all subjects, in which case the teacher was asked to fill out both the English 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 might have an English teacher filling out the English 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 had teachers who 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 English and mathematics, or two teachers, one teaching English 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 English teacher—who filled out the teacher-level questionnaire (132 cases out of 4,114 or 3 percent). This is very different from fifth grade where a much larger number of children had only one teacher who taught both reading and mathematics. In eighth grade, the teaching structure changes for middle schools and almost all children have different teachers for different subjects. If the child had two teachers, then in the majority of cases, both teachers filled out the teacher-level questionnaire (3,810 cases out of 4,114 or 93 percent). There are very few cases where only one of the two teachers filled out the teacher-level questionnaire.

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 2006–07

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		4,114
1	English	132
2	English	89
2	Mathematics	83
2	English and Mathematics	3,810

SOURCE: U.S. Department of Education, National Center for Education Statistics, Early Childhood Longitudinal Study, Kindergarten Class of 1998–99 (ECLS-K), spring 2007.

Table 4-15 shows the same information for science. Since C7CPTM0 and C7CPTS0 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 (4,130 as shown in table 4-17 compared with 4,114 in table 4-14 for C7CPTM0; and 4,164 as shown in table 4-17 compared with 4,151 in table 4-15 for C7CPTS0). Using teacher-level data to define the child-parent-teacher weights is also consistent with previous years' practice.

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 2006–07

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		4,151
1	English	114
2	English	106
2	Science	82
2	English and Science	3,849

SOURCE: U.S. Department of Education, National Center for Education Statistics, Early Childhood Longitudinal Study, Kindergarten Class of 1998–99 (ECLS-K), spring 2007.

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 sections 4.8.4.1 and 4.8.4.2, respectively) was computed as follows:

- 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 follows:

- 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.

4.8.5 Types of Cross-Sectional Weights and Their Use

The different types of cross-sectional weights are described in section 4.8.1 and their use is summarized in exhibit 4-1. They were all created as described in section 4.8.2, but the definition of which children were eligible respondents varied for the different weights. Each weight was adjusted for

unknown eligibility and nonresponse and raked separately. There was no eighth-grade mover adjustment since all movers were followed into their new schools.

4.8.5.1 Cross-Sectional Weights To Be Used With Direct Child Assessment Data (C7CW0)

In spring-eighth grade, responding children for this type of weight were eligible children who had spring-eighth grade scorable direct child cognitive assessment data or student questionnaire data, or children with disabilities who, according to specifications in their Individualized Education Programs (IEPs), could not participate in the assessments. A child was eligible if he or she was a fifth-grade respondent or a fifth-grade eligible nonrespondent.

Table 4-16 shows the number of children who were not assessed in eighth grade due to the following special situations: children with disabilities, children who could not be located, children who had moved outside of the country or who were deceased, children whose parents refused consent, or children whose parents could not be located for consent. Of these, only children with disabilities had weights included in the eighth-grade data file.

Table 4-16. Number of children who were not assessed in spring-eighth grade, by special situations: School year 2006–07

Special situation	Number of children	
	Unweighted	Weighted
Spring-eighth grade		
Children with disabilities[1]	45	14,132
Ineligible (moved out of the country or deceased)	36	13,170
Not located	254	114,816
Parent refused consent	44	18,963
Parent not located for consent	744	323,230

[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 2007.

4.8.5.2 Cross-Sectional Weights To Be Used With Parent Data (C7PW0)

The weight C7PW0 is to be used with parent interview data. In spring-eighth 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 eighth-grade children, not to the parents of eighth-grade children.

4.8.5.3 Cross-Sectional Weights To Be Used With a Combination of Child Direct Assessment Data and Parent Interview Data and Teacher Data for Children With English Teacher Questionnaire (C7CPTE0)

The weight C7CPTE0 is to be used for analysis involving all children with child assessment, parent, and teacher-level data. If child-level data from English teachers are included in the analysis, then the same weight C7CPTE0 should be used. A respondent for this type of weight was defined as a child who had scorable cognitive assessment data or student questionnaire data for spring-eighth grade (or was excluded from direct assessment due to a disability), whose parent completed the FSQ section of the parent interview for spring-eighth grade, and who had completed teacher-level data from either the English teacher and/or the mathematics/science teacher.

4.8.5.4 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 (C7CPTM0)

The weight C7CPTM0 is to be used for analysis involving children who were subsampled to have a mathematics teacher questionnaire and who had child assessment, 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 or student questionnaire data for spring-eighth grade (or was excluded from direct assessment due to a disability), whose parent completed the FSQ section of the parent interview for spring-eighth grade, and who had completed teacher-level data from either the English teacher or the mathematics teacher. If there were mathematics data but no teacher-level data, then C7CPTM0 is zero, and such a case would not be included in the analysis. See section 4.8.1 for how the child-parent-teacher weights were defined.

4.8.5.5 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 (C7CPTS0)

The weight C7CPTS0 is to be used for analysis involving children who were subsampled to have a science teacher questionnaire and who had child assessment, 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 or student questionnaire data for spring-eighth grade (or was excluded from direct assessment due to a disability), whose parent completed the FSQ section of the parent interview for spring-eighth grade, and who had completed teacher-level data from either the English teacher or the science teacher. If there were science data but no teacher-level data, then C7CPTS0 is zero, and such a case would not be included in the analysis. See section 4.8.1 for how the child-parent-teacher weights were defined.

4.8.6 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-eighth 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 C7CW0 are C7CW1 through C7CW90. The methods used to compute the replicate weights and how they are used to compute the sampling errors of the estimates are described in section 4.9.3.

4.8.7 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.

Table 4-17. Characteristics of the eighth-grade cross-sectional child-level weights: School year 2006–07

Sample	Number of cases	Mean	Standard deviation	CV (× 100)	Minimum	Maximum	Skewness	Kurtosis	Sum
C7CW0	9,358	421.44	546.25	129.62	2.19	5479.19	3.44	13.93	3,943,827
C7PW0	8,809	447.74	579.18	129.36	1.91	5626.11	3.52	15.23	3,944,166
C7CPTE0	8,294	475.44	631.93	132.91	2.42	7716.63	3.41	13.91	3,943,318
C7CPTM0	4,130	955.24	1,227.71	128.52	5.30	10,632.36	3.20	11.46	3,945,141
C7CPTS0	4,164	946.51	1,227.76	129.71	6.88	9919.15	3.16	10.87	3,941,257

SOURCE: U.S. Department of Education, National Center for Education Statistics, Early Childhood Longitudinal Study, Kindergarten Class of 1998–99 (ECLS-K), spring 2007.

The difference in the estimate of the population of children (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 children in previous rounds of data collection who became ineligible in eighth grade (due to death or leaving the country) and (2) the adjustment of the weights for the children of unknown eligibility.

4.9 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.

4.9.1 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):

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

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 4.10.2 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 self-representing (SR) strata and 38 non-self-representing (NSR) strata. Among the 38 NSR strata, 11 strata were identified as Durbin strata[22] 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

[22] For a description of the Durbin method, see the *ECLS-K Third Grade Methodology Report* (NCES 2005–018) (Tourangeau, Brick, Byrne, et al. 2004).

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 in which 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.

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 4.8.3.2), 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).

4.9.2 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 exhibit 4-3.

Exhibit 4-3. ECLS-K Taylor Series stratum and first-stage unit identifiers: School year 2006-07

Variable name	Description
C7TCWSTR	Sampling stratum—spring-eighth grade C-weights
C7TCWPSU	First-stage sampling unit within stratum—spring-eighth grade C-weights
C7TPWSTR	Sampling stratum—spring-eighth grade P-weights
C7TPWPSU	First-stage sampling unit within stratum—spring-eighth grade P-weights
C7CPTEST	Sampling stratum—spring-eighth grade CPTE-weights
C7CPTEPS	First-stage sampling unit within stratum—spring-eighth grade CPTE-weights
C7CPTMST	Sampling stratum—spring-eighth grade CPTM-weights
C7CPTMPS	First-stage sampling unit within stratum—spring-eighth grade CPTM-weights
C7CPTSST	Sampling stratum—spring-eighth grade CPTS-weights
C7CPTSPS	First-stage sampling unit within stratum—spring-eighth 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 2007.

4.9.3 Specifications for Computing Standard Errors

Specifications for computing standard errors (SEs) are given in table 4-18. 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-eighth 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, C7CW0 as the full sample weight, C7CW1 to C7CW90 as the replicate weights, and JK2 as the method of replication.

Table 4-18. Specifications for computing standard errors, spring-eighth grade: School year 2006–07

		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-eighth grade cross-sectional	C7CW0	CHILDID	C7CW1 – C7CW90	JK2	WR	C7TCWSTR C7TCWPSU	1.829
	C7PW0	CHILDID	C7PW1 – C7PW90	JK2	WR	C7TPWSTR C7TPWPSU	
	C7CPTE0	CHILDID	C7CPTE1 – C7CPTE90	JK2	WR	C7CPTEST C7CPTEPS	
	C7CPTM0	CHILDID	C7CPTM1-C7CPTM90	JK2	WR	C7CPTMST C7CPTMPS	
	C7CPTS0	CHILDID	C7CPTS1-C7CPTS90	JK2	WR	C7CPTSST C7CPTSPS	

[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 2007.

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 (C7CW0), the stratum variable (C7TCWSTR), and the PSU variable (C7TCWPSU) must be specified. The "with replacement" sample design option, WR, must also be specified if using SUDAAN.

The next to last column in table 4-19 gives the average root design effect (DEFT) that can be used to approximate the SEs for each type of analysis. For a discussion of the use of design effects, see section 4.10.1.

4.10 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.

4.10.1 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*.[23] 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 $\bar{x}_w$ is calculated as follows:

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

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

where p is the weighted estimate of proportion for the characteristic of interest and n is the number of cases in the sample.

[23] Common procedures in SAS, SPSS, and Stata assume simple random sampling. 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.

4.10.2 Median Design Effects for the ECLS-K

In the ECLS-K, a large number of data items were collected from children, 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 4-19 shows estimates, SEs, and design effects for 52 means and proportions that were selected from the ECLS-K eighth-grade child assessment, student questionnaire, parent interview, and child-level teacher questionnaires. It is from this set of selected items that median design effects were computed for subgroups and presented in table 4-20.

For each survey item, table 4-19 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 academic 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 various questionnaires. They include the different scale scores from the direct child assessment, Academic Rating Scale scores from the teachers, characteristics of the children as they reported themselves in the student questionnaires, characteristics of the parents, and characteristics of the students as reported by the parents and teachers. In general, the design effects are lower than in previous years. The median design effect is 3.1 (compared with 4.0 in fifth grade). This is due to a smaller sample size that clustered in a smaller

number of schools; there were fewer middle schools for children to attend when they moved up from elementary schools.

Table 4-19. ECLS-K standard errors and design effects by selected child and parent variables, for the full sample—child assessment, student questionnaire, parent interview, and child-level teacher questionnaire data: School year 2006–07

Survey item	Variable name	Number of cases	Estimate	Design SE[1]	SRS SE[2]	DEFT[3]	DEFF[4]
Scores (mean)							
Reading scale score	C7R4RSCL	9,225	166.51	0.773	0.304	2.539	6.449
Mathematics scale score	C7R4MSCL	9,285	138.70	0.575	0.244	2.352	5.532
Science scale score	C7R2SSCL	9,304	82.24	0.406	0.179	2.270	5.152
Math score by teacher	T7ARSMAT	4,430	2.98	0.025	0.014	1.758	3.089
Oral score by teacher	T7ARSORL	8,908	3.18	0.020	0.010	1.956	3.825
Science score by teacher	T7ARSSCI	4,416	2.90	0.024	0.016	1.534	2.353
Writing score by teacher	T7ARSWRT	8,900	2.92	0.020	0.011	1.797	3.228
Characteristics from student questionnaire (percent)							
Participated in school sports	C7SPORTS	9,212	58.16	1.004	0.514	1.954	3.818
Described as overweight/slightly overweight	C7DESCWT	9,132	29.04	0.810	0.475	1.705	2.906
Tried to change weight	C7TRYWT	9,121	41.52	0.785	0.516	1.522	2.317
Home alone at least once a week	C7HOME	9,187	51.99	0.875	0.521	1.679	2.820
Angry when had trouble learning	C7ANGRY	9,226	79.44	0.685	0.421	1.628	2.649
Liked reading	C7LIKRD	9,186	77.62	0.814	0.435	1.872	3.506
Often felt lonely	C7LONLY	9,166	32.54	0.643	0.490	1.313	1.725
Felt good about self	C7FLGOOD	9,221	93.94	0.455	0.248	1.831	3.354
Parents helped with school work	C7SCHLPA	9,151	57.27	0.941	0.517	1.820	3.313
Parents advised on important decisions	C7ADVIPA	9,166	71.25	0.791	0.473	1.673	2.799
Characteristics from parent interview (percent)							
Lived in single parent family	P7HFAMIL	8,809	26.13	0.875	0.468	1.870	3.496
Lived in two-parent family	P7HFAMIL	8,809	71.01	0.957	0.484	1.979	3.915
Mom worked 35 hours+/week	P7HMEMP	6,765	68.39	1.171	0.565	2.072	4.293
Parents had high school or less	W8PARED	8,809	28.61	0.957	0.482	1.987	3.948
Household income	W8INCCAT	8,809	50.46	1.163	0.533	2.184	4.768
Parent attended PTA	P7ATTENP	6,012	33.38	1.074	0.608	1.765	3.116
Had family TV rule	P7TVRULE	8,679	87.74	0.630	0.352	1.789	3.200
Have someone help with reading homework	P7HELPR	8,531	94.61	0.377	0.245	1.540	2.373
Talk to child about day at school everyday	P7OFTTLK	8,688	78.11	0.756	0.444	1.704	2.902
Talk to child about smoking 3+ times a year	P7TLKSMK	8,679	76.87	0.794	0.453	1.754	3.076
Talk to child about alcohol 3+ times a year	P7TLKALC	8,681	76.47	0.875	0.455	1.921	3.691
Took away privilege when child angry	P7HITPRV	8,646	87.24	0.636	0.359	1.773	3.145
Self-reported in very good health	P7HEALTH	8,491	86.74	0.650	0.368	1.767	3.123
Received food stamps in last 12 months	P7FSTAMP	8,590	15.43	1.136	0.390	2.914	8.492

See notes at end of table.

Table 4-19. ECLS-K standard errors and design effects by selected child and parent variables, for the full sample—child assessment, student questionnaire, parent interview, and child-level teacher questionnaire data: School year 2006–07—Continued

Survey item	Variable name	Number of cases	Estimate	Design SE[1]	SRS SE[2]	DEFT[3]	DEFF[4]
Characteristics from teacher questionnaire (percent)							
Child in eighth grade	T7GLVL	9,358	85.66	0.822	0.362	2.268	5.142
Worked hard for grades-English	G7WRKHRD	8,921	70.04	0.861	0.485	1.775	3.149
Attentive in class-English	G7ATTENT	8,923	73.45	0.736	0.468	1.574	2.479
Was able to organize thoughts-English	G7ORGANZ	8,879	67.37	0.825	0.497	1.659	2.752
Worked hard for grades-Math	M7WRKHRD	3,994	72.42	1.122	0.707	1.586	2.515
Attentive in class-Math	M7ATTENT	3,984	73.84	1.100	0.697	1.579	2.493
Worked hard for grades-Science	N7WRKHRD	4,011	71.07	0.972	0.716	1.358	1.843
Attentive in class-Science	N7ATTENT	4,000	75.22	0.798	0.683	1.169	1.367
Other characteristics (mean)							
Age of child in months	R7AGE	9,351	171.53	0.104	0.049	2.115	4.474
Child's BMI	C7BMI	8,829	23.14	0.098	0.063	1.550	2.403
Hours spent in school activities	C7HRSCLB	8,976	4.80	0.112	0.071	1.586	2.515
Hours spent on non-school reading	C7HRSRD	8,938	3.87	0.176	0.087	2.013	4.052
Hours spent watching TV on weekdays	C7TVWKDY	9,128	3.19	0.049	0.033	1.487	2.211
Hours spent watching TV on weekend	C7TVWKEN	9,101	4.67	0.072	0.043	1.659	2.753
Hours spent playing videogames on weekdays	C7VIDWKD	9,116	1.52	0.042	0.025	1.689	2.854
Hours spent playing videogames on weekend	C7VIDWKN	9,137	2.75	0.077	0.038	2.007	4.027
Hours spent on the internet on weekdays	C7INTWKD	9,060	2.16	0.043	0.025	1.700	2.889
Hours spent on the internet on weekend	C7INTWKN	9,065	3.01	0.061	0.035	1.751	3.067
Child's household size	P7HTOTAL	8,809	4.49	0.030	0.015	2.048	4.193
Number of children <18 in child's HH	P7LESS18	8,809	2.41	0.028	0.013	2.226	4.953
Number of siblings in HH	P7NUMSIB	8,809	1.53	0.026	0.013	2.071	4.288
Median						1.770	3.134
Mean						1.829	3.438
Standard deviation						0.310	1.232
Coefficient of variation						0.170	0.358
Minimum						1.169	1.367
Maximum						2.914	8.492

[1] Design SE is the standard error under the ECLS-K sample design. For an explanation of this statistic, see section 4.10.
[2] SRS SE is the standard error assuming simple random sample. For an explanation of this statistic, see section 4.10.
[3] DEFT is the root design effect. For an explanation of DEFT, see section 4.10.
[4] DEFF is the design effect. For an explanation of DEFF, see section 4.10.
SOURCE: U.S. Department of Education, National Center for Education Statistics, Early Childhood Longitudinal Study, Kindergarten Class of 1998–99 (ECLS-K), spring 2007.

Table 4-20 presents the median design effects from the same survey items for subgroups based on school affiliation, child's sex and race/ethnicity, geographic region, level of urbanicity, and the socioeconomic scale quintile of the parents. Design effects are highest for children in the Midwest and lowest for American Indians. American Indians are the smallest group of children, and they are highly clustered.

Table 4-20. ECLS-K median design effects for subgroups: School year 2006–07

Subgroups	Spring-eighth grade	
	DEFT[1]	DEFF[2]
All children	1.770	3.134
School affiliation[3]		
Public	1.780	3.168
Private	1.859	3.456
Catholic private	1.943	3.776
Other private	1.680	2.820
Sex		
Male	1.680	2.824
Female	1.717	2.946
Race/ethnicity		
White	1.805	3.256
Black	1.515	2.294
Hispanic	1.402	1.965
Asian	1.424	2.027
Pacific Islander	1.337	1.787
American Indian	1.158	1.387
Other	1.600	2.561
Region		
Northeast	1.743	3.040
Midwest	2.036	4.147
South	1.767	3.122
West	1.667	2.779
Urbanicity		
Central city	1.759	3.093
Urban fringe and large town	1.711	2.929
Small town and rural area	1.927	3.711
Socioeconomic quintile		
First (lowest)	1.520	2.309
Second	1.651	2.725
Third	1.627	2.646
Fourth	1.809	3.270
Fifth (highest)	1.637	2.679

[1] DEFT is the root design effect. For an explanation of DEFT, see section 4.10.

[2] DEFF is the design effect. For an explanation of DEFF, see section 4.10.

[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 eighth 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 2007.

In spring-eighth grade, as in first, third, and fifth 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.

This page is intentionally left blank.

5. DATA COLLECTION METHODS AND RESPONSE RATES

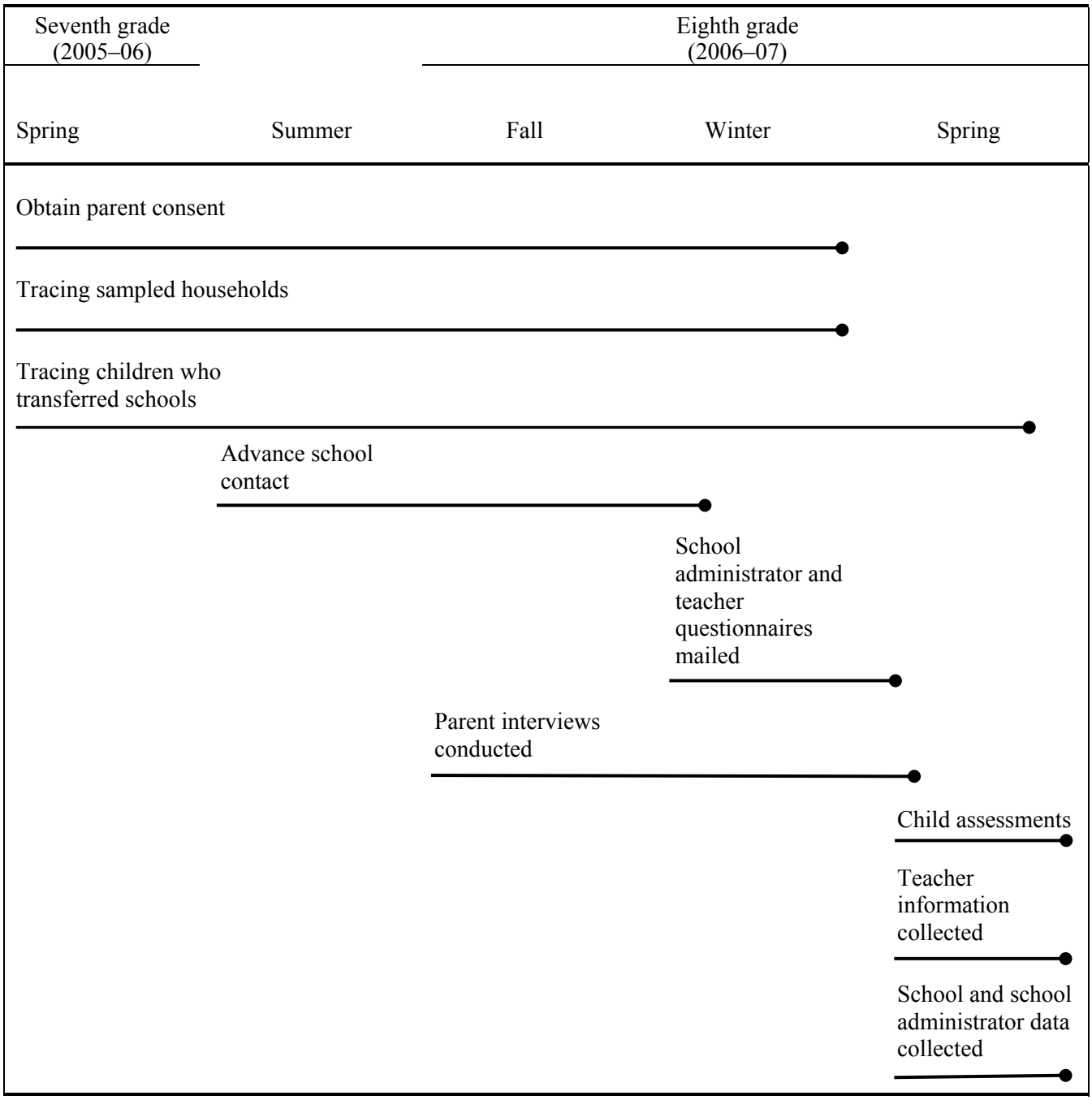
The following sections discuss the data collection procedures and response rates in the eighth-grade data collection phase of the Early Childhood Longitudinal Study, Kindergarten Class of 1998–99 (ECLS-K). Section 5.1 gives an overview of the data collection methods. Detailed information is provided on study training procedures (section 5.2); spring, summer, and fall 2006 data collection activities (section 5.3); tracing activities (section 5.4); spring-eighth grade (spring 2007) data collection (section 5.5); and quality control procedures (section 5.6). Spring-eighth grade completion rates are presented and discussed in section 5.7.

5.1 Overview of Data Collection Methods

The ECLS-K eighth-grade data collection activities began in spring 2006 and continued through spring 2007. Spring 2006 data collection was conducted to obtain consent from parents of sampled children for continued participation in the ECLS-K study and to identify the school their child attended. Fall data collection included conducting parent interviews, obtaining parent consent for outstanding cases, and recruiting schools. Schools were contacted to set appointments to conduct the child assessments in the spring of the 2006–07 school year, link children to teachers, identify children who had withdrawn from the school, and obtain locating information about their new schools. Spring data collection included the direct child assessments, and collection of student, teacher, and school questionnaires. Activities to locate children and confirm or obtain the name of the school in which they were enrolled continued throughout the entire data collection period. The content and timeline of the eighth-grade data collection are shown in exhibit 5-1.

The modes of data collection for obtaining consent and conducting the parent interview was telephone and in-person computer-assisted interviewing (CAI) and mailed, hard-copy consent forms; the child assessments were timed and group-administered using hard-copy assessment booklets; self-administered questionnaires were used to gather information from teachers, school administrators, and children.

Exhibit 5-1. Timeline of eighth-grade data collection: 2006–07



SOURCE: U.S. Department of Education, National Center for Education Statistics, Early Childhood Longitudinal Study, Kindergarten Class of 1998-99, spring 2007.

5.2 Field Staff Training

Several in-person training sessions were conducted to prepare staff for the eighth-grade data collection. In spring 2006, field supervisors and interviewers were trained to contact parents to obtain consent and to identify the school their child would attend in the 2006-2007 school year. In fall 2006, two trainings were held: one to train supervisors and interviewers to conduct the parent interview and one to train supervisors to contact original schools and recruit transfer schools. In spring 2007, two trainings on the administration of the direct child assessments were held: one for field supervisors and one for test administrators. The following sections discuss each specific type of training.

5.2.1 Obtaining Parent Consent Training

Field supervisors and interviewers were trained on obtaining parent consent in May 2006. Prior to the May in-person training session, supervisors and interviewers completed 16 hours of home study training that included reading materials and written exercises on the study design and field procedures as well as extensive individual and role-play practice in refusal aversion techniques to better answer respondent questions and address respondent concerns. The home study practice included role plays on answering respondent concerns and questions over the telephone with another interviewer as well as with a field supervisor.

Field supervisor training. The topics covered in the field supervisor training included debriefing interviewers on the home study exercises that supervisors completed with interviewers, principles of supervision, establishing and monitoring production goals, field management issues, using the automated Field Management System (FMS), and administrative issues.

The FMS was used throughout all phases of data collection to enter information about the sampled children, parents, teachers, and schools and to monitor production on all data collection activities. Field supervisors entered information into the FMS during training presentations, thus acquiring hands-on experience with the FMS and all field procedures prior to data collection. Field supervisor training for the parent consent phase of the study preceded the interviewer training and lasted for one day. Seven field supervisors completed training.

Interviewer training. The topics covered included an overview of study activities to date, a review of the parent folder that included parent contact information, an introductory script for obtaining consent, the CAI parent consent recording application, interactive lectures and role plays on answering respondent's questions or concerns about the study, and the procedures for recording parents' spoken consent on the telephone. A major goal of this training was to train interviewers to be able to respond immediately, directly, and in a fluid and natural way to respondent concerns in order to build consent response rates. The obtaining parent consent training was 1½ days long. A total of 113 interviewers completed training.

5.2.2 Parent Interview Training

Field supervisors and interviewers were trained on conducting the parent interview in August 2006. Prior to the August in-person training session, supervisors and interviewers completed 4 hours of home study training that included reading materials on basic features of the parent interview, CAI, and general interviewing techniques as well as written exercises on the procedures for conducting the parent interview.

Field supervisor training. The field supervisor training preceded the interviewer training and lasted for a half-day. The same seven field supervisors who managed the interviewers who obtained parent consent continued to manage interviewers as they conducted parent interviews. The supervisor training included establishing and monitoring production goals, field management issues, and using the FMS to organize and track production.

Interviewer training. The training sessions included an overview of the content of the parent interview and all of its sections and all procedures associated with conducting the interview. Interviewers practiced using the CAI system on laptop computers during interactive lectures and role plays. Interviewer training was 1½ days long. A total of 108 interviewers completed parent interview training. Ninety-two of the 108 interviewers (85 percent) were continuing from the training on obtaining parent consent. Sixteen interviewers were new hires to fill staffing needs as a result of staff attrition and were trained on obtaining parent consent by their supervisors, outside of the in-person training session. Fifteen of the 108 interviewers (14 percent) were certified as Spanish bilingual interviewers and attended a half-day bilingual training after the parent interview training ended. The bilingual training consisted of interactive lectures and role-plays on conducting the parent interview in Spanish.

5.2.3 Advance School Contact and Recruitment Training

Field supervisors were trained for 2½ days in August 2006 to contact original sampled schools and transfer schools to set up the data collection in the spring. A total of 63 field supervisors and three field managers completed training. Topics included an overview of study activities to date, a review of parent consent procedures, identifying and locating children who had moved from the schools they attended in the fifth grade, identifying the teachers of ECLS-K children and linking them to those children, and exercises on scheduling schools efficiently within an assignment. Prior to in-person training, field supervisors completed 8 hours of home study training that included watching a DVD called "Tips from Experienced Recruiters," reading materials, written exercises, and active practice answering respondent questions and addressing concerns in both written exercises and role-plays with a colleague.

As in the fifth-grade training, advance contact and recruitment training were conducted using the FMS. As noted earlier, the FMS was used during all phases of data collection to enter information about the sampled children, 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, in addition to completing role plays and exercises that involved entering information into the FMS.

5.2.4 Spring-Eighth Grade Direct Child Assessment Training

Field supervisors and test administrators were trained for the spring-eighth grade data collection in March 2007.

Field supervisor training. Field supervisor training preceded the test administrator training and lasted for one day. The topics covered in the field supervisor training session included an overview of study activities to date, a review of assignments, and interactive lectures on labeling and shipping school and teacher questionnaires to newly identified schools and teachers. As in earlier trainings, field supervisors were trained to use the FMS, and they practiced entering information into the FMS during training presentations. Twelve field supervisors completed training.

Test administrator training. The test administrator training sessions included an overview of study activities to date, interactive lectures based on the child assessments, practice scoring the child

assessment routing forms, reviewing materials from the fall school recruitment, role plays to practice contacting school coordinators, identifying and locating children who had moved from their eighth-grade schools identified in the fall, identifying the regular and special education teachers of ECLS-K children and linking them to those children, and distributing and following up on teacher questionnaires and school administrator questionnaires. A major goal of the test administrator training was to train field staff to properly conduct the assessments. This included reading the script word for word, correctly scoring the assessment routing forms, and identifying the appropriate second-stage form and labeling it correctly. Test administrators had multiple sessions to practice scoring the assessment routing forms, and identifying and labeling the second-stage form. The sessions provided trainees with hands-on experience with all the child assessment materials and procedures prior to data collection. Trainees practiced entering information into the FMS on laptop computers during training presentations. Test administrator training lasted 2 days. Field supervisors were also trained to perform all test administrator activities. A total of 217 test administrators and 12 field supervisors completed training.

5.3 2006 Eighth-Grade Data Collection Activities

Data collection activities in 2006 included obtaining parents' consent for their children to continue participating in the study and the schools they would attend, tracing households with outdated address information, conducting the parent interview, and contacting schools to recruit them into the study and arrange the spring data collection. The following sections discuss each of these data collection activities.

5.3.1 Obtaining Parent Consent

In mid-April 2006, advance packages were mailed to the 11,924 households eligible to participate in this round of the study. The package included a letter to the parents on ECLS-K stationery, a parent consent form that asked permission for continued participation in the study and asked the parent to confirm or provide school contact information for the school their sampled child would be attending in the upcoming school year (2006–07), and a parent newsletter with study results from elementary school years. Three weeks after mailing the parent advance package, a reminder postcard was mailed to all parents. By the second week in May, hard-copy consent forms had been received from 36 percent (4,265) of the eligible households.

Beginning the second week in May and continuing through the end of December, interviewers telephoned all parents who had not responded to the advance mailing, obtained parent consent, and confirmed or updated school contact information. During this data collection period, parent consent was obtained either by the parent signing and returning the consent form or by recording spoken consent on the interviewer's laptop. Spoken consent was obtained by reading the permission form to the parent and asking her for consent to record her response to the request. If the parent agreed to give spoken consent, the interviewer read a statement from her laptop that identified the parent and child and stated that the parent had given permission to record her spoken consent. All consent recordings were verified by home office staff who listened to the recordings and, when verified, generated a hard-copy parent consent form with a proxy auto-signature of the verifier. For those parents from whom consent was not received and who did not have a telephone, in-person visits to the home were made to obtain their consent. By the end of December 2006, consent had been obtained from approximately 83 percent (9,835) of eligible households.

5.3.2 Conducting the Parent Interview

Parent interview procedures mirrored those of previous rounds of data collection. The parent interview was conducted in the fall and winter of 2006 in order to first obtain parent consent and school information for the sampled child for any outstanding cases.

The parent interview was administered, primarily as a CAI telephone interview, from September 2006 through January 2007. For cases with parent consent still needed, interviewers attempted to obtain consent and complete a parent interview during the same call. Slightly over 34 percent of the parent interviews were completed in September, 34 percent in October, 18 percent in November, and over 6 percent in December and January. The parent interview averaged 45 minutes. As in previous rounds of data collection, the parent interview was conducted in person if the respondent did not have a telephone. Table 5-1 presents the number of parent interviews completed by mode and language. In eighth grade, slightly over 2 percent of all completed parent interviews were conducted in person; 9 percent of all completed parent interviews were conducted in a language other than English; and 89.4 percent of the latter were conducted in Spanish.

Table 5-1. Number and percent of completed parent interviews by data collection mode and language: School year 2006–07

	Spring-eighth grade	
Parent interviews	Number	Percent
Total interviews	8,809	100.0
Complete	8,610	97.7
Partial	199	2.3
Mode of data collection		
In person	193	2.2
By telephone	8,417	95.6
Mode unknown	199	2.3
Language of parent interview		
English	7,827	88.9
Spanish	701	8.0
Other language	82	0.9
Language unknown	199	2.3

NOTE: Cases where mode and language of parent interview are unknown are cases that did not complete the parent interview. Since the mode and language of parent interview is the last question of the parent interview, cases that terminate early do not have these data recorded.
SOURCE: U.S. Department of Education, National Center for Education Statistics, Early Childhood Longitudinal Study, Kindergarten Class of 1998–99 (ECLS-K), spring 2007.

5.3.3 Fall Preassessment Contact

Beginning in September 2006, all schools confirmed or identified by parents while obtaining consent 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.

Advance mailings. In September 2006, an advance package was mailed via Federal Express to all identified 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 previous rounds and an overview of eighth-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.

Preassessment contact. The fall preassessment contact was made by telephone between September and December 2006. 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 parent-reported schools and entered eighth-grade transfer schools. Schools were determined to be ineligible for eighth-grade data collection if no ECLS-K sampled children were currently enrolled. More original 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 English teacher and mathematics or science teacher, and special education teacher, obtain information on special accommodations[24] during assessment for the enrolled sampled child, and answer any questions that the school coordinator might have about the study.

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 supervisor asked the school coordinator to provide the names, addresses, and telephone numbers of these transfer schools. Field supervisors entered this information into the FMS and the updated information was distributed to parent interviewers if the parent interview was not completed. Parent interviewers also contacted field supervisors when they were unable to locate a sampled child's parent/guardian after having exhausted all leads and asked the supervisor for any leads they may have received during the school recruitment phase. All children who transferred were followed to their new school and not subsampled as in previous years. (Refer to the *ECLS-K Fifth-Grade Methodology Report* (NCES 2006-037) (Tourangeau, Lê, and Nord 2005) for additional details on how transfer children were subsampled in prior rounds.) 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 sampled children still enrolled in the school, including the child's current grade; the name and classroom for the child's English teacher and mathematics or science teacher; and whether or not the child had an Individualized Education Plan (IEP).

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

If the child had an IEP, then the name and classroom of the child's special education teacher was collected, along with whether the child required any accommodations to participate in the direct cognitive assessment. The accommodations in the eighth-grade direct cognitive assessment included all of those for the kindergarten, first-grade, third-grade, and fifth-grade direct cognitive assessments, with the addition of large print. Field supervisors contacted the teachers of the ECLS-K children as necessary for any of this information.

Contacting families of homeschooled children. As part of obtaining parent consent, the status of homeschooled children who were identified in rounds 1 through 6 was confirmed with their parents and updated as necessary. As parents of these children were contacted to obtain consent, they were asked to confirm that the child was still homeschooled or if the child 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 fifth grade, 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. The design of the eighth-grade child-level teacher questionnaire followed this model although English teachers rather that reading teachers were contacted. In elementary schools, children were primarily taught in intact classrooms, and teachers only reported classroom level information once for the classroom. Due to the design change in fifth grade, the teacher-child links were broadened to include the domain (reading, mathematics, or science) as well as information to identify the English, mathematics, or science classroom. In order to reduce data collection burden for teachers who were linked to multiple sampled 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 that 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 hard-copy teacher questionnaires (see section 5.5.3 for more information on teacher questionnaire data collection). 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 Eighth-Grade Data Collection

In order to ensure that as many of the sampled children as possible were contacted for eighth-grade data collection, tracing activities were ongoing through all phases of data collection. Tracing began in April of 2006 when the parent consent packages were mailed and continued through the spring data collection. If the parent advance package was returned as undeliverable but had new address information, it was remailed to the parent at the new address and the updated address was added the ECLS-K tracking database. If the package was returned as undeliverable with no updated address information, this information was entered into the tracking database and appeared on the parent locating form generated for each case. Interviewers used the parent locating form to attempt to obtain updated telephone numbers and addresses while prompting for consent and conducting the parent interview. Locating efforts included calling all contacts identified on the locating form, using directory assistance and the Internet resources, and in person-visits to the last known address of the case to attempt to collect updated address information from neighbors.

5.5 Spring-Eighth 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-eighth 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 eighth-grade sample for whom there were neither third-grade nor fifth-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 eighth grade or had been retained.

Test administrators received school assignments with a set of schools in or around a particular geographic area. An average assignment consisted of 13 schools. Each test administrator was responsible for all data collection activities in his or her assignment; they conducted the direct child assessments and collected all school and teacher questionnaires. A majority of the field staff hired for eighth-grade assessments were continuing from fall school recruiting or had worked on previous rounds

of ECLS-K data collection. Any staff hired with no prior experience on the study had experience on the National Assessment of Educational Progress (NAEP) in conducting group assessments.

5.5.1 Preassessment School Contact

Based on the information collected in the fall of 2006, packets of hard-copy teacher and school administrator questionnaires and instructions were assembled and mailed to schools beginning in January 2006, along with letters confirming the scheduled visits to the school. Teachers and school administrators were asked to complete the questionnaires and turn them in to the school coordinator for pickup by test administrators on assessment day.

Test administrators conducted preassessment activities by telephone starting in March 2007. The preassessment activities for these schools were similar to those conducted in previous rounds of data collection and included confirming the assessment date, the school's 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 March through early June 2007, the same time of year as in prior spring data collections. About 81 percent of the assessments were completed in March and April, about 18 percent were completed in May, and less than one percent were completed in June. In year-round schools, multiple assessment visits to the school were done, as needed, to assess all of the sampled children in each track.

The direct child assessments were usually conducted in a school classroom or library. Before conducting the assessments, test administrators set up the room for the assessments. The test administrator followed procedures for meeting the child(ren) at the test area as agreed upon during the preassessment contact with the school. 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. As noted earlier, the eighth-grade direct child assessments for reading, mathematics, and science were timed, two-stage, group-administered assessments. Test administrators read from a script for each

component of the assessment. The assessment routing forms were administered first in the following order: reading, mathematics, and science, and were timed for a total of 29 minutes. While the test administrators scored the assessment routing forms and identified and labeled the appropriate second-stage form for each domain, children were given 20 minutes to complete the student questionnaire. The second-stage assessments were administered in the following order: reading, mathematics and science, and were timed for a total of 51 minutes. The assessment session also included measurements of the sampled children's height and weight. The total time to complete all activities in an assessment session averaged slightly less than 2 hours. Participating children received a $15 honorarium.

Table 5-2 displays the total number of completed child assessments during spring-eighth grade data collection. All of the assessments were completed in reading: 94.6 percent of assessments were completed with no accommodations required; 4.9 percent completed the assessment with some accommodation, and less than 0.5 percent were excluded from participating in the assessments.

Table 5-2. Completed child assessments, by accommodation, spring-eighth grade data collection: School year 2006–07

	Spring-eighth grade	
Characteristic	Number	Percent
Child assessments completed	9,358	100.0
No accommodation[1]	8,853	94.6
With accommodation	460	4.9
Excluded	45	0.5

[1]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). "Excluded" is a subcategory in this table because the assessment status of these children is a result of their IEP requirements and not due to a refusal or failure to locate.
NOTE: This table does not include children who were subsampled out in fall- and spring-first grade and spring-third grade. These numbers should not be used to estimate child mobility.
SOURCE: U.S. Department of Education, National Center for Education Statistics, Early Childhood Longitudinal Study, Kindergarten Class of 1998–99 (ECLS-K), spring 2007.

Accommodations and exclusions. Less than 1 percent of participating children in eighth grade 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, if their IEP prevented their participation in assessments, or they required an accommodation not offered by the ECLS-K assessments. Less than 5 percent of participating children required accommodations. Accommodations offered by the ECLS-K assessments in this round were as follows: alternative setting (e.g., special lighting, adaptive chair), scheduling, or timing; health care aide

present; the use of a personal assistive device, and large print. Table 5-3 presents the number of children excluded from or requiring an accommodation to the direct child assessment in the spring of eighth grade.

Table 5-3. Number of children excluded from or accommodated in the spring-eighth grade assessments: School year 2006–07

Category	Number of children
Exclusions	
Excluded for disability	41
Accommodation[1]	
Alternative setting accommodation	116
Scheduling/timing accommodation	150
Health care aide present	7
Personal assistive device	7
Large print	3

[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 2007.

5.5.3 Teacher and School Data Collection

Data were collected from school administrators, regular classroom teachers, and special education teachers from March through June 2007.

The school and teacher questionnaires were mailed to the school coordinators beginning in late January 2007 on a flow basis, depending on the school's scheduled assessment date. Using the teacher-child-domain-classroom linkage information collected in the fall, a packet of questionnaires was assembled for each English, mathematics, science, and special education teacher. The customized teacher questionnaire materials included a cover letter and a $25 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). A packet of materials was also assembled for the school administrator. 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 field supervisor gathered the relevant information during the spring preassessment call and mailed the packets.

All teachers, including special education/service providers, received $25 for completing child-level instruments for sampled children in their classrooms. Teachers completing questionnaires for more than 10 children in their classes received remunerations of up to $55. Over 97 percent of teachers had fewer than 10 ECLS-K children.

On assessment day, after collecting completed questionnaires, the test administrator (TA) scanned the questionnaires to ensure that there were no missing critical items. During the field period, the TAs followed up with the school coordinator by making an in-person visit to the school or prompting by telephone to review the status of the incomplete or missing questionnaires.

5.6 Data Collection Quality Control

Continuous quality assurance procedures were employed during all data collection activities, with a particular focus on the assessments. The procedures were incorporated throughout all stages of the study (i.e., during instrument development, in the staff training program, and through parent validations).

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 practiced multiple exercises on scoring the first stage of each assessment and affixing labels to the second stage of each assessment. When the fieldwork began, field supervisors made telephone calls to parents to validate the interview. The teacher and school questionnaire packages were reviewed for accuracy at 100 percent to ensure the correct questionnaires were sent to the schools for distribution and completion.

5.6.1 Quality Control on the Child Assessment

The mode of assessment administration changed in eighth-grade from a one-on-one, CAI-with-easels assessment administration to a group-administered, timed, hard-copy assessment. The hard-copy assessment was a two-stage assessment with a routing assessment for each of three domains, reading, mathematics, and science, and two levels of the second-stage assessment for each domain. TAs had to administer the routing assessment, score the three domains, and identify the appropriate second-stage assessment by domain and affix a label with a child's name and identification number. In the training session, TAs practiced this process multiple times to be able to quickly and accurately score and label assessment forms in the field. All trainees were proficient on the process after completing training.

TAs accuracy in identifying the appropriate assessment forms was examined during the field period by comparing the child's' routing test score and the assessment form the TA labeled for the child. TAs identified the appropriate second-stage assessment with over 99 percent accuracy for each assessment domain: 99.2 percent accuracy for the reading assessment; 99.3 percent accuracy for the science assessment; and 99.5 percent accuracy for the mathematics assessment.

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 interviewer was always selected for validation. Over the course of the field period, a running count of an interviewer's completed parent interviews was maintained, and each tenth completed parent interview was selected for validation, thus ensuring that 10 percent of each interviewer's cases were selected for validation. The parent validation was approximately 5 minutes long and was conducted by telephone. In spring-eighth grade, a total of 834 parent interviews were validated with 75.8 percent reporting the same answers as in the original interview. 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 full name, date of birth, and sex; and
- seven questions repeated from the parent interview.

Field supervisors noted if the validation check was completed with no changes, with "minor" changes, or with "major" changes. "Minor" changes include spelling of parent name, child's name, parent's address or telephone number, child's date of birth, or child's gender. "Major" changes include any changes to the question responses.

Table 5-4 shows the results of parent interview validations. Discrepancies between parents' responses during the original parent interview and those during the validation may reflect differences in respondent recall, respondent interpretation of the question, or actual change in the data, rather than a validation issue. Feedback from supervisors indicated that two validation items may reflect some of these differences, rather than true validation issues. As a result, the results for major changes may be overreported.

Table 5-4. Results of parent interview validations: School year 2006–07

Parent interview	September		October		November		December		January	
	Number	Percent	Number	Percent	Number	Percent	Number	Percent	Number	Percent
Total complete	2,727	100.0	5,621	100.1	7,610	100.0	7,973	100.0	8,681	100.0
Validation cases generated	272		593	10.5	763	10.0	807	10.1	898	10.3
Validation cases receipted	93	34.2	388	65.4	677	88.7	734	91.0	834	92.9
No changes	83	89.2	31	80.7	516	76.2	556	75.7	632	75.8
Minor changes	3		26	6.7	44	6.5	44	6.0	45	5.4
Major changes	7		49	12.6	117	17.3	134	18.3	157	18.8
Other (specify)	0		0	0.0	0	0.0	0	0.0	0	0.0
Cases pending	179	65.8	20	34.6	86	11.3	73	9.0	64	7.1

SOURCE: U.S. Department of Education, National Center for Education Statistics, Early Childhood Longitudinal Study, Kindergarten Class of 1998–99 (ECLS-K), spring 2007.

5.7 Spring-Eighth 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, eighth-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 eighth-grade data collection were computed with the same procedures used for spring-first grade, spring-third grade, and spring-fifth grade to allow for comparisons of completion rates for the four rounds 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: (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. For spring-eighth grade, only the 12,129 children who were still eligible after the fifth-grade data collection were fielded; they included both fifth grade respondents and eligible nonrespondents. Weighted completion rates were computed using the eighth-grade base weight (i.e., inverse of selection probabilities) adjusted for previous round movers, but not adjusted for nonresponse.

5.7.1 Children Sampled in Kindergarten

Tables 5-5 to 5-7 present weighted and unweighted child-level completion rates for spring-eighth grade data collection, broken out by school characteristics.[25] 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 eighth-grade data collection.

In general, completion rates for eighth grade are lower than in previous year. Even though hard-to-field cases[26] from the fifth-grade collection were excluded, the completion rates are lower for three main reasons: (1) the eighth-grade data collection occurred three years after the fifth-grade data collection, making it harder to find respondents, (2) the children were older and could refuse to cooperate at a much higher rate than younger children, and (3) the change in the field procedure in which explicit parent consent had to be obtained before the children could be approached.

Table 5-5 shows that the completion rates for the child assessment are higher in public schools than in private schools. Within the private school category, the difference in the rates is not as large. Excluding the "unknown" category, the complete for the child assessment rates range from 82.7 percent for children in non-Catholic private schools to 97.1 percent for children in schools in small towns. The pattern of completion rates is similar or the parent interviews, ranging from 76.5 percent for children in non-Catholic private schools to 89.2 percent for children in schools in large towns, excluding the "unknown" category. The "unknown" category includes children who were unlocatable as their whereabouts were unknown. The category "unknown" also includes 48 children who were homeschooled and thus had no information concerning schools.

Table 5-6 shows that the overall weighted completion rates are 75.3 for the student questionnaire, 73.3 percent for the school administrator questionnaire, and 74.5 for the teacher-level questionnaire. Excluding the "unknown" category, the completion rates for the student questionnaire follow the same pattern of the rates for child assessment with the lowest rate for children in non-Catholic schools (82.0 percent) to the highest rate for children who were not in schools in cities or their surrounding areas (in this case 96.2 percent in the rural area outside the Metropolitan Statistical Areas).

[25] 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 eighth grade with those in previous years.

[26] 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.

The pattern of completion rates for the school and teacher instruments is somewhat different. For the school administrator questionnaire, the rates range from 80.4 percent for schools with the highest minority enrollment to 97.0 percent for schools with the lowest minority enrollment. This is a phenomenon observed in previous rounds for the school administrator questionnaire.

Table 5-7 shows that the rates for the child-level teacher questionnaires. All three of these subject-specific teacher questionnaires were completed at an overall rate of 72 or 73 percent. Excluding the "unknown" category, the completion rates for the child-level teacher questionnaires are as follows: 80.4 percent (large city) to 97.3 percent (small town) for English; 77.2 percent (large city) to 97.3 percent (small town) for mathematics, and 79.2 percent (non-Catholic private or large city) to 93.7 percent (high total enrollment) for science. These rates are not as high as in fifth grade but higher than in third grade, most likely due to the higher incentives employed in fifth grade and carried on to eighth grade.

Table 5-5. 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 2006–07

School characteristic[1]	Child assessment			Parent interview		
	Completes[2]	Completion rate		Completes[3]	Completion rate	
		Weighted	Unweighted		Weighted	Unweighted
All schools	9,296	75.7	77.9	8,755	71.7	73.4
School affiliation						
Public	7,662	93.6	93.9	6,968	85.2	85.4
Private	1,576	83.9	85.6	1,483	77.8	80.5
Catholic	963	85.0	84.4	911	78.9	79.8
Other private	613	82.7	87.4	572	76.5	81.6
Unknown	58	3.1	3.0	304	16.8	15.8
Type of locale						
Large city	1,250	88.0	87.4	1,105	78.4	77.3
Mid-size city	1,434	93.5	92.0	1,327	85.9	85.1
Urban fringe of large city	2,291	89.3	90.1	2,106	82.9	82.8
Urban fringe of mid-size city	938	94.6	95.0	869	84.2	88.0
Large town	212	95.5	95.5	201	89.2	90.5
Small town	903	97.1	96.6	812	88.3	86.8
Rural—outside MSA	1,040	95.0	94.4	966	86.6	87.7
Rural—inside MSA	984	94.8	95.6	905	87.3	87.9
Unknown	244	9.6	11.5	464	21.3	21.9
School size (total enrollment)						
1 to 299	1,470	89.8	88.9	1,377	83.1	83.3
300 to 499	1,816	90.8	89.9	1,661	82.3	82.3
500 to 749	2,326	92.6	92.5	2,134	84.3	84.9
750 or more	1,883	95.4	96.5	1,708	86.5	87.5
Unknown	1,801	44.3	47.5	1,875	48.2	49.5

See notes at end of table.

Table 5-5. 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 2006–07—Continued

School characteristic[1]	Child assessment			Parent interview		
	Completes[2]	Completion rate		Completes[3]	Completion rate	
		Weighted	Unweighted		Weighted	Unweighted
Percent non-White enrolled						
0–10	2,654	94.3	93.7	2,475	86.9	87.4
11–49	3,573	93.5	93.5	3,338	86.9	87.3
50–89	1,672	92.7	92.1	1,492	83.7	82.2
90–100	1,320	90.0	89.4	1,128	76.7	76.4
Unknown	77	3.5	3.9	322	17.2	16.2
Region						
Northeast	1,710	92.3	91.3	1,560	83.0	83.3
Midwest	2,590	93.6	93.1	2,443	87.7	87.8
South	3,022	91.6	92.4	2,734	82.4	83.6
West	1,941	93.1	92.2	1,734	85.1	82.3
Unknown	33	1.8	1.7	284	16.2	15.0

[1] School characteristics are for schools attended by children in the ECLS-K eighth-grade sample and are based on ECLS-K survey data, not data from the sampling frame.
[2] English, mathematics, or science assessment was scorable, or child was disabled and could not be assessed, or child had student questionnaire data or height and weight data.
[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 2007.

Table 5-6. Number of completed child-level cases and child-level completion rates for the student questionnaire, school administrator questionnaire, and teacher-level questionnaire for children sampled in the base year, by school characteristics: School year 2006–07

School characteristic[1]	Student questionnaire			School administrator questionnaire			Teacher-level questionnaire		
		Completion rate			Completion rate			Completion rate	
	Completes[2]	Weighted	Unweighted	Completes[2]	Weighted	Unweighted	Completes[2]	Weighted	Unweighted
All schools	9,244	75.3	77.5	9,200	73.3	77.0	9,147	74.5	77.0
School affiliation									
Public	7,617	93.2	93.3	7,434	90.0	90.9	7,560	92.1	92.6
Private	1,569	83.6	85.2	1,749	90.4	95.0	1,563	83.1	84.9
Catholic	961	84.9	84.2	1,086	92.9	95.2	960	84.7	84.1
Other private	608	82.0	86.7	663	87.6	94.6	603	81.3	86.0
Unknown	58	3.1	3.0	17	0.8	0.9	24	1.3	1.3
Type of locale									
Large city	1,241	87.5	86.8	1,258	81.4	87.1	1,194	83.0	83.5
Mid-size city	1,429	93.3	91.7	1,455	92.8	93.3	1,419	92.2	91.0
Urban fringe of large city	2,276	88.9	89.5	2,280	87.4	89.6	2,256	87.6	88.7
Urban fringe of mid-size city	937	94.5	94.9	940	93.6	95.2	932	93.6	94.4
Large town	211	95.2	95.0	216	96.3	97.3	212	95.5	95.5
Small town	895	96.2	95.7	854	92.6	91.3	904	97.9	96.7
Rural—outside MSA	1,031	94.1	93.6	1,030	93.6	93.5	1,042	94.9	94.6
Rural—inside MSA	980	94.7	95.2	959	91.7	93.2	978	93.6	95.0
Unknown	244	9.6	11.5	208	7.7	9.8	210	8.2	10.1
School size (total enrollment)									
1 to 299	1,463	89.4	88.5	1,586	93.7	95.9	1,456	88.4	88.0
300 to 499	1,805	90.3	89.4	1,864	89.5	91.7	1,798	89.5	89.1
500 to 749	2,312	92.2	92.0	2,298	91.1	91.4	2,301	91.7	91.5
750 or more	1,870	94.9	95.8	1,805	90.7	92.5	1,883	95.2	96.5
Unknown	1,794	44.2	47.3	1,647	40.2	43.5	1,709	42.2	45.7

See notes at end of table.

Table 5-6. Number of completed child-level cases and child-level completion rates for the student questionnaire, school administrator questionnaire, and the teacher-level questionnaire for children sampled in the base year, by school characteristics: School year 2006–07—Continued

School characteristic[1]	Student questionnaire			School administrator questionnaire			Teacher-level questionnaire		
		Completion rate			Completion rate			Completion rate	
	Completes[2]	Weighted	Unweighted	Completes[2]	Weighted	Unweighted	Completes[2]	Weighted	Unweighted
Percent non-White enrolled									
0–10	2,642	93.9	93.3	2,754	97.0	97.2	2,649	94.3	93.5
11–49	3,548	93.0	92.8	3,601	91.9	93.8	3,562	92.9	93.2
50–89	1,664	92.3	91.6	1,607	87.0	88.5	1,630	89.5	89.8
90–100	1,315	89.8	89.1	1,209	80.4	81.9	1,270	86.2	86.0
Unknown	75	3.5	3.8	29	1.2	1.5	36	1.6	1.9
Region									
Northeast	1,705	92.1	91.1	1,729	90.1	91.7	1,677	89.7	89.6
Midwest	2,569	93.0	92.3	2,677	94.8	96.2	2,583	93.1	92.8
South	3,004	91.2	91.8	2,935	87.9	89.7	2,995	90.4	91.6
West	1,933	92.9	91.8	1,859	88.0	88.3	1,892	90.9	89.8
Unknown	33	1.8	1.7	0	0.0	0.0	0	0.0	0.0

[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 2007.

Table 5-7. 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 2006–07

School characteristic[1]	Child-level English teacher questionnaire			Child-level mathematics teacher questionnaire			Child-level science teacher questionnaire		
	Completes[2]	Completion rate		Completes[2]	Completion rate		Completes[2]	Completion rate	
		Weighted	Unweighted		Weighted	Unweighted		Weighted	Unweighted
All schools	8,957	73.2	75.4	4,449	71.6	75.2	4,459	73.3	74.8
School affiliation									
Public	7,394	90.5	90.6	3,670	89.4	90.0	3,664	89.1	89.7
Private	1,539	81.5	83.6	769	82.6	84.0	781	82.1	84.3
Catholic	935	81.7	81.9	459	83.6	82.1	489	84.6	84.0
Other private	604	81.3	86.2	310	81.6	86.8	292	79.2	84.9
Unknown	24	1.3	1.3	10	0.9	1.1	14	1.8	1.5
Type of locale									
Large city	1,158	80.4	81.0	557	77.2	79.9	582	79.2	79.4
Mid-size city	1,391	90.8	89.2	710	90.1	88.9	685	91.6	90.1
Urban fringe of large city	2,228	86.8	87.6	1,097	84.2	86.0	1,105	86.8	87.2
Urban fringe of mid-size city	894	90.1	90.6	448	90.3	91.2	453	88.5	91.3
Large town	200	88.1	90.1	109	96.2	94.8	99	84.9	92.5
Small town	894	97.3	95.6	427	97.3	94.7	451	91.7	93.2
Rural—outside MSA	1,025	93.6	93.0	519	93.8	93.5	498	91.6	91.0
Rural—inside MSA	959	92.3	93.2	483	92.1	93.2	477	92.9	93.3
Unknown	208	8.1	10.1	99	6.9	9.7	109	9.4	10.4
School size (total enrollment)									
1 to 299	1,429	86.1	86.4	709	87.0	87.0	721	85.3	85.9
300 to 499	1,774	88.2	87.9	890	87.8	88.1	882	87.9	87.4
500 to 749	2,247	90.0	89.4	1,077	88.1	86.7	1,134	87.6	89.2
750 or more	1,848	94.1	94.7	944	93.7	95.4	903	93.7	93.8
Unknown	1,659	41.4	44.4	829	38.7	44.5	819	42.9	43.7

See notes at end of table.

Table 5-7. 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 2006–07—Continued

School characteristic[1]	Child-level English teacher questionnaire			Child-level mathematics teacher questionnaire			Child-level science teacher questionnaire		
	Completes[2]	Completion rate		Completes[2]	Completion rate		Completes[2]	Completion rate	
		Weighted	Unweighted		Weighted	Unweighted		Weighted	Unweighted
Percent non-White enrolled									
0–10	2,640	94.1	93.2	1,295	93.8	92.5	1,324	91.9	92.5
11–49	3,477	90.9	91.0	1,759	90.7	91.7	1,730	91.1	90.9
50–89	1,599	88.3	88.1	784	85.0	85.3	785	87.5	87.5
90–100	1,206	82.4	81.7	592	81.3	81.7	603	79.5	80.3
Unknown	35	1.6	1.8	19	1.7	2.0	17	1.6	1.7
Region									
Northeast	1,651	88.5	88.2	825	89.3	88.6	840	88.5	89.3
Midwest	2,539	91.6	91.3	1,233	90.3	90.9	1,308	92.9	91.8
South	2,956	89.5	90.4	1,486	88.9	90.0	1,430	86.9	88.3
West	1,811	87.3	86.0	905	85.1	85.0	881	85.0	84.6
Unknown	0	0.0	0.0	0	0.0	0.0	0	0.0	0.0

[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 2007.

Tables 5-8 to 5-10 show the completion rates by mover status. Unlike previous years in which only a subsample of movers was followed into their new schools, the eighth-grade data collection followed all movers. The number of movers is larger than the number of nonmovers as children left their elementary schools for middle schools. Because of these changes, the rates are no longer comparable to rates in earlier years. In earlier years, nonmovers responded at a higher rate than movers. This is not the case for eighth grade. Not only the number of nonmovers is much smaller, but they also responded at a lower rate, 73.4 percent compared with 81 percent for movers, in the case of the child assessment. Since all movers were followed and highly successfully located, the difference between the completion rates of located movers and unlocated movers was not as large as shown in previous years. Of those who moved, 97 percent were located. There are cases whose mover status was unknown. These are children whose parents refused consent for their children to be approached for data collection, and the whereabouts of the children were not traced. The parent interview completion rates are 67.8 percent for nonmovers and 76.6 percent for movers. The difference in the rates between located movers and all movers is minimal, again because almost all movers were successfully located. There is the peculiar case of a high completion rate of unlocated movers. Even though children could not be located for the child assessment, a parent interview was conducted by telephone, leading to the 91 percent response rate for this category. The same is true for the cases of children with unknown mover status; 43 cases had parent interviews that apparently did not have information about where their children went to school. The pattern of completion rates by mover status is the same for the student questionnaire and the teacher questionnaires. The school administrator questionnaire is the only one where the completion rate for nonmovers is higher than for movers, a 10 percent difference. This can be explained by the fact that movers were not always assessed in schools so that the school administrator questionnaire could be administered; schools where nonmovers attended had been in the sample for a long time and tend to cooperate more than schools that were new to the sample, had a lower level of commitment to the ECLS-K, and often refused to complete the school administrator questionnaire.

Table 5-8. 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 mover's status: School year 2006–07

Mover status	Child assessment			Parent interview		
		Completion rate			Completion rate	
	Completes[1]	Weighted	Unweighted	Completes[2]	Weighted	Unweighted
All children	9,296	75.7	77.9	8,755	71.7	73.4
Mover status						
Mover	7,868	81.0	88.4	7,385	76.6	83.0
Located	7,868	83.3	90.6	7,204	76.2	82.9
Not located	0	0.0	0.0	181	91.2	85.8
Nonmover	1,428	73.4	75.7	1,327	67.8	70.3
Unknown	0	0.0	0.0	43	3.7	3.8

[1] English, mathematics, or science assessment was scorable, or child was disabled and could not be assessed, or child had student questionnaire data or height and weight data.
[2] 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 2007.

Table 5-9. Number of completed child-level cases and child-level completion rates for the student questionnaire, school administrator questionnaire, and the teacher-level questionnaire for children sampled in the base year, by mover's status: School year 2006–07

Mover status	Student questionnaire			School administrator questionnaire			Teacher-level questionnaire		
		Completion rate			Completion rate			Completion rate	
	Completes[1]	Weighted	Unweighted	Completes[1]	Weighted	Unweighted	Completes[1]	Weighted	Unweighted
All children	9,244	75.3	77.5	9,200	73.3	77.0	9,147	74.5	77.0
Mover status									
Mover	7,824	80.7	87.9	7,498	77.1	84.3	7,719	79.8	87.3
Located	7,824	83.0	90.1	7,498	79.6	86.6	7,719	82.1	89.4
Not located	0	0.0	0.0	0	0.0	0.0	0	0.0	0.0
Nonmover	1,420	72.9	75.3	1,702	87.1	90.2	1,428	73.4	75.7

[1] 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 2007.

Table 5-10. 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 2006–07

Mover status	Child-level English teacher questionnaire			Child-level mathematics teacher questionnaire			Child-level science teacher questionnaire		
	Completes[1]	Completion rate		Completes[1]	Completion rate		Completes[1]	Completion rate	
		Weighted	Unweighted		Weighted	Unweighted		Weighted	Unweighted
All children	8,957	73.2	75.4	4,449	71.6	75.2	4,459	73.3	74.8
Mover status									
Mover	7,542	78.3	85.3	3,753	76.2	84.6	3,746	78.6	85.0
Located	7,542	80.6	87.4	3,753	78.5	86.6	3,746	80.8	87.2
Not located	0	0.0	0.0	0	0.0	0.0	0	0.0	0.0
Nonmover	1,415	72.7	75.0	696	72.3	74.4	713	72.6	74.9
Unknown	0	0.0	0.0	0	0.0	0.0	0	0.0	0.0

[1] 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 2007.

Table 5-11. 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 2006–07

Child characteristics[1]	Child assessment			Parent interview		
	Completes[2]	Completion rate		Completes[3]	Completion rate	
		Weighted	Unweighted		Weighted	Unweighted
All children	9,296	75.7	77.9	8,755	71.7	73.4
Sex						
Male	4,684	75.6	77.5	4,434	72.1	73.4
Female	4,612	75.7	78.3	4,321	71.2	73.4
Race/ethnicity						
White, non-Hispanic	5,719	80.9	83.9	5,480	78.0	80.4
Black, non-Hispanic	951	66.6	70.2	834	59.4	61.5
Hispanic	1,602	71.0	71.6	1,486	65.8	66.5
Asian	516	59.9	61.0	474	55.6	56.0
Pacific Islander	107	73.1	69.9	90	59.3	58.8
American Indian or Alaska Native	183	80.1	81.7	184	82.7	82.1
Other	210	71.9	74.2	198	69.8	70.0
Unknown	8	50.0	47.1	9	71.2	52.9
Year of birth						
1992	2,733	74.3	78.0	2,586	71.0	73.8
1993	6,513	76.3	77.9	6,122	72.0	73.2
Other/unknown	50	70.1	78.1	47	65.8	73.4

[1] Based on ECLS-K survey data and not on data from the sampling frame.
[2] English, mathematics, or science assessment was scorable, or child was disabled and could not be assessed, or child had student questionnaire data or height and weight data.
[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 2007.

Table 5-12. Number of completed child-level cases and child-level completion rates for the student questionnaire, school administrator questionnaire, and teacher-level questionnaire for children sampled in the base year, by child characteristics: School year 2006–07

Child characteristic[1]	Student questionnaire			School administrator questionnaire			Teacher-level questionnaire		
		Completion rate			Completion rate			Completion rate	
	Completes[2]	Weighted	Unweighted	Completes[2]	Weighted	Unweighted	Completes[2]	Weighted	Unweighted
All children	9,244	75.3	77.5	9,200	73.3	77.0	9,147	74.5	77.0
Sex									
Male	4,653	75.3	77.0	4,622	72.7	76.7	4,608	74.5	76.8
Female	4,591	75.4	78.0	4,578	74.0	77.6	4,539	74.6	77.3
Race/ethnicity									
White, non-Hispanic	5,684	80.5	83.4	5,798	80.5	85.2	5,673	80.5	83.6
Black, non-Hispanic	947	66.4	69.9	873	60.1	64.3	924	64.6	68.4
Hispanic	1,595	70.7	71.3	1,526	66.7	68.1	1,549	68.8	69.6
Asian	512	59.6	60.5	514	58.7	60.8	503	59.1	59.7
Pacific Islander	107	73.1	69.9	91	63.5	59.1	108	74.2	71.1
American Indian or Alaska Native	181	79.2	80.8	184	79.8	81.8	180	78.1	80.7
Other	210	71.9	74.2	208	69.3	73.2	204	70.3	72.3
Unknown	8	50.0	47.1	6	22.9	20.7	6	36.3	35.3
Year of birth									
1992	2,708	73.8	77.3	2,756	73.0	78.8	2,696	73.5	77.4
1993	6,490	76.1	77.6	6,393	73.5	76.4	6,404	75.1	76.9
Other/unknown	46	65.6	71.9	51	55.1	68.9	47	63.2	74.6

[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.
SOURCE: U.S. Department of Education, National Center for Education Statistics, Early Childhood Longitudinal Study, Kindergarten Class of 1998–99 (ECLS-K), spring 2007.

Table 5-13. Number of completed child-level cases and child-level completion rates for the teacher-level questionnaires for children sampled in the base year, by child's mover status: School year 2006–07

Child characteristic[1]	Child-level English teacher questionnaire			Child-level mathematics teacher questionnaire			Child-level science teacher questionnaire		
		Completion rate			Completion rate			Completion rate	
	Completes[2]	Weighted	Unweighted	Completes[2]	Weighted	Unweighted	Completes[2]	Weighted	Unweighted
All children	8,957	73.2	75.4	4,449	71.6	75.2	4,459	73.3	74.8
Sex									
Male	4,511	73.1	75.1	2,240	71.9	75.3	2,255	73.2	74.5
Female	4,446	73.3	75.7	2,209	71.4	75.1	2,204	73.3	75.2
Race/ethnicity									
White, non-Hispanic	5,600	79.6	82.6	2,787	78.7	82.4	2,792	79.5	82.1
Black, non-Hispanic	912	64.1	67.6	435	58.0	65.8	449	65.2	65.2
Hispanic	1,485	66.0	66.7	730	65.5	66.1	748	65.4	66.7
Asian	489	57.8	58.0	247	59.2	59.1	238	55.1	56.0
Pacific Islander	102	70.4	67.1	60	73.3	69.0	46	73.5	70.8
American Indian or Alaska Native	164	71.9	73.5	94	75.2	79.0	78	74.9	75.0
Other	199	69.2	70.6	93	64.2	66.9	105	69.7	73.4
Unknown	6	36.3	35.3	3	51.5	37.5	3	18.9	33.3
Year of birth									
1992	2,656	72.2	76.3	1,315	71.0	76.8	1,331	73.1	75.3
1993	6,256	73.8	75.1	3,114	72.0	74.6	3,105	73.4	74.7
Other/unknown	45	60.8	71.4	20	48.5	69.0	23	63.9	67.6

[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.
SOURCE: U.S. Department of Education, National Center for Education Statistics, Early Childhood Longitudinal Study, Kindergarten Class of 1998–99 (ECLS-K), spring 2007.

Tables 5-11 to 5-13 present child-level weighted and unweighted completion rates for the spring-eighth 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 within 2 percentage points, but for race and ethnicity they are more substantial. Table 5-11 shows that the child assessment completion rate was highest for Whites (80.9 percent) and lowest for Asians (59.9 percent), a reverse in the trend of earlier years. The low response rate for Asians persists for other instruments as well. The unweighted sample of Asians is about 8 percent, about the same proportion as in earlier years. Their moving pattern is the same as in previous years; their relative moving rate is about 50 percent higher compared with fifth grade, the same as their minority counterparts (Hispanics and American Indians[27]). Therefore, the drop in the completion rates cannot be attributed to a change in the sample. The highest completion rate is for White, uniform across all instruments. American Indians have a higher completion rate for the parent interview, but the sample size for this group is so small that it should not be compared with Whites.

In addition to the child assessment, parent interview, student questionnaire, school administrator questionnaire, and teacher questionnaires (for which completion rates have been summarized in the preceding paragraphs), data were also collected in eighth grade from special education teacher questionnaires for children who had special education teachers. Table 5-14 presents counts of completes and weighted and unweighted completion rates at the overall child level for the special education teacher questionnaires A and B. Although the number of special education teacher questionnaires is small, its completion rates are high, 93.9 percent for part A, which captures teacher information, and 94.7 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 5-14. Number of completed instruments and child-level completion rates for the special education teacher questionnaires for children sampled in the base year: School year 2006–07

Category	Completes	Completion rates	
		Weighted	Unweighted
Special education part A[1]	812	93.9	94.5
Special education part B[1]	820	94.7	95.5

[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 2007.

[27] American Indian includes Alaska Native.

5.7.2 Children Sampled in First Grade

In spring-first grade the child 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-eighth grade, unless they belonged in the fifth-grade excluded groups or they became ineligible after fifth grade. Nonresponse in the freshened child 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-eighth 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 5-15 presents weighted and unweighted completion rates for freshened children. The two components of the completion rates are presented separately in table 5-15. The overall completion rates (i.e., the third set of rates in the table) 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 77.6 percent. As part of the freshening process, schools were asked to prepare an alphabetic roster of children enrolled in first grade. These schools were also requested to identify which children did not attend kindergarten the previous year. Schools did not participate in the freshening process because they either refused or were unable to provide the requested information. Within the schools that agreed to freshen, the freshening completion rate is 99.2 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 77 percent. Note that the first component rate for spring-eighth grade is not identical to the first component rate for earlier grades because of the exclusion of children in special groups as explained in section 4.7.

Table 5-15. Number of completed child-level cases and child-level completion rates for children sampled in first grade: School year 2006–07

Category	Completes	Completion rate[1]	
		Weighted	Unweighted
First component (first-grade sample freshening)	5,384	77.0	85.9
School term[2]	5,405	77.6	86.2
Child term[3]	5,384	99.2	99.6
Second component (eighth-grade data collection)			
Child assessment[4]	62	60.9	63.9
Parent interview[5]	54	51.5	55.7
Student questionnaire[6]	62	60.9	63.9
School administrator questionnaire[6]	62	54.4	62.6
Teacher-level questionnaire[6]	63	60.1	64.9
English teacher questionnaire (child level)[6]	61	58.6	62.9
Mathematics teacher questionnaire (child level)[6]	33	56.8	66.0
Science teacher questionnaire (child level)[6]	27	54.5	57.4
Special education part A[6]	10	86.1	83.3
Special education part B[6]	10	86.1	83.3
Overall completion rate			
Child assessment[4]	62	46.9	54.9
Parent interview[5]	54	39.6	47.8
Student questionnaire[6]	62	46.9	54.9
School administrator questionnaire[6]	62	41.9	53.7
Teacher-level questionnaire[6]	63	46.3	55.7
English teacher questionnaire (child level)[6]	61	45.1	54.0
Mathematics teacher questionnaire (child level)[6]	33	43.7	56.7
Science teacher questionnaire (child level)[6]	27	42.0	49.3
Special education part A[6]	10	66.3	71.5
Special education part B[6]	10	66.3	71.5

[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] English, 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 2007.

The second component varies by survey instrument, and is much lower than in previous years. As discussed before, the completion rates dropped in general due to the time gap between the fifth-grade and eighth-grade data collections and the introduction of the explicit parent consent into eighth grade. Also, the number of children sampled is much smaller than in the past, a drop of 40 percent; there were 165 children sampled in first grade in the fifth-grade data collection; there were 100 of such children in the eighth-grade data collection. The rates for the paper-and-pencil instruments range from 54.5 percent for the child-level science teacher questionnaire to 86.1 percent for the special education questionnaire part A or B. The rate for the child assessment, at 60.9 percent, is almost 15 points lower than for the kindergarten sample, and the parent interview, at 51.5 percent, is about 20 points lower. The rates for the school instrument and the teacher instruments are all lower between 14 and 19 points, except for the special education teacher questionnaires where the difference is about 8 percentage points. 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 66.3 percent for the special education questionnaire part A or B to a low of 39.6 percent for the parent interview.

5.7.3 Spring-Eighth Grade Completion Rates—All Children

Table 5-16 presents final spring-eighth 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 rates for all children by less than half a percent relative to the rates for children sampled in kindergarten, even though the completion rates for children sampled in first grade are much lower than the kindergarten rates. The spring-eighth grade overall completion rates for the child assessment and the parent interview are 75 percent and 70.9 percent, respectively.

The unweighted completion rates are almost always higher than the weighted completion rates, by as much as 13 percent at the overall level. Where there is a large difference, it is due to fifth-grade movers who have larger weights than fifth-grade nonmovers. The weights of the fifth-grade movers had been increased in fifth grade to account for the subsampling of movers in fifth grade. This difference is not as large as in fifth grade, because movers in eighth grade were not subsampled out and no mover adjustment was applied to the weight. The fifth-grade mover adjustment, however, did apply to the eighth grade as explained in chapter 7.

Table 5-17 shows the completion rates for the child assessment, the parent interview, the student questionnaire, and the school and teacher instruments for children who have nonzero child weights (C7CW0>0). These are children whose spring-eighth grade English, mathematics, or science assessments were scorable, children who could not be assessed because of disabilities, or children who completed a student questionnaire. These conditioned completion rates are useful to analysts who want to assess the relationship between the different instruments in term of participation. The completion rates from the different instruments are dependent in that if data from one instrument are missing (e.g., parent instrument) it is likely that data from other instruments are also missing. (e.g., school administrator questionnaire). The conditioned completion rate for the child assessment is by definition 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.

When the completion rates are conditioned on the presence of the child weight, they are at least 17.5 points higher than the unconditional completion rates for all instruments but 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 all other instruments, the conditional completion rates are higher by 16.9 points for the parent interview, and as high as 21.2 points for the teacher-level questionnaire. The rate for the student questionnaire is not part of this comparison because almost all children who were assessed also completed the student questionnaire.

Table 5-16. Number of completed child-level cases and child-level completion rates, for children sampled in kindergarten and first grade, by survey instruments: School year 2006–07

Survey instrument	Children sampled in kindergarten			Children sampled in first grade			All children		
		Completion rate			Completion rate			Completion rate	
	Completes	Weighted	Unweighted	Completes	Weighted	Unweighted	Completes	Weighted	Unweighted
Child assessment[1]	9,296	75.7	77.9	62	46.9	54.9	9,358	75.0	77.7
Parent interview[2]	8,755	71.7	73.4	54	39.6	47.8	8,809	70.9	73.2
Student questionnaire[3]	9,244	75.3	77.5	62	46.9	54.9	9,306	74.6	77.4
School administrator questionnaire[3]	9,200	73.3	77.0	62	41.9	53.7	9,262	72.5	76.8
Teacher-level questionnaire[3]	9,147	74.5	77.0	63	46.3	55.7	9,210	73.8	76.9
English teacher questionnaire (child level)[3]	8,957	73.2	75.4	61	45.1	54.0	9,018	72.5	75.3
Mathematics teacher questionnaire (child level)[3]	4,449	71.6	75.2	33	43.7	56.7	4,482	70.9	75.1
Science teacher questionnaire (child level)[3]	4,459	73.3	74.8	27	42.0	49.3	4,486	72.5	74.6
Special education part A[3]	812	93.9	94.5	10	66.3	71.5	822	93.2	94.3
Special education part B[3]	820	94.7	95.5	10	66.3	71.5	830	94.0	95.3

[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 2007.

Table 5-17. 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 2006–07

Survey instrument	Children sampled in kindergarten			Children sampled in first grade			All children		
		Completion rate			Completion rate			Completion rate	
	Completes	Weighted	Unweighted	Completes	Weighted	Unweighted	Completes	Weighted	Unweighted
Child assessment[1]	9,296	100	100	62	78.7	86.7	9,358	99.5	99.9
Parent interview[2]	8391	89.8	90.3	51	63.6	71.4	8,442	89.1	90.2
Student questionnaire[3]	9244	99.6	99.4	62	78.7	86.7	9,306	99.1	99.3
School administrator questionnaire[3]	8741	94.1	94.4	58	70.4	81.1	8,799	93.5	94.3
Teacher-level questionnaire[3]	9090	97.8	98.1	61	76.8	85.4	9,151	97.3	98.0
English teacher questionnaire (child level)[3]	8914	96.2	96.2	59	74.8	82.6	8,973	95.7	96.1
Mathematics teacher questionnaire (child level)[3]	4426	95.5	95.8	31	76.4	84.1	4,457	95.0	95.7
Science teacher questionnaire (child level)[3]	4444	94.9	95.7	27	66.1	78.1	4,471	94.2	95.6
Special education part A[3]	803	94.2	95.0	9	67.6	71.0	812	93.5	94.8
Special education part B[3]	811	95.0	96.0	9	67.6	71.0	820	94.3	95.8

[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 2007

5.7.4 Overall Response Rates

The ECLS-K overall response rate can be computed by the product of the school-level response rate from the base year and the completion rates from each round of data collection after the base year. Table 5-18 presents the overall response rate after data collection for 5 school years: base year, first grade, third grade, fifth grade, and eighth grade, and for each study instrument that is common to all rounds of data collection: child assessment, parent interview, school administrator questionnaire, teacher-level questionnaires A and B (replaced by one single teacher-level questionnaire in fifth and eighth grade), child-level teacher questionnaire part C (replaced by the reading/English child-level questionnaire in fifth and eighth grade), and the two special education questionnaires A and B.

The instrument-specific overall response rates are driven by the school-level response rate in the base year. Since the overall school response rate is low at 74 percent, overall response rates for all instruments cannot be higher than 74 percent. In fact, they range between 62 and 70 percent in the base year, and steadily drop each year until they range only between 17 and 38 percent in eighth grade. Leaving aside the special education questionnaires that were administered to a small selected sample, the instrument with the highest overall response rate by the end of the study in eighth grade is the child assessment, followed by the parent interview. The school and teacher questionnaires have about the same overall response rates. The drop in the overall response rate from year to year is natural in a longitudinal study.

Table 5-18. Kindergarten to eighth grade overall response rate: School year 2006–07

Data collection	Completes[2]	Completion rate		Overall response rate	
		Weighted	Unweighted	Weighted	Unweighted
Kindergarten, school level	1,014	74.0	73.7	74.0	73.7
Kindergarten, child level					
Child assessment[1]	19,967	88.0	88.3	65.1	65.1
Parent interview[2]	18,950	83.9	83.8	62.1	61.8
School administrator questionnaire[3]	19,282	85.9	85.4	63.6	62.9
Teacher questionnaire part A[3]	15,389	86.9	86.9	64.3	64.0
Teacher questionnaire part B[3]	15,880	89.7	89.6	66.4	66.0
Teacher questionnaire part C[3]	15,233	85.9	86.0	63.6	63.4
Special education part A[3]	737	94.1	92.2	69.6	68.0
Special education part B[3]	698	87.2	87.4	64.5	64.4
First grade, child level					
Child assessment[1]	16,727	87.2	91.6	56.8	59.6
Parent interview[2]	15,626	83.5	85.6	51.8	52.9
School administrator questionnaire[3]	14,764	75.9	81.3	48.2	51.2
Teacher questionnaire part A[3]	15,166	77.6	83.5	49.9	53.5
Teacher questionnaire part B[3]	15,022	77.0	82.7	51.1	54.6
Teacher questionnaire part C	15,123	77.4	83.3	49.2	52.8
Special education part A[3]	708	88.1	88.4	61.3	60.1
Special education part B[3]	664	82.4	82.9	53.2	53.4

See notes at end of table.

Table 5-18. Kindergarten to eighth grade overall response rate: School year 2006–07—Continued

Data collection	Completes[2]	Completion rate		Overall response rate	
		Weighted	Unweighted	Weighted	Unweighted
Third grade, child level					
Child assessment[1]	14,470	80.1	85.9	45.5	51.2
Parent interview[2]	13,489	76.9	80.1	39.9	42.3
School administrator questionnaire[3]	12,463	65.5	73.1	31.6	37.4
Teacher questionnaire part A[3]	11,856	61.7	69.6	30.8	37.2
Teacher questionnaire part B[3]	11,826	61.6	69.4	31.5	37.9
Teacher questionnaire part C[3]	11,884	62.0	69.7	30.5	36.8
Special education part A[3]	887	72.3	74.8	44.4	44.9
Special education part B[3]	883	72.2	74.5	38.4	39.8
Fifth grade, child level					
Child assessment[1]	11,346	83.9	93.4	38.2	47.8
Parent interview[2]	10,996	88.3	90.5	35.2	38.3
School administrator questionnaire[3]	11,023	76.4	89.4	24.1	33.4
Teacher-level questionnaire[3,4]	10,959	79.3	90.4	25.0	34.3
English teacher questionnaire (child level)[3,5]	10,877	78.7	89.8	24.0	33.0
Special education part A[3]	975	91.6	93.7	40.6	42.1
Special education part B[3]	981	92.9	94.2	35.7	37.5

See notes at end of table.

Table 5-18. Kindergarten to eighth grade overall response rate: School year 2006–07—Continued

Data collection	Completes[2]	Completion rate		Overall response rate	
		Weighted	Unweighted	Weighted	Unweighted
Eighth grade, child level					
Child assessment[1]	9,358	75.0	77.7	28.6	37.2
Parent interview[2]	8,809	70.9	73.2	25.0	28.1
School administrator questionnaire[3]	9,262	72.5	76.8	17.5	25.7
Teacher-level questionnaire[3,4]	9,210	73.8	76.9	18.4	26.3
English teacher questionnaire (child level)[3,5]	9,018	72.5	75.3	17.4	24.9
Special education part A[3]	822	93.2	94.3	37.9	39.7
Special education part B[3]	830	94.0	95.3	33.5	35.7

[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.
[4] Teacher questionnaires part A and part B were replaced by the teacher-level questionnaire in fifth and eighth grade.
[5] Teacher questionnaire part C was replaced by the subject-specific teacher questionnaire in fifth and eighth grade.
SOURCE: U.S. Department of Education, National Center for Education Statistics, Early Childhood Longitudinal Study, Kindergarten Class of 1998–99 (ECLS-K), spring 2007.

5.7.5 Nonresponse Bias Analysis

Estimates from nearly all surveys are potentially subject to bias due to nonresponse. Two aspects of the ECLS-K that increased the concern about nonresponse bias were its longitudinal design and the use of multiple sources for acquiring data about the sampled children. In the ECLS-K, nonresponse occurred in the initial base year of collecting data, and then attrition occurred in subsequent rounds of data collection. As in most longitudinal surveys, nonresponse in the ECLS-K generally increased as the sample aged. The use of multiple sources in the ECLS-K (e.g., direct child assessment, parent interview, teacher interview) provided the opportunity to obtain valuable data about the child, but it also presented multiple chances for nonresponse. For example, even if the child could be assessed, the parent might decline to be interviewed, and estimates using the parent data were subject to nonresponse. Chapter 6 of the *ECLS-K Methodology Report for the Eighth Grade* (NCES 2009–003) (Tourangeau et al. forthcoming) includes an examination of the potential for nonresponse bias using three methods: (1) comparison of respondents and nonrespondents using the available sample frame, (2) multivariate analysis to identify the characteristics of cases most likely to respond, and (3) analysis of attrition bias applicable to longitudinal studies. Nonresponse bias of the estimates from the eighth grade was present but small. In most cases, the use of a mover status category in the fifth-grade nonresponse adjustment weighting helped reduce the bias, and the sample-based raking to the characteristics of the base-year children further reduced the nonresponse bias and variance of the estimates. The proper use of the ECLS-K weights in data analysis will minimize the effect of nonresponse bias.

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 spring-eighth grade: computer-assisted interviews (CAI) and self-administered paper forms (hard copy). The data preparation approach differed with the mode of data collection. The parent interview was conducted using CAI techniques. Editing specifications were built into the computer programs used by interviewers to collect these data. The child assessments and student questionnaires were administered as hard-copy forms and were completed in a group setting. The teacher and school administrator questionnaires were self-administered on hard-copy forms. When the field supervisors returned these forms, coders recorded the receipt of these forms into a project-specific forms tracking system. The forms were then sent to a scanning subcontractor for transfer into an electronic format. After the data were scanned, upcoding was conducted, and the data were 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 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. 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 times a child changed from one school to another since spring 2004 had a soft range of 0 to 3. A value outside this range could be entered and confirmed as correct by the interviewer as long as it was within the hard range of values (0 to 5).

"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.

Parent interview items on house value. No hard coding range was specified for items asking about the remaining principal on the house (PAQ020). In 82 cases, the remaining principal on the house (PAQ200) was greater than the house value (PAQ150). For some of these cases, the difference between the principal and value was less than $10,000; however, in other cases the discrepancy seemed unusually high. For example, 48 cases had principal values that exceeded the home value by at least $50,000. Therefore, analysts are advised to scrutinize those cases having remaining principal on the house greater than the house value and use judgment when working with these cases.

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 interviewer recorded the problem either in a comment or on a problem report.

6.1.3 Additional Coding

Additional coding was required for some of the items collected in the CAI instrument. These items included "Other, specify" text responses, occupation, and race/ethnicity. 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.

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 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. In some cases, the post-data collection "Other, specify" text upcoding resulted in a routing question being set to a category that would route to another item that was correctly skipped during the interview. In those cases, the skipped item was set to -9. Users should be aware that in these cases, the values of -9 are due to the post-data collection "Other, specify" text upcoding and not due to early termination of the telephone interview.

Other cases of which users should be aware in which a value of -9 was set during the post-data collection editing are in twin households where a non-English language was spoken in the home (PLQ020 = 1). There are 12 records on the data file in which PLQ083 = -9 and PLQ090 = -9 for the second child of a set of twins. The Blaise CAPI program did not collect child-level language data for the twins in households speaking any language other than English. As a result, the child-level PLQ variables were set to -9 (Not Ascertained) for the 12 twins.

Parent involvement. In the eighth-grade data collection, parent data was collected in the fall rather than in the spring, as was the method in previous rounds. Because the data were collected at the beginning of the school year, items tapping parent involvement (PIQ020) in various school functions were followed by a question asking whether parents had yet had an opportunity to be involved in those functions. When indicated, responses were treated as "Other, specify" items and upcoded to "No opportunity yet" for PIQ020 in the data set.

Parent occupation coding. As in the kindergarten, first-grade, third-grade, and fifth-grade data collections, occupations were coded using the Industry and Occupation Coding Manual (NCES 2000-077) (U.S. Department of Education, National Center for Education Statistics 1999). 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, Bureau of the Census 1982) and *Standard Occupational Classification Manual—1980* (U.S. Department of Commerce, Office of Federal Statistical Policy and Planning 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 7.6.7 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 (36.8 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 (63.2 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. Coders also independently assigned a second code for autocoded cases. A coding supervisor adjudicated disagreements between the initial code and the verification code. The adjudication by the supervisor served to further train coders by presenting concrete examples of appropriate coding. Of the cases that were autocoded, 16.6 percent required adjudication because the verifier disagreed with the autocoding. Of the cases that were manually coded, 28.3 percent required adjudication because the manual coder and the verifier disagreed. After

coding, verification, and adjudication were completed, all of the data were sorted by job title and code to check the coding one last time for consistency and to catch any coding errors that may have been overlooked.

Race/ethnicity coding. The same coding rules used since the kindergarten year were used to code all race/ethnicity variables for children, resident parents, and nonresident parents. (See chapter 7, sections 7.6.1.4 and 7.6.2.9 for details on how the race and ethnicity variables were coded and how the race/ethnicity composite was created.)

Partially complete parent interviews. All "completed" parent instruments (i.e., had completed all sections of the parent interview) were retained in the final data file. A small number of interviews in eighth grade (199, less than 3 percent) terminated the parent interview after the Family Structure (FSQ) section but before the end of the instrument. These interviews were considered as "partially complete" cases and were also included in the data file. All instrument items after the interview termination point were set to -9 for "Not Ascertained."

Parent interviews in which the respondent terminated the interview prior to the FSQ section were considered "incomplete" and not retained on the data file.

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, third-grade, and fifth-grade files. One flag was used to identify cases that were edited for any of the reasons described below. The flag is P7EDIT; the flag was set to "1" if the case was edited in the given wave. There were 347 cases requiring edits in eighth 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. The relationship of an individual to both the focal child and the reference person was also 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 addition, the number of household roster errors by the interviewer was counted. For example, a household roster error would occur if an interviewer entered the same sibling into the household roster twice. In that instance, the interviewer would set the duplicate entry to "no longer in the household," and the reason departed would be set to "roster error." In the eighth-grade data, there are 14 cases with these types of errors after the roster tests were run; the cases can be identified by the flag "P7ERRFLG."

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 chapter 5, section 5.3.3 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 composition of the classroom and the percent of instructional time devoted to various subjects.

In the second step, all returned instruments were selected for classrooms with multiple children that had missing data for the race and percent of instructional time 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 30 English records, 13 mathematics records, and 20 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 14 such cases for English, 5 such cases for mathematics, and 10 such cases for science. In some cases, the teacher had reported the classroom characteristic items on a questionnaire other than that designated for the key child, and those data were used for that classroom.

6.2 Coding and Editing Specifications for Hard-Copy Questionnaires and Assessments

6.2.1 Receipt Control

In order to monitor the almost 96,000 documents that were to be received in the eighth-grade year, the project-specific receipt and document control system developed in the kindergarten year was used, with modifications to track hard-copy questionnaires sent to and received from the scanning subcontractor. 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, they were sorted by instrument type and ID number. Batch forms were generated and included in the batch to indicate which questionnaires were included in the batch. The child assessment forms, the student questionnaire, the teacher questionnaires, and the school administrator questionnaires were batched and sent to the scanning subcontractor to be scanned into electronic form. When these instruments were returned from the scanning subcontractor, 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 batch was returned. When the clerks verified that the correct questionnaires were returned, they were logged into the receipt and document control system.

Data from two hard-copy forms, the English Stage 1 Routing test and the Mathematics/ Science Stage 1 Routing test, were keyed into electronic format by Westat data entry staff. 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. In the case of differences, 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 study staff and loaded into the computer system for data editing. When these instruments were returned from the Westat data entry staff, the identification number on each form was also matched against the identification number in the tracking system to verify that the appropriate number of forms for each batch was returned. When the clerks verified that the correct forms were returned, they were logged into the receipt and document control system.

The following sections describe the coding and editing processes for hard-copy questionnaires.

6.2.2 Coding

6.2.2.1 Coding of Questionnaires

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 process. This staff also edited the data after scanning and the data were loaded into the system. Senior coders verified coding.

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.

6.2.2.2 Coding of Reading and Mathematics Assessment Forms

The hard-copy assessments required coding of open-ended items on the reading and mathematics assessment forms (the science forms had only multiple choice items that were scored programmatically). The coding staff were trained on the coding procedures and had coding manuals to support the process. All open-ended items were coded twice by different coding staff members and compared for agreement. Percent agreement for the open-ended reading items, across the Red and Orange Reading forms, was 95 percent. Percent agreement for the open-ended mathematics items, across the Blue and Green Math forms, was 98 percent. Discrepancies were adjudicated by a senior coder.

6.2.3 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 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 dataset. The frequencies were then rerun and reviewed. This was an iterative process; it was repeated until no further inconsistencies were found.

School participation in breakfast program (school administrator questionnaire). Item data on the school administrator questionnaire (SAQ) tapping school participation in the U.S. Department of Agriculture (USDA) School Breakfast Program were edited to enforce a skip pattern that was not clearly marked in the questionnaire. Based on their response to S7USDABR (q13: Does your school participate in USDA's School Breakfast Program?), respondents were to be routed to different sets of items. If S7USDABR = 1 (yes), respondents were supposed to skip S7FEWSTD, S7COSTLY, S7LATEST, S7NOFACL, S7NOSTAF, and S7OTHER (i.e., q14). If S7USDABR = 2 (no), respondents were supposed to skip S7BRKSTR, S7BRKEND, S7BRKLOC, S7BRKCLR, S7PRABRK, S7ELIBRK, S7PARBRK, S7ELRPBK, and S7PARPBK (i.e., q15–q18). This skip was enforced in post-collection data editing.

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.