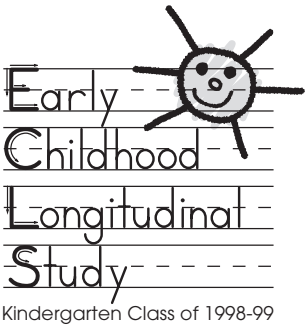
National Center for Education Statistics



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

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

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

February 2006

Karen Tourangeau
Christine Nord
Thanh Lê
Westat

Judith M. Pollack
Educational Testing Service

Sally Atkins-Burnett
University of Toledo

Elvira Germino Hausken
Project Officer
National Center for Education Statistics

U.S. Department of Education
Margaret Spellings
Secretary

Institute of Education Sciences
Grover J. Whitehurst
Director

National Center for Education Statistics
Mark Schneider
Commissioner

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

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

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

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

February 2006

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

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

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

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

TABLE OF CONTENTS

TABLE OF CONTENTS (continued)

TABLE OF CONTENTS (continued)

TABLE OF CONTENTS (continued)

TABLE OF CONTENTS (continued)

TABLE OF CONTENTS (continued)

TABLE OF CONTENTS (continued)

TABLE OF CONTENTS (continued)

TABLE OF CONTENTS (continued)

TABLE OF CONTENTS (continued)

List of Tables

TABLE OF CONTENTS (continued)

List of Tables (continued)

TABLE OF CONTENTS (continued)

List of Tables (continued)

TABLE OF CONTENTS (continued)

List of Tables (continued)

List of Tables (continued)

TABLE OF CONTENTS (continued)

List of Tables (continued)

TABLE OF CONTENTS (continued)

List of Tables (continued)

List of Tables (continued)

TABLE OF CONTENTS (continued)

List of Tables (continued)

TABLE OF CONTENTS (continued)

List of Exhibits

List of Exhibits (continued)

List of Exhibits (continued)

List of Exhibits (continued)

GETTING STARTED

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

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

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

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

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

Major Differences in the Fifth-Grade Data Collection

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

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

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

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

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

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

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

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

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

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

- **In fifth grade, a different approach from previous rounds was used to collect information from teachers.** The approach for administering teacher questionnaires differed from that of previous rounds because many fifth-grade children were

expected to have different teachers for different subject areas. In earlier rounds, all questions pertaining to the core academic subjects were asked in a single questionnaire and given to teachers who had sample children in their homeroom class. In the fifth grade, however, separate questionnaires were given to sample children's reading/language arts, mathematics, and science teachers. Teacher questionnaire content changes are described in section 2.3. Information about how to use the variables in the teacher questionnaire is presented in section 7.2.

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

Cautions and Caveats

Users of previous rounds of the ECLS-K data have frequently asked certain questions. For example, can school-level and teacher-level estimates be made with the ECLS-K data? Or, did the ECLS-K sample whole classrooms? NCES has developed a set of responses to users' most common questions. Please see the NCES web site for commonly asked questions and responses: http://nces.ed.gov/ecls.

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

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

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

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

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

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

- **Use of weights.** The fifth-grade data file contains 5 sets of cross-sectional weights and 10 longitudinal (panel) weights. Although there are a variety of weights on the file, there are scenarios for which there may not be a perfect weight. For a discussion of

the weights and guidance in selecting an appropriate one, refer to sections 4.7 and 9.3.1.

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

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

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

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

Electronic Codebook Reference Guide

- **Electronic Codebook (ECB).** The ECB is designed to run under Windows 95®, Windows 98®, Windows 2000®, Windows XP®, or Windows NT® on a Pentium-class or higher personal computer (PC). The PC should have a minimum of 20 megabytes (MB) of available disk space. The ECB offers the most convenient way to access the data because it enables users to search the names and labels of variables, to examine question wording and response categories for individual items, and to generate SAS,

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

- **Data file.** The fifth-grade child catalog contains one record for each of 11,820 responding students in spring-fifth grade. Data collected from teachers and schools are stored in the child catalog. The file, named child5p.dat, is stored in the root directory of the CD-ROM as an ASCII file. It is strongly recommended, however, that users access the data using the ECB software available on the CD-ROM rather than access the ASCII file directly. Appendix B on the CD-ROM contains the record layout for the child catalog.

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

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

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

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

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

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

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

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

1. INTRODUCTION

This manual provides guidance and documentation for users of the fifth-grade data[1] of the Early Childhood Longitudinal Study, Kindergarten Class of 1998–99 (ECLS-K). It begins with an overview of the ECLS-K study. Subsequent chapters provide details on the instruments and measures used, the sample design, weighting procedures, response rates, data collection and processing procedures, and the structure of the data file.

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

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

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

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

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

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

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

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

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

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

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

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

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

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

The ECLS-K is part of a longitudinal studies program comprising two cohorts—a kindergarten cohort and a birth cohort. The birth cohort (ECLS-B) is following a national sample of children born in the year 2001 from birth to kindergarten. The ECLS-B examines how early learning environments are associated with early cognitive, physical, and socioemotional development and thus

prepare children for kindergarten success. Together these cohorts will provide the depth and breadth of data required to more fully describe and understand children's early learning, development, and education experiences.

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

1.1 Background

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

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

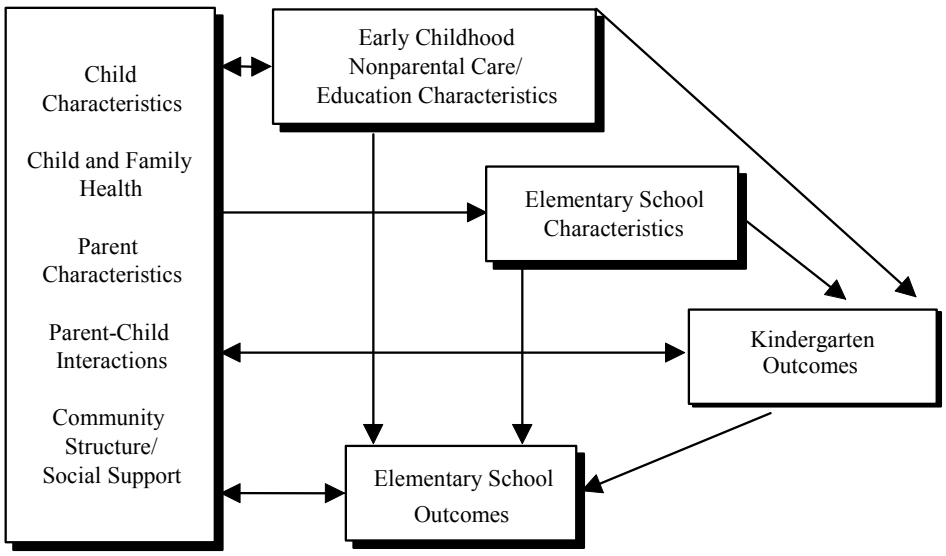
The spring-fifth-grade data collection can be used to describe the diversity of the children in the study and the classrooms and schools they attend. It can also be used to study children's academic

gains in the years following kindergarten. The ECLS-K sample includes substantial numbers of children from various minority groups. Thus, the ECLS-K data present many possibilities for studying cultural and ethnic differences in the educational preferences, home learning practices, and school involvement of families; the developmental patterns and learning styles of children; and the educational resources and opportunities that different groups are afforded in the United States.

1.2 Conceptual Model

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

Exhibit 1-2. ECLS-K conceptual model



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

1.3 Study Components

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

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

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

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

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

from school administrators via self-administered questionnaires during each spring data collection.

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

1.4 ECLS-K Data Files

The ECLS-K data are released in restricted-use and public-use versions. A brief overview of the differences between the restricted-use and public-use files is provided here, followed by a description of the data files that are currently available.

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

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

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

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

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

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

Variable X frequency distribution

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

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

Masked variable X frequency distribution

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

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

1.4.2 Overview of Available Data Files

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

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

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

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

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

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

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

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

  The first-grade data file can be used not only to analyze data collected in the first grade but also to provide weights and variables that can be used in longitudinal data analysis of both kindergarten and first grade. In addition to the cross-sectional weights, cross-year (kindergarten–first grade) weights have been added to the first-

grade data file for those analysts who wish to examine children's learning across school years. A longitudinal public-use file, however, is available that combines the base year and first-grade data (see next bullet). Most analysts will find it more convenient to use the already created longitudinal file described below. For more information about the first-grade file, see the *User's Manual for the ECLS-K First Grade Public-Use Data Files and Electronic Codebook* (NCES 2002–135) (Tourangeau, Burke, et al. 2002).

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

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

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

- The **Salary and Benefits File** is collected at the school level and contains information on the base salary, merit pay, and benefit pay of teachers and principals. The salary and benefits data, when combined with other ECLS-K data, can be used to examine, for example, the relationship between student outcomes and school resource allocation

and use. This file is only available as a restricted-use file. For more information about this file, see the *ECLS-K Base Year Restricted-Use Salary and Benefits File* (NCES 2001–014) (Tourangeau, Burke, et al. 2001b).

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

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

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

1.5 Contents of Manual

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

The manual contains information about the data collection instruments (chapter 2) and the psychometric properties of these instruments (chapter 3). It describes the ECLS-K sample design and weighting procedures (chapter 4); data collection procedures and response rates (chapter 5); and data processing procedures (chapter 6). In addition, this manual shows the structure of the fifth-grade data file and provides definitions of composite variables (chapter 7); describes how to install and use the Electronic Codebook (chapter 8); and describes how to use and merge the base year, first-grade, third-grade, and fifth-grade files (chapter 9). Finally, chapter 10 presents information on the longitudinal kindergarten–fifth grade public-use data file. The Electronic Codebook contains unweighted frequencies for all variables. Because this manual focuses on the fifth-grade data collection, minimal information is provided about the base year, first-grade, or third-grade data. Users who wish to learn more about these data collections should refer to the *ECLS-K Base Year Public-Use Data Files and Electronic Codebook: User's Manual* (NCES 2001–029r) (Tourangeau, Burke, Lê, et al. 2004); the *User's Manual for the ECLS-K First Grade Public-Use Data Files and Electronic Codebook* (NCES 2002–135) (Tourangeau, Burke, et al. 2002) or the *User's Manual for the ECLS-K Third Grade Public-Use Data File and Electronic Code Book* (NCES 2004–001) (Tourangeau, Brick, Lê, et al. 2004). Additional information about the ECLS program can be found on the World Wide Web at http://nces.ed.gov/ecls.

This page is intentionally left blank.

2. DESCRIPTION OF DATA COLLECTION INSTRUMENTS

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

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

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

[1] Adapted with permission from *Social Skills Rating System, Elementary Scale A ("How Often?")*, F.M. Gresham and S.N. Elliott. (1990). Circle Pines, MN: American Guidance Services, Inc.

[2] Adapted with permission from *Self-Description Questionnaire I,* H.W. Marsh. (1992). Campbelltown, N.S.W.: Australia: University of Western Sydney, Macarthur.

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

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

X Round that included the instrument.

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

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

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

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

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

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

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

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

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

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

2.1 Direct Child Assessments

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

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

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

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

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

X Round that included the instrument.

/ OLDS (Oral Language Development Scale) was given to language-minority students new to the study in the spring, or who did not pass the cut score in the English version during the previous OLDS administration. The screener determined if the children understood English well enough to receive the direct child assessments in English. For further information on the language screener, please refer to the *ECLS-K Base Year Public-Use Data Files and Electronic Code Book: User's Manual* (NCES 2001-029r) (Tourangeau, Burke, Lê, et al. 2004). The screener was not used in third or fifth grade because the majority of minority-language children (86 percent) passed it by spring-first grade.

[2] In spring-third grade, the general knowledge assessment was replaced with a science assessment. Children received a science assessment in third and fifth grade that measured their understanding of science concepts and scientific investigation skills.

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

2.1.1 Socioemotional Development

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

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

six scales.

- **SDQ Reading** scale includes eight items on reading grades, the difficulty of reading work, and their interest in and enjoyment of reading.
- **SDQ Mathematics** scale includes eight items on mathematics grades, the difficulty of mathematics work, and their interest in and enjoyment of mathematics.
- **SDQ School** scale includes six items on how well they do in "all school subjects" and their enjoyment of "all school subjects."
- **SDQ Peer** scale includes six items on how easily they make friends and get along with children as well as their perception of their popularity.
- **SDQ Anger/Distractibility** scale includes six items on externalizing problem behaviors such as fighting and arguing "with other kids," talking and disturbing others, and problems with distractibility.
- **SDQ Sad/Lonely/Anxious** scale includes eight items on internalizing problem behaviors such as feeling "sad a lot of the time," feeling lonely, feeling ashamed of mistakes, feeling frustrated, and worrying about school and friendships.

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

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

2.1.2 Cognitive Components

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

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

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

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

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

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

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

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

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

[5] Woodcock, McGrew, and Werder (1994). *Woodcock-McGrew-Werder Mini Battery of Achievement*, Itasca, IL: Riverside Publishing.
[6] See section 3.1.4.3 for a discussion of highest proficiency mastered.

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

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

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

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

and expository text (evaluating nonfiction). Children were required to identify the tone of a remark, the author's purpose for a selection, and evidence for and against theories discussed in the text.

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

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

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

The third- and fifth-grade mathematics assessments addressed the following content strands: number sense, properties, and operations; measurement; geometry and spatial sense; data analysis, statistics, and probability; and pattern, algebra, and functions. The cognitive processes (conceptual,

procedural, and problem solving) were assessed in each of the strands. Some of the items drew upon knowledge from more than one strand. For example, an item might require that a child apply knowledge about geometry, measurement, and number operations to answer the question correctly. Proficiency levels defined in the third-grade assessment included levels 4 and 5 retained from the earlier test forms, plus two new levels: place value, and rate and measurement. The fifth-grade mathematics assessment retained level 5 from K-1, levels 6 and 7 from third grade, and added two new levels, 8 and 9.

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

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

Science. The K-1 assessment battery differed from the third- and fifth-grade batteries. The K-1 battery included a measure of general knowledge whereas the third and fifth grade included a measure of science. The K-1 general knowledge assessment battery consisted of items that measured knowledge in the natural sciences and social studies in a single scale. The science subdomain measured two broad classes of science competencies: (1) conceptual understanding of scientific facts, and (2) skills

and abilities to form questions about the natural world, to answer such questions on the basis of the tools and the evidence collected, to communicate answers and to explain how the answers were obtained. The social studies subdomain included questions that measured children's knowledge in a wide range of disciplines such as history, government, culture, geography, economics, and law. The science subdomain included questions from the fields of earth, space, physical, and life sciences. The assessment items drew on children's experiences with their environment, and many questions related to more than one of the categories. The items captured information on children's conception and understanding of the social, physical, and natural world and of their ability to draw inferences and comprehend implications. The skills children need to establish relationships between and among objects, events, or people and to make inferences and to comprehend the implications of verbal and pictorial concepts were measured.

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

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

2.1.3 Physical Components

In the fall of the base year there were two parts to the physical component of the child assessment: psychomotor and anthropometric. The psychomotor component (fine and gross motor) was not administered beyond fall kindergarten. The anthropometric component was administered in all six rounds. The anthropometric component consisted of recording the children's height (in inches to the nearest quarter inch) and weight (in pounds) to measure their physical growth and development. A Shorr Board (for measuring height) and a digital scale were used to obtain the height and weight measurements, which were recorded on a height and weight recording form and entered into a laptop computer by field

staff. Height and weight were measured twice. For additional detail on the procedures used to collect height and weight, see the *ECLS-K Fifth-Grade Methodology Report* (NCES 2006–037) (Tourangeau, Lê, and Nord 2005).

2.1.4 Food Consumption

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

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

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

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

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

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

2.2 Parent Interview

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

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

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

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

See notes at end of exhibit.

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

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

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

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

2.3 Teacher Questionnaires

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

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

The reading teacher questionnaire had three different sections. The first section included questions from the Social Rating Scale (SRS) that collected data on five areas of children's social skills. The second section had questions from the Academic Rating Scale (ARS) and gathered data on each sampled child's skills in areas of language and literacy. The third section asked child-specific instructional information (e.g., child's grade, additional tutoring or services the child received), asked the

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

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

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

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

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

See notes at end of exhibit.

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

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

See notes at end of exhibit.

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

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

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

2.3.1 Content of the Academic Rating Scale (ARS)

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

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

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

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

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

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

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

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

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

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

2.3.2 Teacher Social Rating Scale

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

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

- The **Approaches to Learning** scale (Teacher SRS) measures behaviors that affect the ease with which children can benefit from the learning environment. It includes six items that rate the child's attentiveness, task persistence, eagerness to learn, learning independence, flexibility, and organization. In the third- and fifth-grade

administration, an item "child follows classroom rules" was added to the SRS to increase variance in the self-control scale.

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

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

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

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

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

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

These measures are adapted with permission from the instrument *Social Skills Rating Scale: Elementary Scale A ("How Often?")* (SSRS) by Gresham and Elliott (1990). The order of presentation of items was different on the SRS. Unlike the SSRS, the SRS did not separate the problem behavior items from the social skill items. On the SRS, the problem behavior items were interspersed throughout the SRS questionnaire to break any response sets. The SSRS uses a three point response scale while the SRS used a four point scale (never, sometimes, often, very often) and allowed respondents to indicate "no opportunity to observe." Only three of the SSRS social skills items are the same on the SRS. The remainder of the social skills items were adapted (N=6) or new (N=7). Some items were adapted completely to tap a wider representation of the skill (e.g., "keeps belonging organized," "forms and maintains friendships," "easily adapts to changes in routine," "pays attention well," "follows classroom rules"). One item was abbreviated to cover a wider range of situations ("controls temper"). Seven of the

social skills items were new items developed for ECLS-K (i.e., "is sensitive to the feelings of others," "respects the property rights of others," "shows eagerness to learn new things," "persists in completing tasks," "works independently," "expresses own feelings, opinions, and ideas without putting down those of others," "comforts or helps other children"). The SRS problem behavior scales were much shorter than the SSRS (ten items on the SRS compared with eighteen on the SSRS). Seven of the items on the SRS problem scales are identical to SSRS problem behavior items. The remaining three items are new (i.e., "worries about things," "talks during quiet study time") or adapted from the SSRS ("shows low self-esteem").

2.4 Special Education Teacher Questionnaires

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

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

Except for one change, the spring-fifth grade special education teacher questionnaires were identical to the ones used in spring-third grade. A question on the receipt of special education or related services due to an attention deficit/hyperactivity disorder (ADHD) was added to the spring-fifth grade questionnaire. Exhibit 2-5 provides a summary of the content areas addressed in the special education teacher questionnaires in spring-fifth grade and in the previous rounds. The questionnaires addressed

topics such as the child's disability, IEP goals, the amount and type of services used by sampled students, and communication with parents and general education teachers.

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

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

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

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

2.5 School Administrator Questionnaire

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

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

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

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

2.6 School Facilities Checklist

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

2.7 Student Records Abstract Form

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

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

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

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

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

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

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

3.1 Direct Cognitive Assessment

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

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

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

[1] This user's manual is applicable to the data gathered during the 2003–04 school year; information contained in this manual about data gathered during the 1998–1999 school year (base year of the study), 1999–2000 school year (first grade), and 2001-02 school year (third grade) is provided primarily for background and comparison purposes.

[2] See chapter 2, section 2.1 of the *ECLS-K Psychometric Report for the Fifth Grade* (NCES 2006–036) (Pollack et al. 2005) for additional information on the two-stage process for the direct cognitive assessments.

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

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

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

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

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

The scores used to describe children's performance on the direct cognitive assessment included broad-based measures that reported performance in each domain as a whole, as well as targeted scores reflecting knowledge of selected content or mastery within a set of hierarchical skill levels. Some of the scores were simple counts of correct answers, while others were based on Item Response Theory (IRT), which uses patterns of correct and incorrect answers to obtain estimates that are comparable across different assessment forms. The different types of scores that were used to describe children's performance on the direct cognitive assessment are described in detail in this chapter. Number-right

scores and IRT scale scores measured children's performance on a set of questions with a broad range of difficulty. Standardized scores (T-scores) reported children's performance relative to their peers. Criterion-referenced proficiency scores and item cluster scores evaluated children's performance with respect to subsets of items that mark specific skills.

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

3.1.1 Number-Right Scores

Number-right scores are counts of the raw number of items a child answered correctly. These scores are useful for descriptive purposes only for assessments that are the same for all children. However, when these scores are for assessments that differ in difficulty, they are not comparable to each other. For example, a student who took the middle difficulty mathematics second-stage form would probably have answered more questions correctly if he or she had taken the easier low form, and fewer if the more difficult high form had been administered. For this reason, raw number-right scores were reported in the database only for the first-stage (routing) tests, which were the same for all children being assessed in that round of data collection. The routing test in each subject area consisted of sets of items spanning a wide range of skills. For example, the kindergarten-first grade (K-1) reading routing test emphasized pre-reading skills, while the routing tests in third and fifth grade contained questions based on reading passages as well as progressively more difficult decoding words. An analyst might use the routing

test number-right scores to report actual performance on these particular sets of tasks. Note that because the same routing test was used for the fall-kindergarten through spring-first grade data collections, rounds 1 through 4, score comparisons *may* be made among these rounds. However, the routing test scores in the third and fifth grades, which contained more difficult items, are *not* comparable with the kindergarten or first-grade number-right scores, nor with each other. The third-grade routing test number-right scores should be used only for comparisons *within third grade*, and the fifth-grade scores only *within fifth grade*, but not across grades.

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

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

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

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

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

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

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

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

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

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

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

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

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

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

each other because they are based on different numbers of questions and content that is not necessarily equivalent in difficulty (i.e., it would not be correct to assume that a child is doing better in reading than in mathematics because his or her IRT scale score in reading is higher than in mathematics).

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

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

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

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

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

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

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

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

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

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

3.1.3 Item Cluster Scores

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

3.1.3.1 Reading

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

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

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

A score based on four relatively difficult decoding items was reported for the third- and fifth-grade assessments. These were words that were unlikely to be in most children's everyday

vocabulary but could be sounded out phonetically. All four words were present in the third-grade reading assessment but only three of the four were in the fifth-grade forms. In order to make the fifth-grade decoding score comparable to the third-grade score for longitudinal comparisons, an estimate of the probability of a correct answer on the missing item was obtained, based on overall performance, for each fifth-grade test taker.

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

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

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

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

3.1.3.2 Science

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

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

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

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

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

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

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

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

3.1.4 Proficiency Levels

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

3.1.4.1 Reading

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

3.1.4.2 Mathematics

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

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

Two types of scores are reported with respect to the proficiency levels: a single indicator of highest level mastered, and a set of IRT-based probability scores, one for each proficiency level. More information on each of these types of scores is provided below. As for the other IRT-based scores (scale

scores and T-scores), re-estimated values for earlier rounds should be used when making comparisons of proficiency levels across rounds.

3.1.4.3 Highest Proficiency Level Mastered

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

Pass or fail for each proficiency level was based on actual counts of correct or incorrect responses, if they were present. If too few items were administered or answered to determine mastery of a level, a pass/fail score was assigned based on the remaining proficiency scores only if they indicated a pattern that was unambiguous. That is, a "fail" was inferred for a missing level if there were easier cluster(s) that had been failed *and* no higher cluster passed; or a "pass" was assumed if harder cluster(s) were passed *and* no easier one failed. In the case of ambiguous patterns (e.g., pass, missing, fail, where the missing level could legitimately be either a pass or a fail), an additional imputation step was undertaken that relied on information from the child's performance on all of the domain items answered in that round of data collection. IRT-based estimates of the probability of a correct answer were computed for each missing assessment item. Then a random number was generated and compared with the computed probability. A right answer was imputed if the random number was less than or equal to the probability; a wrong answer if the random number was greater than the probability. At a low level of ability (and low probability of a correct answer), at least some test takers could be expected to give a correct answer, even if only by guessing. Conversely, some children with a high probability of a right

answer (based on their other responses) might get any given item wrong. The imputation procedure employed took this into account, and thus preserved variance better than simply rounding the probability or setting an arbitrary cut point for imputation of right/wrong answers. These imputed responses were then aggregated in the same manner as actual responses to determine mastery at each of the missing levels. About 67 percent of the "highest level" scores in reading and 80 percent in mathematics were determined on the basis of item response data alone; the rest utilized IRT-based probabilities for some or all of the missing items. Scores were not imputed for missing levels that included a reversal (e.g., fail, blank, pass) because no resolution of the missing data could result in a consistent hierarchical pattern.

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

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

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

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

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

3.1.4.4 Proficiency Probability Scores

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

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

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

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

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

See notes at end of table.

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

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

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

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

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

See notes at end of table.

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

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

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

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

- At entry to kindergarten about 68 percent (mean probability = .68) of children were proficient at letter recognition (C1R3RPB1).
- The largest gains between spring-kindergarten and spring-first grade were made in reading simple sight words, with 15 percent of children having mastered this skill at the end of kindergarten (C2R3RPB4) compared with 74 percent a year later (C4R3RPB4), 98 percent in third grade (C5R3RPB4) and 100 percent in fifth grade (C6R3RPB4).
- There were only small gains in letter recognition after spring-kindergarten, because most children, 93 percent, knew their letters by this time (C2R3RPB1).
- Children's skills in making inferences based on cues directly stated in text (literal inference) increased between first and third grade, from 16 percent (C4R3RPB6) to 68 percent (C5R3RPB6).
- In spring-third grade, most children had not yet demonstrated understanding of the author's craft or making connections between a problem in the narrative and similar life problems. While 24 percent mastered the evaluation level in third grade (C5R3RPB8), 44 percent demonstrated mastery in fifth grade (C6R3RPB8).

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

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

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

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

3.1.5 Choosing the Appropriate Score for Analysis

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

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

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

- A criterion-referenced versus norm-referenced interpretation.

3.1.5.1 Item Response Theory-Based Scores

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

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

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

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

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

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

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

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

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

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

3.1.6 Measuring Gains

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

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

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

The use of adaptive assessments increases the reliability of individual assessment scores by removing the sources of floor and ceiling effects. When assessment forms are matched to children's ability levels, all students have an equal chance to gain on the vertical scale, that is, a scale that spans several time points. Depending on how adaptive the measure is, how the scale is constructed, and how even-handed the educational treatment, one may not observe large differences in individual children's amounts of gain in total scale score points. Individual and group differences in the *amount* of gain given a fairly standard treatment (e.g., a year of schooling) can be relatively trivial compared to individual and group differences in *where* the gains take place. It is more likely that one will see substantial subgroup differences in initial status than in gains, suggesting that the gains being made by individuals at different

points on the score scale are qualitatively different. Thus analysis of the total IRT scale score without explicitly taking into consideration where the gain takes place tells only part of the story.

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

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

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

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

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

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

3.1.7 Reliability

Reliability statistics assess consistency of measurement, in other words, the extent to which a set of test items is related to each other and to the score scale as a whole. For tests of equal length, reliability estimates can be expected to be higher for sets of items that are closely related to the underlying

construct than for tests with more diversity of content. Conversely, for tests with similar levels of diversity in content, reliabilities tend to be higher for longer tests compared with shorter tests. In general, the most diverse subject, science, had lower reliability coefficients than reading and mathematics. Reliabilities for scores using the greatest number of test items, the IRT ability estimates that are based on all items taken by each child, were highest. Reliabilities for scores based on the fewest items, the item cluster scores in reading and science, were lowest. Reliability statistics appropriate for each type of score were computed for each subject area for each round of data collection.

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

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

The definition of reliability–consistency of measurement under different circumstances–suggested an appropriate way to assess the reliability of the "highest proficiency level mastered" score. The score denoting the highest level mastered reduces the series of pass/fail scores on the hierarchical set of proficiency levels to a single score. For example, a student demonstrating mastery of the first five reading levels but not the remaining four would be said to have a "highest proficiency mastered" score of

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

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

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

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

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

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

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

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

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

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

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

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

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

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

3.1.8 Validity

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

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

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

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

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

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

3.2 Indirect Cognitive Assessment

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

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

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

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

3.2.1 Comparison to Direct Cognitive Assessment

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

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

3.2.2 Rasch Scores Available for the Academic Rating Scale

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

Tables 3-22 to 3-24 provide standard errors (SE) for each of the Rasch scores for fifth grade. The "Score" column is the sum of the raw score ratings. "Measure" is the Rasch-based score. The column

labeled "SE" is the corresponding standard error of measurement for those scores. These standard errors can be used in analytic models to correct for the heteroskedasticity of scores.

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

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

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

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

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

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

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

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

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

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

3.3 Teacher Social Rating Scale

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

[1] T6LEARN = T6 Approaches to Learning
T6CONTRO = T6 Self-control
T6INTERP = T6 Interpersonal
T6EXTERN = T6 Externalizing Problem Behaviors
T6INTERN = T6 Internalizing Problem Behaviors
T6SCINT = T6 Combo of Self-Control & Interpersonal (Peer Relations)
NOTE: See chapter 7, section 7.4 for variable naming conventions.
SOURCE: U.S. Department of Education, National Center for Education Statistics, Early Childhood Longitudinal Study, Kindergarten Class of 1998–99 (ECLS-K), spring 2004.

3.4 Self-Description Questionnaire

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

are skewed (negatively skewed for the positive social behavior scales, and positively skewed for the problem behavior scales). The reliability is lower for scales with only six items (see table 3-29). The means and standard deviations for the scales are presented in table 3-30.

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

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

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

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

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

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

4. SAMPLE DESIGN AND IMPLEMENTATION

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

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

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

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

for analytic reasons. Children not in the language minority group continued to be subsampled for followup if they moved in third grade.

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

4.1 Base Year Sample

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

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

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

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

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

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

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

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

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

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

The third stage sampling units were children of kindergarten age, selected within each sampled school. The goal of the student sample design was to obtain an approximately self-weighting sample of students and at the same time to achieve a minimum required sample size for APIs who were

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

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

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

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

4.2 Fall-First Grade Subsample

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

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

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

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

See notes at end of table.

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

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

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

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

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

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

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

The fall-first grade data collection consisted of the direct child assessment and the parent interview. Data collection was attempted for every eligible child found still attending the school in which he or she had been sampled during kindergarten and a subset of eligible children who had transferred from the school in which he or she was originally sampled. "Eligible" is defined as a base year respondent (i.e., a child who had either a fall- or spring-kindergarten child assessment or parent interview or was excluded from assessment because of a disability or because the child belonged in the language minority,

not Spanish group). To contain the costs of data collection, a random 50 percent of children were flagged to be followed for fall-first grade data collection in the event that they had transferred.

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

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

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

4.3 Spring-First Grade Sample

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

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

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

See notes at end of table.

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

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

† Not applicable.
[1] School offers kindergarten and at least another grade between 1st and 12th.
NOTE: School characteristics (i.e., region, locale, school affiliation, and school type) describe the school the child attended in kindergarten.
SOURCE: U.S. Department of Education, National Center for Education Statistics, Early Childhood Longitudinal Study, Kindergarten Class of 1998–99 (ECLS-K), fall 1999 and spring 2000.

4.3.1 Subsampling Movers

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

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

having its ECLS-K transfer students followed during spring-first grade, and any transfer student who had been followed in fall-first grade would still be followed in spring-first grade.

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

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

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

See notes at end of table.

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

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

† Not applicable.
[1] School offers kindergarten and at least another grade between 1st and 12th.
NOTE: School characteristics (i.e., region, locale, school affiliation, and school type) describe the school the child attended in kindergarten.
SOURCE: U.S. Department of Education, National Center for Education Statistics, Early Childhood Longitudinal Study, Kindergarten Class of 1998–99 (ECLS-K), fall 1999 and spring 2000.

4.3.2 Student Freshening

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

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

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

4.4 Spring-Third Grade Sample

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

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

4.4.1 Subsampling Movers

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

4.4.2 Language Minority Children

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

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

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

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

See notes at end of table.

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

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

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

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

4.5 Spring-Fifth Grade Sample

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

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

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

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

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

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

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

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

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

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

language minority children in third grade. For fifth grade, these mover followup activities targeted a 42 percent subsample of movers where were eligible to be fielded in fifth grade and resulted in a 41 percent subsample.

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

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

4.6 Sample Attrition

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

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

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

See notes at end of table.

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

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

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

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

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

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

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

† Not applicable.

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

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

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

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

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

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

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

† Not applicable.

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

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

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

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

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

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

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

† Not applicable.

[1] 535 private school children were sampled from refusal-converted schools.

[2] 4,554 private school children remained eligible after the base year. In addition, 19 private school children were sampled via the sample freshening procedure.

[3] 1,201 children from original private schools were excluded from the fifth-grade data collection. They were children who became ineligible in an earlier round, movers not subsampled to be followed in previous rounds, and hard-to-field cases such as hard refusals, and children with neither first-grade nor third-grade data.

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

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

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

4.7 Calculation and Use of Sample Weights

As in previous years, the ECLS-K data were weighted to compensate for differential probabilities of selection at each sampling stage and to adjust for the effects of nonresponse. In the ECLS-K base year, weights were computed at the child, school and teacher levels. Estimates using the base year weights are representative of all kindergarten children, all schools with kindergarten programs and all kindergarten teachers. After the base year, only child-level weights were computed. The use of these weights is essential to produce estimates that are representative of the cohort of children who were in kindergarten in 1998–99 or in first grade in 1999–2000. Since the sample was not freshened after the first-grade year with third or fifth-graders who did not have a chance to be sampled in kindergarten or first grade (as was done in first grade), estimates from the ECLS-K third- and fifth-grade data are

representative of the population cohort rather than all third-graders in 2001–02 or all fifth-graders in 2003–04. The estimated number of third-graders from the third-grade ECLS-K data collection is approximately 86 percent of all third-graders. From the fifth-grade ECLS-K data collection, the estimated number of fifth-graders is approximately 83 percent of all fifth-graders. While the vast majority of children in third grade in the 2001–02 school year and in fifth grade in the 2003–04 school year are members of the cohort, third-graders who repeated second or third grade, fifth-graders who repeated third or fourth grade, and recent immigrants are not covered. Data were collected from teachers and schools to provide important contextual information about the environment for the sampled children. The teachers and schools are not representative of third-grade teachers and schools in 2001–02, nor of fifth-grade teachers and schools in 2003–04. For this reason, the weights produced from the study after the kindergarten year are for making statements about children, including statements about the teachers and schools of those children.

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

4.7.1 Types of Cross-Sectional Sample Weights

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

Careful consideration should be given to which set of weights is appropriate for the desired analysis. Using the wrong weights will result in more biased or inefficient estimates. For example, if C6CPTR0 were used in an analysis of child and teacher-level data only, then the resulting estimates will be inefficient compared to estimates using C6CW0. The lower parent response causes C6CPTR0 to result in a smaller sample with positive weights. If using C6CPTR0 with child-level data from the questionnaire filled out by the mathematics teacher, then there will be missing mathematics-related data for

approximately half of the children. There may be combinations of data for which no weights were specifically developed, but all analyses should incorporate whichever weight that matches most closely.

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

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

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

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

4.7.2 Weighting Procedures

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

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

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

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

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

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

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

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

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

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

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

4.7.3 Computation of Spring-First Grade Initial Child Weights

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

4.7.3.1 Base Year Nonresponse-Adjusted School Weights

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

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

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

4.7.3.2 Base Year Child Weights

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

Base year child weights for base year respondents. As previously described, a base year respondent was defined as one with at least one direct cognitive test score in fall- or spring-kindergarten or was excluded from assessment because of a disability or because the child belonged in the language minority/not Spanish group, or whose parent responded to the family structure section of the parent instrument in fall- or spring-kindergarten. In terms of weights, a base year respondent is a sampled child with a positive fall- or spring-kindergarten weight (i.e., C1CW0, C2CW0, C1PW0 or C2PW0 weights). The base year child weight is the product of the base year nonresponse-adjusted school weight and the inverse of the within school selection probability of the child, adjusted for child-level nonresponse. The nonresponse weighting classes included school characteristics from the school nonresponse adjustments such as school affiliation, locale, region, school enrollment class, and child characteristics such as age group, sex, and race/ethnicity. These weighting classes are similar to those used for the original child weights in fall- and spring-kindergarten. For a description of the computation of child weights in fall- and

spring-kindergarten, see chapter 4, section 4.3.4 of *Early Childhood Longitudinal Study, Kindergarten Class of 1998–99 (ECLS-K) Base Year Public-Use Data Files and Electronic Codebook: User's Manual* (NCES 2001–029r) (Tourangeau, Burke, Lê, et al. 2004).

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

4.7.4 Computation of Spring-Fifth Grade Child Weights

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

4.7.4.1 Adjustment for Movers

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

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

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

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

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

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

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

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

4.7.4.2 Adjustment for Nonresponse

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

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

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

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

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

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

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

4.7.4.3 Raking to Sample-Based Control Totals

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

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

SES, language minority status, whether sampled in kindergarten or first grade, and if sampled in kindergarten, mover status.

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

4.7.5 Types of Cross-Sectional Weights and Their Use

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

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

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

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

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

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

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

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

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

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

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

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

disability), whose parent completed the FSQ section of the parent interview for spring-fifth grade, and who had completed teacher-level data either from the reading teacher and/or the mathematics/science teacher.

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

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

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

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

4.7.6 Replicate Weights

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

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

4.7.7 Characteristics of Cross-Sectional Sample Weights

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

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

4.8 Variance Estimation

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

with two units per stratum, the paired jackknife replication method (JK2) is recommended. This section describes the JK2 and the Taylor Series estimation methods.

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

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

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

4.8.1 Paired Jackknife Replication Method

In this method, a survey estimate of interest is calculated from the full sample. Subsamples of the full sample are then selected to calculate subsample estimates of the same parameter. The subsamples are called *replicates*, and the subsample estimates are called *replicate estimates*. The variability of the replicate estimates about the full sample estimate is used to estimate the variance of the full sample estimate. The variance estimator is computed as the sum of the squared deviations of the replicate estimates from the full sample estimate (Wolter 1985):

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

where

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

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

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

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

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

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

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

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

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

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

4.8.2 Taylor Series Method

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

type of weight, the sample size was examined, and then strata were combined when the sample size was not adequate. The sequential numbering of strata and first-stage units was done separately for each weight. Consequently, there is a different set of stratum and first-stage unit identifiers for each set of weights.

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

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

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

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

4.8.3 Specifications for Computing Standard Errors

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

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

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

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

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

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

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

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

4.9 Design Effects

An important analytic device is to compare the statistical efficiency of survey estimates from a complex sample survey such as the ECLS-K, with what would have been obtained in a hypothetical and usually impractical simple random sample (SRS) of the same size. In a stratified clustered design like the ECLS-K, stratification generally leads to a gain in efficiency over simple random sampling, but clustering has the opposite effect because of the positive intracluster correlation of the units in the cluster. The basic measure of the relative efficiency of the sample is the *design effect*, defined as the ratio, for a given statistic, of the variance estimate under the actual sample design to the variance estimate that would be obtained with an SRS of the same sample size:

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

The root design effect, *DEFT*, is defined as:

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

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

4.9.1 Use of Design Effects

Methods of computing SEs for the ECLS-K are jackknife replication and Taylor Series linearization. If statistical analyses are conducted using software packages that assume the data were collected using simple random sampling, the SEs will be calculated under this assumption and should be corrected using *DEFT*.[6] The SE of an estimate under the actual sample design can be approximated as follows:

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

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

Means:
$$Var_{SRS} = \frac{1}{n}\frac{\sum_{1}^{n} w_i \left(x_i - \overline{x_w}\right)^2}{\sum_{1}^{n} w_i} = SE_{SRS}^2\,,$$

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

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

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

4.9.2 Median Design Effects for the ECLS-K

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

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

For each survey item, Table 4-20 presents the number of cases for which data are nonmissing, the estimate, the standard error taking into account the actual sample design (Design SE), the standard error assuming SRS (SRS SE), the root design effect (DEFT), and the design effect (DEFF). Standard errors (Design SE) were produced in WesVar using JK2 based on the actual ECLS-K complex design. For each survey item, the variable name as it appears in the ECLS-K fifth-grade Electronic Codebook (ECB) is also provided in the table. For more information on the variables used in this section, refer to chapter 3, which describes the assessment and rating scale scores used in the ECLS-K, and chapter 7, which has a detailed discussion of the other variables.

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

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

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

See notes at end of table.

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

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

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

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

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

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

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

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

5. DATA COLLECTION METHODS AND RESPONSE RATES

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

5.1 Overview of Data Collection Methods

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

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

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

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

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

5.2 Field Staff Training

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

5.2.1 Advance Contact and Recruitment Training

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

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

5.2.2 Spring-Fifth Grade Training

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

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

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

5.2.2.1 Certification of the Child Assessors

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

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

The written exercises were used to ensure that each trainee understood the coding rules for selected open-ended questions with particularly complex scoring rubrics. Each exercise included certain assessment items from the level form that was just discussed, with an assortment of possible responses. The trainees were instructed to score each response as either correct or incorrect. The exercises were then

scored by the co-trainer during the next training session. Trainees who did not achieve a passing score were asked to attend a training session in the evening to review the items. These trainees then re-took the same exercises that they had previously failed to pass.

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

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

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

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

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

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

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

5.3 Fall Preassessment School Contact

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

5.3.1 Advance Mailings

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

5.3.2 Preassessment Contact

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

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

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

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

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

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

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

Identifying the Key Child in Classrooms with Multiple Study Children. In grade 5, the design of the child-level teacher questionnaire was changed to include collecting data about the child's reading class and mathematics or science class. In previous rounds, children had been taught primarily in intact classrooms and teachers only reported classroom level information once for the classroom. Due to the design change, the teacher-child links were broadened to include the domain (reading, mathematics, or science) as well as information to identify the reading, mathematics, or science classroom. In order to reduce data collection burden for teachers who were linked to multiple sample children in the same class, a "Key Domain Child" was identified for each separate subject and class that each teacher taught. The teachers would be asked to report classroom level information only once in the questionnaire for the key domain child and child-level information for all sampled children in their class. Field supervisors

collected the teacher-child-domain-classroom link information about each child and entered the information into the FMS. The information was used to generate the hardcopy teacher questionnaires (see section 5.5.1 for more information). Refer to the *ECLS-K Fifth-Grade Methodology Report* (NCES 2006–037) (Tourangeau, Lê, and Nord 2005) for additional detail on the Key Child concept.

5.4 Tracing Activities during the Fifth-Grade Data Collection

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

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

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

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

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

5.5 Spring-Fifth Grade Data Collection

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

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

5.5.1 Preassessment School Contact

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

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

5.5.2 Conducting the Direct Child Assessments

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

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

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

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

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

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

†Not applicable.

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

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

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

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

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

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

5.5.3 Conducting the Parent Interview

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

5.5.5 Teacher and School Data Collection

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

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

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

During the field period, field supervisors followed up with school administrators and teachers in visits to the schools to conduct assessments and by telephone to collect completed

questionnaires, ensuring that questionnaires were not missing critical information and that completed questionnaires were mailed to Westat.

5.6 Data Collection Quality Control

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

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

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

5.6.1 Child Assessments Observations

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

and 100 percent for second observations), but some assessors were observed only once because they completed their assignments in April before the second observation could be scheduled.

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

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

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

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

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

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

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

[1] Two hundred and forty-two assessors were to be observed coming out of training; assessors completing only parent interviews were not observed. 239 assessors had initial observations; 227 assessors had second observations. Only 1 assessor failed to pass the observation and was released from the project.

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

5.6.2 Validation of Parent Interviews

Approximately 10 percent of the respondents who completed parent interviews were selected for a short re-interview conducted by a field supervisor (i.e., a "validation" interview). The first parent

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

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

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

5.6.3 Validation of School Visits

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

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

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

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

5.6.4 Assessor Interrater Reliability

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

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

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

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

† Not applicable

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

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

5.7 Spring-Fifth Grade Completion Rates

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

In this section, fifth-grade completion rates are presented for three groups of children: (1) children sampled in kindergarten, (2) children sampled in first grade through the freshening procedure, and (3) both groups combined. Completion rates for the fifth-grade data collection were computed with the same procedures used for spring-first grade and spring-third grade to allow for comparisons of completion rates for the three years of data collection following the base year. For spring-first grade and spring-third grade, the sample of children is the same: base year respondents (i.e., children who had either a fall- or spring-kindergarten child assessment or parent interview) and children sampled in spring-first

grade as part of sample freshening as described in section 4.3.2. For spring-fifth grade, the sample of children was reduced to exclude base year respondents who belonged in the following special groups as described in section 4.5: (1) children who became ineligible in an earlier round (because they died or moved out of the country), (2) children who were subsampled out in previous rounds because they moved out of the original schools and were not subsampled to be followed, (3) children whose parents emphatically refused to cooperate (hard refusals) in any of the data collection rounds since spring-kindergarten, and (4) children eligible for the third-grade sample for whom there are neither first-grade nor third-grade data. Among the 21,357 children who were eligible for the study after the base year, 16,143 were part of the fifth-grade data collection. Weighted completion rates were computed using the base weight (i.e., inverse of selection probabilities) adjusted for movers, but not adjusted for nonresponse.

5.7.1 Children Sampled in Kindergarten

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

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

Table 5-9 shows that the completion rates for the child assessment are quite high and uniform across school characteristics. Excluding the "unknown" category, the rates vary from 93.1 percent in non-Catholic private schools to 99.7 percent in schools in large towns. Similarly, the completion rates for the parent interviews were uniform across school characteristics ranging from 87.2 percent for children in schools with 750 or more students and in schools where 50 to 89 percent of level

[2] The categories of school affiliation in the tables in this chapter do not match categories of school affiliation in the tables in chapter 4. This is to allow users to compare completion rates in fifth grade with those in previous years.

children belong to the minority groups, to 94.3 percent for children in small towns (excluding the "unknown" category). The "unknown" category includes children who could not be located and those children who had moved into a nonsampled county. The category "unknown" also includes 35 children who were homeschooled and thus had no information concerning schools.

The "unknown" category aside, both the child assessment and the parent interview completion rates increased between third grade and fifth grade for all school characteristics. The completion rates by mover status are discussed later, but the rates of completing all the instruments are much lower for children who moved than for those who did not move.

Table 5-10 shows that the overall weighted completion rate is 77.1 percent for the school administrator questionnaire and 78.8 percent for the facilities checklist. The rate for school administrator questionnaires is 11 percentage points higher than the corresponding rate in third grade. The rate for facilities checklist is only about 2 percent higher. The completion rates for the school administrator questionnaire range from 87.4 percent for schools with 750 or more students to 100 percent for those in large towns (excluding the "unknown" category). Rates for the facilities checklist range from 90.3 percent for schools in the urban fringe of mid-size cities to 100 percent for schools in large towns. It is worth noting that the completion rates for the school administrator questionnaire are lower for schools with higher percentages of minorities, a phenomenon also observed in previous rounds for the school administrator questionnaire. However, this disparity decreased considerably after the base year, reflecting the success of increased data collection efforts targeted toward these schools.

Table 5-11 shows that the rates for the student records abstract are the lowest of all the instruments, as they were in previous years of the ECLS-K. For fifth grade, this rate is about 70 percent compared with 67 percent in third grade. The "unknown" category aside, the completion rates of the student records abstract range from 71.8 percent in the northeast region to 93.7 percent for children in large towns.

All four of the teacher questionnaires were completed at an overall rate of 78 to 80 percent, much higher than the 62 to 63 range achieved in third grade. The completion rates for the teacher-level questionnaire in table 5-11 are uniform across school characteristics, ranging from 86.9 percent for schools in the northeast and schools with 750 or more students to 99.7 percent for schools in large towns (excluding the "unknown" category). The same uniform rates are found for the subject-specific child-

Table 5-9. Number of completed child-level cases and child-level completion rates for the child assessment and parent interview for children sampled in the base year, by school characteristics: School year 2003–04

School characteristics[1]	Child assessment			Parent interview		
		Completion rates			Completion rates	
	Completes[2]	Weighted	Unweighted	Completes[3]	Weighted	Unweighted
All schools	11,260	84.7	93.6	10,913	89.1	90.7
School affiliation						
Public	9,187	96.2	97.7	8,518	89.9	90.6
Private	2,049	95.2	97.7	1,977	92.0	94.3
Catholic	1,313	97.1	98.3	1,260	94.0	94.3
Other private	736	93.1	96.7	717	89.5	94.2
Unknown	24	5.8	4.6	418	82.5	79.5
Type of locale						
Large city	1,863	97.2	97.6	1,697	89.6	88.9
Mid-size city	1,863	97.6	98.3	1,733	90.8	91.5
Urban fringe of large city	3,286	94.8	96.6	3,085	88.6	90.7
Urban fringe of mid-size city	767	93.5	96.8	725	90.3	91.5
Large town	283	99.7	99.6	270	91.9	95.1
Small town	825	94.1	98.8	781	94.3	93.5
Rural – outside MSA	1,286	97.2	97.9	1,205	88.1	91.7
Rural – inside MSA	922	97.7	99.4	871	93.4	93.9
Unknown	165	17.0	24.6	546	83.4	81.5
School size (total enrollment)						
1 to 299	2,359	96.7	98.2	2,255	93.4	93.8
300 to 499	3,703	96.5	98.0	3,453	89.9	91.4
500 to 749	3,167	97.1	97.6	2,945	90.8	90.8
750 or more	1,963	94.2	97.0	1,798	87.2	88.9
Unknown	68	9.7	11.7	462	82.1	79.8

See notes at end of table.

Table 5-9. Number of completed child-level cases and child-level completion rates for the child assessment and parent interview for children sampled in the base year, by school characteristics: School year 2003–04—Continued

School characteristics[1]	Child assessment			Parent interview		
		Completion rates			Completion rates	
	Completes[2]	Weighted	Unweighted	Completes[3]	Weighted	Unweighted
Percent non-White enrolled						
0 – 10	3,509	97.9	98.3	3,352	93.2	93.9
11 – 49	3,705	94.5	97.2	3,522	90.5	92.4
50 – 89	1,956	96.5	97.4	1,768	87.2	88.0
90 – 100	1,997	97.0	98.0	1,786	88.2	87.7
Unknown	93	11.3	15.4	485	82.4	80.3
Region						
Northeast	2,080	95.5	96.8	1,956	90.5	91.1
Midwest	2,957	97.9	98.8	2,803	93.0	93.6
South	3,614	95.1	97.5	3,334	87.6	90.0
West	2,585	96.5	97.3	2,402	91.2	90.4
Unknown	24	5.8	4.6	418	82.5	79.5

[1] School characteristics are for schools attended by children in the ECLS-K fifth-grade sample and are based on ECLS-K survey data, not data from the sampling frame.
[2] Reading, mathematics, or science assessment was scorable, or child was disabled and could not be assessed.
[3] Family structure portion of parent interview was completed.
SOURCE: U.S. Department of Education, National Center for Education Statistics, Early Childhood Longitudinal Study, Kindergarten Class of 1998–99 (ECLS-K), spring 2004.

Table 5-10. Number of completed child-level cases and child-level completion rates for the school administrator questionnaire and facilities checklist for children sampled in the base year, by school characteristics: School year 2003–04

School characteristics[1]	School administrator questionnaire			Facilities checklist		
		Completion rates			Completion rates	
	Completes[2]	Weighted	Unweighted	Completes[2]	Weighted	Unweighted
All schools	10,937	77.1	89.6	11,154	78.8	91.4
School affiliation						
Public	8,884	90.7	94.5	9,084	92.7	96.6
Private	2,053	93.1	97.9	2,070	95.7	98.7
Catholic	1,323	96.8	99.0	1,328	98.1	99.4
Other private	730	88.8	95.9	742	92.8	97.5
Type of locale						
Large city	1,775	92.2	93.0	1,862	96.4	97.6
Mid-size city	1,825	93.6	96.3	1,844	94.9	97.3
Urban fringe of large city	3,173	87.7	93.2	3.240	90.4	95.2
Urban fringe of mid-size city	762	90.9	96.2	761	90.3	96.1
Large town	284	100.0	100.0	284	100.0	100.0
Small town	823	91.2	98.6	823	91.1	98.6
Rural – outside MSA	1,274	92.0	97.0	1,292	93.6	98.3
Rural – inside MSA	904	93.9	97.4	911	94.6	98.2
Unknown	117	8.2	13.9	137	9.4	16.3
School size (total enrollment)						
1 to 299	2,360	95.1	98.2	2,368	95.5	98.5
300 to 499	3,662	92.0	96.9	3.684	92.6	97.5
500 to 749	3,071	92.3	94.6	3,136	94.3	96.6
750 or more	1,844	87.4	91.2	1,939	91.9	95.8
Unknown	0	0.0	0.0	27	1.7	3.6

See notes at end of table.

Table 5-10. Number of completed child-level cases and child-level completion rates for the school administrator questionnaire and facilities checklist for children sampled in the base year, by school characteristics: School year 2003–04—Continued

School characteristics[1]	School administrator questionnaire			Facilities checklist		
		Completion rates			Completion rates	
	Completes[2]	Weighted	Unweighted	Completes[2]	Weighted	Unweighted
Percent non-White enrolled						
0 – 10	3,502	95.0	98.1	3,515	95.5	98.5
11 – 49	3,633	89.8	95.4	3,674	91.1	96.4
50 – 89	1,864	90.4	92.8	1,920	92.9	95.6
90 – 100	1,916	92.4	94.1	1,996	96.2	98.0
Unknown	22	0.9	2.8	49	2.5	6.3
Region						
Northeast	2,008	87.7	93.5	2,048	90.4	95.3
Midwest	2,920	95.5	97.5	2,938	95.9	98.1
South	3,547	90.7	95.7	3,598	92.1	97.1
West	2,462	89.1	92.7	2,570	93.9	96.8

[1] School characteristics are for schools attended by children in the ECLS-K third-grade sample and are based on ECLS-K survey data, not data from the sampling frame.
[2] A completed questionnaire was defined as one that was not completely left blank.
SOURCE: U.S. Department of Education, National Center for Education Statistics, Early Childhood Longitudinal Study, Kindergarten Class of 1998–99 (ECLS-K), spring 2004.

Table 5-11. Number of completed child-level cases and child-level completion rates for the student records abstract and teacher-level questionnaire for children sampled in the base year, by school characteristics: School year 2003–04

School characteristics[1]	Student records abstract			Teacher-level questionnaire		
		Completion rates			Completion rates	
	Completes[2]	Weighted	Unweighted	Completes[2]	Weighted	Unweighted
All schools	10,015	69.9	82.1	10,872	79.9	90.6
School affiliation						
Public	8,177	82.3	86.9	8,849	90.2	94.1
Private	1,838	84.6	87.6	2,023	93.6	96.5
Catholic	1,209	90.3	90.5	1,313	97.0	98.3
Other private	629	77.9	82.7	710	89.4	93.3
Type of locale						
Large city	1,585	78.9	83.1	1,758	89.8	92.1
Mid-size city	1,688	85.3	89.1	1,819	94.1	96.0
Urban fringe of large city	2,778	76.8	81.6	3,131	87.4	92.0
Urban fringe of mid-size city	698	82.1	88.1	751	90.2	94.8
Large town	272	93.7	95.8	283	99.7	99.6
Small town	758	85.7	90.8	819	91.1	98.1
Rural – outside MSA	1,227	89.3	93.4	1,273	91.0	96.9
Rural – inside MSA	884	89.4	95.3	903	93.4	97.3
Unknown	125	8.8	14.9	135	12.0	21.3
School size (total enrollment)						
1 to 299	2,114	85.2	88.0	2,340	94.8	97.4
300 to 499	3,385	84.1	89.6	3,628	91.3	96.0
500 to 749	2,808	83.3	86.5	3,039	91.4	93.7
750 or more	1,687	78.6	83.4	1,844	86.9	91.2
Unknown	21	1.3	2.8	21	1.7	3.9

See note at end of table.

Table 5-11. Number of completed child-level cases and child-level completion rates for the student records abstract and teacher-level questionnaire for children sampled in the base year, by school characteristics: School year 2003–04—Continued

School characteristics[1]	Student records abstract			Teacher-level questionnaire		
		Completion rates			Completion rates	
	Completes[2]	Weighted	Unweighted	Completes[2]	Weighted	Unweighted
Percent non-White enrolled						
0 – 10	3,274	89.0	91.7	3,465	93.9	97.1
11 – 49	3,303	81.4	86.7	3,608	89.2	94.7
50 – 89	1,686	80.2	84.0	1,854	90.5	92.3
90 – 100	1,712	80.2	84.0	1,902	91.4	93.4
Unknown	40	2.1	5.2	43	2.8	7.6
Region						
Northeast	1,635	71.8	76.1	1,970	86.9	91.7
Midwest	2,788	90.4	93.1	2,913	94.7	97.3
South	3,359	83.4	90.7	3,531	90.1	95.3
West	2,233	81.3	84.1	2,458	90.0	92.5

[1] School characteristics are for schools attended by children in the ECLS-K third-grade sample and are based on ECLS-K survey data, not data from the sampling frame.
[2] A completed questionnaire was defined as one that was not completely left blank.
SOURCE: U.S. Department of Education, National Center for Education Statistics, Early Childhood Longitudinal Study, Kindergarten Class of 1998–99 (ECLS-K), spring 2004.

Table 5-12. Number of completed child-level cases and child-level completion rates for the child-level teacher questionnaires for children sampled in the base year, by school characteristics: School year 2003–04

School characteristics[1]	Child-level reading teacher questionnaire			Child-level mathematics teacher questionnaire			Child-level science teacher questionnaire		
		Completion rates			Completion rates			Completion rates	
	Completes[2]	Weighted	Unweighted	Completes[2]	Weighted	Unweighted	Completes[2]	Weighted	Unweighted
All schools	10,793	79.3	90.0	5,339	78.1	89.3	5,405	79.5	89.9
School affiliation									
Public	8,780	89.5	93.3	4,357	87.9	92.6	4,376	89.8	93.1
Private	2,013	92.9	96.0	982	93.6	96.0	1,029	92.1	95.8
Catholic	1,306	96.7	97.8	635	96.5	97.2	670	96.7	98.1
Other private	707	88.5	92.9	347	90.2	93.8	359	86.6	91.8
Type of locale									
Large city	1,753	90.2	91.9	853	90.3	91.8	880	88.0	89.9
Mid-size city	1,811	93.6	95.6	893	90.3	94.3	898	94.0	94.7
Urban fringe of large city	3,111	86.5	91.4	1,547	84.4	90.4	1,551	87.9	91.7
Urban fringe of mid-size city	741	88.3	93.6	366	87.6	93.1	374	88.0	93.7
Large town	283	99.7	99.6	140	99.4	99.3	143	100.0	100.0
Small town	808	90.3	96.8	383	85.0	96.5	432	96.8	98.6
Rural – outside MSA	1,258	90.1	95.7	634	92.8	95.5	620	86.8	95.4
Rural – inside MSA	893	91.8	96.2	456	90.5	95.6	441	94.3	97.8
Unknown	135	12.0	21.3	67	13.1	20.9	66	10.5	21.0
School size (total enrollment)									
1 to 299	2,323	93.8	96.7	1,144	94.2	96.6	1,175	93.7	96.4
300 to 499	3,605	91.0	95.4	1,830	90.7	95.8	1,768	90.2	94.6
500 to 749	3,014	90.6	92.9	1,462	88.8	91.7	1,537	91.3	93.2
750 or more	1,830	85.7	90.5	892	81.7	88.0	917	88.0	90.9
Unknown	21	1.7	3.9	11	2.2	4.0	8	0.7	3.0

See notes at end of table.

Table 5-12. Number of completed child-level cases and child-level completion rates for the child-level teacher questionnaires for children sampled in the base year, by school characteristics: School year 2003–04—Continued

School characteristics[1]	Child-level reading teacher questionnaire			Child-level mathematics teacher questionnaire			Child-level science teacher questionnaire		
		Completion rates			Completion rates			Completion rates	
	Completes[2]	Weighted	Unweighted	Completes[2]	Weighted	Unweighted	Completes[2]	Weighted	Unweighted
Percent non-White enrolled									
0 – 10	3,444	93.7	96.5	1,695	93.2	96.3	1,745	94.1	96.5
11 – 49	3,584	88.1	94.1	1,791	87.2	93.8	1,793	89.0	94.4
50 – 89	1,823	89.1	90.8	919	85.8	89.7	893	90.8	90.8
90 – 100	1,899	91.3	93.2	911	90.4	91.7	956	88.8	91.6
Unknown	43	2.8	7.6	23	3.3	7.9	18	1.9	6.5
Region									
Northeast	1,956	86.7	91.1	967	84.6	90.6	969	87.0	89.6
Midwest	2,909	94.4	97.2	1,423	94.5	96.9	1,482	94.5	97.1
South	3,498	89.0	94.4	1,747	87.5	93.6	1,743	89.2	94.8
West	2,430	89.0	91.5	1,202	87.4	90.5	1,211	89.1	91.2

[1] School characteristics are for schools attended by children in the ECLS-K third-grade sample and are based on ECLS-K survey data, not data from the sampling frame.
[2] A completed questionnaire was defined as one that was not completely left blank.
SOURCE: U.S. Department of Education, National Center for Education Statistics, Early Childhood Longitudinal Study, Kindergarten Class of 1998–99 (ECLS-K), spring 2004.

level teacher questionnaires in table 5-12: 85.7 to 99.7 percent for reading, 81.7 to 99.4 percent for mathematics, and 86.8 to 100 percent for science. These rates are higher than in any previous years of the ECLS-K, in all likelihood due to the higher incentives employed in fifth grade.

As noted above, the rate at which the survey instruments were completed varies markedly by mover status and, within movers, by whether the child was located and followed. As shown in table 5-13, the completion rate for the child assessment was 98.2 percent for children still enrolled in their base year school. For movers it dropped by about 6 points to 91.9 percent for those who were located and followed, and for those not located or followed due to a move to a non-ECLS-K PSU, it was zero. The parent interview completion rates varied from 91.6 percent for nonmovers to 87.1 percent for movers who were located and followed for the purposes of the child assessment, to 85.7 percent for movers who could either not be located or were not followed for the purposes of the child assessment. Even though children who had moved to a non-ECLS-K PSU were not administered the child assessment, a parent interview was conducted by telephone wherever possible, leading to the 86 percent response rate for this category.

The school administrator questionnaire completion rate is 15 points lower for movers, even when the children were located and followed; for the facilities checklist, it is 14 points lower (table 5-14). There are several reasons for this difference: located movers were not always assessed in schools; new schools in which movers enrolled had a lower level of commitment to the ECLS-K and often refused to complete the school administrator questionnaire; and some of these schools were contacted too late in the school year for them to consider completing it. The completion rate for nonmovers was 97.1 percent for the school administrator questionnaire and 98.8 percent for the facilities checklist. For located and followed movers it was 82.4 and 84.8 percent for the school administrator questionnaire and for the facilities checklist, respectively. The rates for the student records abstract are 90.1 percent for nonmovers and 72.3 percent for movers who were located and followed (table 5-15).

The teacher-level questionnaire completion rate, as shown in table 5-15, is about 14 points lower for movers who could be located and followed (82.2 percent) than for nonmovers (96.5 percent). Movers who could not be located were all nonrespondents for this instrument, pulling the overall completion rate for movers downward to 63.9 percent.

Table 5-13. Number of completed child-level cases and child-level completion rates for the child assessment and parent interview for children sampled in the base year, by child's mover status: School year 2003–04

Mover status[1]	Child assessment			Parent interview		
		Completion rates			Completion rates	
	Completes[2]	Weighted	Unweighted	Completes[3]	Weighted	Unweighted
All children	11,260	84.7	93.6	10,913	89.1	90.7
Mover status						
Mover	1,814	71.8	74.7	2,094	86.8	86.2
Located, followed	1,814	91.9	92.1	1,704	87.1	86.5
Other[4]	0	0.0	0.0	390	85.7	85.0
Nonmover	9,446	98.2	98.4	8,819	91.6	91.9

[1] This is the mover status used in weighting, which does not consider children who moved into identified destination schools as movers. A destination school is a school that received at least four students from the school where they had just completed the highest grade.
[2] Reading, mathematics, or science assessment was scorable, or child was disabled and could not be assessed.
[3] Family structure portion of parent interview was completed.
[4] This category includes movers who could not be located, and movers who moved into nonsampled PSUs.
SOURCE: U.S. Department of Education, National Center for Education Statistics, Early Childhood Longitudinal Study, Kindergarten Class of 1998–99 (ECLS-K), spring 2004.

Table 5-14. Number of completed child-level cases and child-level completion rates for the school administrator questionnaire and facilities checklist for children sampled in the base year, by child's mover status: School year 2003–04

Mover status[1]	School administrator questionnaire			Facilities checklist		
		Completion rates			Completion rates	
	Completes[2]	Weighted	Unweighted	Completes[2]	Weighted	Unweighted
All children	10,937	77.1	89.6	11,154	78.8	91.4
Mover status						
Mover	1,589	59.2	61.1	1,647	61.0	63.3
Located, followed	1,589	82.4	82.1	1,647	84.8	85.1
Other[3]	0	0.0	0.0	0	0.0	0.0
Nonmover	9,348	97.1	97.4	9,507	98.8	99.0

[1] This is the mover status used in weighting, which does not consider children who moved into identified destination schools as movers. A destination school is a school that received at least four students from the school where they had just completed the highest grade.
[2] A completed questionnaire was defined as one that was not completely left blank.
[3] This category includes movers who could not be located, and movers who moved into nonsampled PSUs.

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

Table 5-15. Number of completed child-level cases and child-level completion rates for the student records abstract and teacher-level questionnaire for children sampled in the base year, by child's mover status: School year 2003–04

Mover status[1]	Student records abstract			Teacher-level questionnaire		
		Completion rates			Completion rates	
	Completes[2]	Weighted	Unweighted	Completes[2]	Weighted	Unweighted
All children	10,015	69.9	82.1	10,872	79.9	90.6
Mover status						
Mover	1,393	52.0	53.6	1,587	63.9	66.3
Located, followed	1,393	72.3	72.0	1,587	82.2	82.0
Other[3]	0	0.0	0.0	0	0.0	0.0
Nonmover	8,622	90.1	89.8	9,285	96.5	96.7

[1] This is the mover status used in weighting, which does not consider children who moved into identified destination schools as movers. A destination school is a school that received at least four students from the school where they had just completed the highest grade.
[2] A completed questionnaire was defined as one that was not completely left blank.
[3] This category includes movers who could not be located, and movers who moved into nonsampled PSUs.
SOURCE: U.S. Department of Education, National Center for Education Statistics, Early Childhood Longitudinal Study, Kindergarten Class of 1998–99 (ECLS-K), spring 2004.

Table 5-16 shows the completion rates for all three child-level teacher questionnaires. These rates are between 95 and 96 percent for nonmovers, and between 80 and 82 percent for movers who were located and followed. Children who could not be located were all nonrespondents for the child-level teacher instruments. The reasons for lower completion rates from teachers if the child moved are similar to the reasons that affected the school administrator questionnaire and facilities checklist completion rates for movers.

Tables 5-17 to 5-20 present child-level weighted and unweighted completion rates for the spring-fifth grade data collection for children who were sampled as part of the kindergarten cohort in the base year, this time broken out by child characteristics. When the "unknown" categories are not included, the differences in completion rates by sex and by year of birth are inconsequential, but for race and ethnicity they are more substantial. Table 5-17 shows that for the child assessment the completion rate was highest for Asians (87.6 percent) and lowest for American Indians or Alaska Natives (78.3 percent). For the parent interview it is the opposite; the rate was highest for American Indians or Alaska Natives (95.2 percent) and lowest for Asian children (82.8 percent).

Table 5-18 shows that, excluding the "unknown" categories, the highest completion rates for the school administrator questionnaire and for the facilities checklist are for Pacific Islanders (85.7 percent and 86.7 percent, respectively), and the lowest completion rates are for American Indians or Alaska Natives (65.8 percent and 71.1 percent, respectively). Table 5-19 shows that the completion rate for the student records abstract is highest for children with "other" race (72.8 percent) and the lowest is for Black (63.9), excluding the "unknown" categories.

For the teacher-level questionnaires (table 5-19), the highest rate is for Pacific Islanders (84 percent) and the lowest rate is for American Indians or Alaska Natives (77 percent), excluding the "unknown" categories. For the child-level reading teacher questionnaire (table 5-20), the highest rate is for Asians (82 percent) and the lowest rate is for American Indians or Alaska Natives (76 percent). For the child-level mathematics teacher questionnaire, the highest rate is for Hispanic and for children whose race/ethnicity is not among the listed (80 percent) and the lowest rate is for Black and American Indians or Alaska Natives (76 percent). For the child-level science teacher questionnaire, the highest rate is for Asians (86 percent) and the lowest rate is for Blacks and for children whose race/ethnicity is not among the listed (76 percent).

In addition to the child assessment, parent interview, school administrator questionnaire, facilities checklist, student records abstract, and teacher questionnaires, whose completion rates have been summarized in the preceding tables, data were also collected in fifth grade from the special education teachers for children who followed individualized special education programs. Table 5-21 presents counts of completes and weighted and unweighted completion rates at the overall student level for the special education questionnaires A and B. The number of special education teacher questionnaires is small but their completion rates are high, 92.2 percent for part A, which captures teacher information, and 93.7 percent for part B, which relates to children who receive individualized special education services. These rates are not broken down by school and child characteristics because of the small sample sizes.

Table 5-16. Number of completed child-level cases and child-level completion rates for the child-level teacher questionnaires for children sampled in the base year, by child's mover status: School year 2003–04

Mover status[1]	Child-level reading teacher questionnaire			Child-level mathematics teacher questionnaire			Child-level science teacher questionnaire		
		Completion rates			Completion rates			Completion rates	
	Completes[2]	Weighted	Unweighted	Completes[2]	Weighted	Unweighted	Completes[2]	Weighted	Unweighted
All children	10,793	79.3	90.0	5,339	78.1	89.3	5,405	79.5	89.9
Mover status									
Mover	1,568	63.3	65.5	818	62.2	65.2	734	63.2	64.4
Located, followed	1,568	81.4	81.0	818	79.7	80.4	734	81.7	80.0
Other[3]	0	0.0	0.0	0	0.0	0.0	0	0.0	0.0
Nonmover	9,225	95.8	96.1	4,521	95.3	95.7	4,671	95.6	95.8

[1] This is the mover status used in weighting, which does not consider children who moved into identified destination schools as movers. A destination school is a school that received at least four students from the school where they had just completed the highest grade.
[2] A completed questionnaire was defined as one that was not completely left blank.
[3] This category includes movers who could not be located, and movers into nonsampled PSUs.
SOURCE: U.S. Department of Education, National Center for Education Statistics, Early Childhood Longitudinal Study, Kindergarten Class of 1998–99 (ECLS-K), spring 2004.

Table 5-17. Number of completed child-level cases and child-level completion rates for the child assessment and parent interview for children sampled in the base year, by child characteristics: School year 2003–04

Child characteristics[1]	Child assessment			Parent interview		
		Completion rates			Completion rates	
	Completes[2]	Weighted	Unweighted	Completes[3]	Weighted	Unweighted
All children	11,260	84.7	93.6	10,913	89.1	90.7
Sex[4]						
Male	5,675	84.1	93.3	5,525	89.9	90.8
Female	5,585	85.4	94.1	5,388	88.5	90.8
Race/ethnicity						
White, non-Hispanic	6,466	84.8	94.4	6,394	91.3	93.4
Black, non-Hispanic	1,273	83.3	93.3	1,151	83.4	84.3
Hispanic	2,093	85.9	92.4	2,036	89.5	89.9
Asian	788	87.6	92.5	707	82.8	83.0
Pacific Islander	144	85.4	92.3	136	87.7	87.2
American Indian or Alaska Native	210	78.3	92.1	222	95.2	97.4
Other	272	86.3	93.8	253	84.5	87.2
Unknown	14	55.3	50.0	14	58.0	50.0
Year of birth						
1992	3,307	83.9	93.6	3,211	88.7	90.8
1993	7,896	85.1	93.8	7,646	89.4	90.8
Other/unknown	57	71.0	76.0	56	76.0	74.7

[1] Based on ECLS-K survey data and not on data from the sampling frame.
[2] Reading, mathematics, or science assessment was scorable, or child was disabled and could not be assessed.
[3] Family structure portion of parent interview was completed.
[4] There is a small number of children whose gender is unknown and who did not have completed child assessment and parent interview. The completion rates for these children, being zero, are not included in the table.
SOURCE: U.S. Department of Education, National Center for Education Statistics, Early Childhood Longitudinal Study, Kindergarten Class of 1998–99 (ECLS-K), spring 2004.

Table 5-18. Number of completed child-level cases and child-level completion rates for the school administrator questionnaire and facilities checklist for children sampled in the base year, by child characteristics: School year 2003–04

Child characteristics[1]	School administrator questionnaire			Facilities checklist		
		Completion rates			Completion rates	
	Completes[2]	Weighted	Unweighted	Completes[2]	Weighted	Unweighted
All children	10,937	77.1	89.6	11,154	78.8	91.4
Sex[3]						
Male	5,517	77.1	89.5	5,621	78.6	91.2
Female	5,420	77.2	90.0	5,533	79.2	91.9
Race/ethnicity						
White, non-Hispanic	6,356	78.8	92.6	6,413	80.1	93.4
Black, non-Hispanic	1,256	74.5	89.9	1,266	75.4	90.6
Hispanic	1,957	75.4	83.5	2,061	79.0	88.0
Asian	752	77.9	85.1	776	80.9	87.8
Pacific Islander	145	85.7	92.4	146	86.7	93.0
American Indian or Alaska Native	191	65.8	82.3	208	71.1	89.7
Other	266	77.7	90.8	270	79.1	92.2
Unknown	14	50.1	46.7	14	50.1	46.7
Year of birth						
1992	3,223	77.0	90.0	3,275	78.7	91.5
1993	7,660	77.2	89.7	7,825	79.1	91.6
Other/unknown	54	64.0	70.1	54	59.5	70.1

[1] Based on ECLS-K survey data and not on data from the sampling frame.
[2] A completed questionnaire was defined as one that was not completely left blank.
[3] The completion of the school-level instruments does not depend on whether the child has completed assessment or parent interview data. Hence, while all children with completed assessment or parent interview data have known value of gender; there are children with completed school-level data whose gender is unknown.
SOURCE: U.S. Department of Education, National Center for Education Statistics, Early Childhood Longitudinal Study, Kindergarten Class of 1998–99 (ECLS-K), spring 2004.

Table 5-19. Number of completed child-level cases and child-level completion rates for the student records abstract and teacher-level questionnaire for children sampled in the base year, by child characteristics: School year 2003–04

Child characteristics[1]	Student records abstract			Teacher-level questionnaire		
		Completion rates			Completion rates	
	Completes[2]	Weighted	Unweighted	Completes[2]	Weighted	Unweighted
All children	10,015	69.9	82.1	10,872	79.9	90.6
Sex[3]						
Male	5,034	69.6	81.6	5,482	79.3	90.4
Female	4,981	70.4	82.7	5,390	80.7	91.0
Race/ethnicity						
White, non-Hispanic	5,878	72.6	85.6	6,300	80.6	92.3
Black, non-Hispanic	1,101	63.9	78.8	1,238	77.1	90.8
Hispanic	1,769	68.4	75.5	1,962	80.2	86.9
Asian	692	70.7	78.3	751	82.6	88.2
Pacific Islander	125	69.6	79.6	138	84.0	88.5
American Indian or Alaska Native	197	66.0	84.9	206	77.0	91.6
Other	242	72.8	82.6	263	81.9	91.3
Unknown	11	30.5	36.7	14	55.3	50.0
Year of birth						
1992	2,996	71.5	83.7	3,213	80.1	91.2
1993	6,976	69.4	81.7	7,610	80.1	90.6
Other/unknown	43	57.4	55.8	49	54.4	65.3

[1] Based on ECLS-K survey data and not on data from the sampling frame.
[2] A completed questionnaire was defined as one that was not completely left blank.
[3] The completion of the school-level instruments does not depend on whether the child has completed assessment or parent interview data. Hence, while all children with completed assessment and parent interview data have known value of gender; there is one child with completed facilities checklist data whose gender is unknown.
SOURCE: U.S. Department of Education, National Center for Education Statistics, Early Childhood Longitudinal Study, Kindergarten Class of 1998–99 (ECLS-K), spring 2004.

Table 5-20. Number of completed child-level cases and child-level completion rates for the child-level teacher questionnaires for children sampled in the base year, by child characteristics: School year 2003-04

Child characteristics[1]	Child-level reading teacher questionnaire			Child-level mathematics teacher questionnaire			Child-level science teacher questionnaire		
		Completion rates			Completion rates			Completion rates	
	Completes[2]	Weighted	Unweighted	Completes[2]	Weighted	Unweighted	Completes[2]	Weighted	Unweighted
All children	10,793	79.3	90.0	5,339	78.1	89.3	5,405	79.5	89.9
Sex[3]									
Male	5,436	78.7	89.7	2,670	78.0	88.9	2,739	78.2	89.5
Female	5,357	80.0	90.5	2,669	78.4	89.9	2,666	81.1	90.4
Race/ethnicity									
White, non-Hispanic	6,255	80.0	91.7	3,111	78.6	91.4	3,147	81.4	92.0
Black, non-Hispanic	1,231	76.7	90.2	591	75.7	88.5	629	76.2	90.4
Hispanic	1,952	79.5	86.4	959	79.6	85.2	963	77.2	84.9
Asian	743	81.8	87.3	362	77.8	85.8	378	85.1	88.1
Pacific Islander	132	79.1	84.6	75	77.5	85.2	59	81.1	86.8
American Indian or Alaska Native	204	76.4	90.7	109	75.6	91.6	93	77.0	87.7
Other	262	81.4	91.0	127	80.2	89.4	129	76.2	88.4
Unknown	14	55.3	50.0	5	47.5	35.7	7	31.2	50.0
Year of birth									
1992	3,186	79.1	90.4	1,565	77.9	90.1	1,618	79.9	90.5
1993	7,559	79.6	90.0	3,751	78.4	89.1	3,762	79.7	89.8
Other/unknown	48	53.5	64.0	23	55.0	63.9	25	52.5	64.1

[1] Based on ECLS-K survey data and not on data from the sampling frame.
[2] A completed questionnaire was defined as one that was not completely left blank.
[3] There is a small number of children whose gender is unknown and who did not have completed teacher questionnaire data. The completion rates for these children, being zero, are not included in the table.
SOURCE: U.S. Department of Education, National Center for Education Statistics, Early Childhood Longitudinal Study, Kindergarten Class of 1998–99 (ECLS-K), spring 2004.

Table 5-21. Number of completed instruments and child-level completion rates for the special education teacher questionnaires for children sampled in the base year: School year 2003–04

Category	Completes	Completion rates	
		Weighted	Unweighted
Special education part A[1]	960	92.2	93.8
Special education part B[1]	967	93.7	94.4

[1] A completed instrument was defined as one that was not completely left blank.
SOURCE: U.S. Department of Education, National Center for Education Statistics, Early Childhood Longitudinal Study, Kindergarten Class of 1998–99 (ECLS-K), spring 2004.

5.7.2 Children Sampled in First Grade

In spring-first grade the student sample was freshened to include first-graders who had no chance of selection in the base year because they had not attended kindergarten in the United States or had been in first grade in the fall of 1998. (For a detailed description of the freshening procedure see section 4.3.2.) This same group of children was followed into spring-fifth grade, unless they belonged in the excluded groups. Nonresponse in the freshened student sample could occur at two stages: during the procedure for sampling schools for freshening and identifying children to be used as freshening links in spring-first grade (first component) and then during data collection from the freshened children in spring-fifth grade (second component). The first component alone can further be decomposed into two sources: attrition due to the refusal of entire schools to implement the freshening procedure (*the school term*), and attrition because ECLS-K sampled children had moved to other schools (*the child term*). To contain costs, children who transferred from schools targeted for freshening were not used as links to identify freshened children, even when they were otherwise followed for data collection. These movers were considered freshening nonrespondents in the *child term*.

Table 5-22 presents weighted and unweighted completion rates for freshened children who were fielded in fifth grade. The two components of the completion rates are presented separately in table 5-22. The overall completion rates are the products of the two components. The first component is separated into *a school term* and *a child term* as described earlier. For this component, the completion rate is defined as the freshening completion rates, as opposed to the survey instrument completion rates found in the second component. The weighted freshening completion rate for children in schools targeted for freshening (*the school term*) is 67.2 percent. The reasons for non-participation in the freshening process included refusal or inability to provide the requested information in order to complete the procedures.

Within the schools that agreed to freshen, the freshening completion rate is 98.2 percent, the slight loss due to children who transferred to other schools (*the child term*). Multiplying these two terms together gives a first component completion rate of 66.0 percent. Note that the first component rate for spring-fifth grade is not identical to the first component rate for spring-first grade and spring-third grade because of the exclusion of children in special groups as explained in section 4.5.

The second component varies by survey instrument. The rates for the paper-and-pencil instruments range from 67.0 percent for the student records abstract to 100.0 percent for the special education questionnaire part A. The rate for the child assessment at 78.6 percent is about 6 points lower than for the kindergarten sample and the parent interview, at 81.9 percent, is about 7 points lower. The rates for the school instruments and the student records abstract are also lower than for the kindergarten sample, but by a smaller amount. The rates for some of the teacher instruments are higher than for the kindergarten sample. The final completion rate for each instrument is the product of the two components. Because of the low rates at the first stage, these range from a high of 53.6 percent for the teacher-level questionnaire to a low of 44.2 percent for the student records abstract.

5.7.3 Spring-Fifth Grade Completion Rates—All Children

Table 5-23 presents final fifth-grade completion rates for children sampled in kindergarten, children sampled in first grade, and all children combined. Because children sampled in first grade represent such a small fraction of the total population of children, their inclusion in the computation of the completion rate brings down the combined rate by less than one percent relative to the rates of children sampled in kindergarten, even though the completion rates for children sampled in first-grade rates are much lower than the kindergarten rates. The spring-fifth grade overall completion rates for the child assessment and the parent interview are 83.9 percent and 88.3 percent, respectively. These rates are higher than in third grade by about 4 percentage points for the child assessment and by about 11 percentage points for the parent interview. The unweighted completion rates are almost always higher than the weighted completion rates, by as much as 23 percent at the overall level. Where there is a large difference, it is due to movers who have larger weights and higher nonresponse rates than nonmovers. The weights of the movers were increased to account for the subsampling of movers. They also responded at a much lower rate than nonmovers, as shown earlier in table 5-6. This difference is larger than in previous years because movers in fifth grade have much larger weights than in previous years (many more movers were not included in fifth grade, necessitating larger adjustment factors). Note that the

unweighted completion rates follow the traditional ECLS-K pattern, that is, rates for the child assessment are higher than rates for the parent interview (93.4 percent for the child assessment and 90.5 percent for the parent interview). This is again due to movers with large weights and to the fact that there are more parent-responding movers than child-responding movers. Thus, the weighted completion rates are higher for the parent interview than for the child assessment.

Table 5-24 shows the completion rates for the parent interviews and the school and teacher instruments for children who have nonzero child weights (C6CW0>0). These are children whose fifth-grade reading, mathematics or science assessment were scorable, or children who could not be assessed because of disabilities. For these children, the completion rate for the child assessment should be 100 percent. The rate slightly less than 100 percent, shown when children sampled in kindergarten are combined with children sampled in first grade, is due to the school freshening nonresponse for children sampled in first grade.

Table 5-22. Number of completed child-level cases and child-level completion rates for children sampled in first grade: School year 2003-04

Category	Completes	Completion rates[1]	
		Weighted	Unweighted
First component (first-grade sample freshening)	7,032	66.0	79.1
School term[2]	7,089	67.2	79.7
Child term[3]	7,135	98.2	99.2
Second component (fifth-grade data collection)			
Child assessment[4]	86	78.6	86.0
Parent interview[5]	83	81.9	83.0
School administrator questionnaire[6]	86	74.4	83.5
Facilities check list[6]	89	77.8	86.4
Student records abstract[6]	75	67.0	72.8
Teacher-level questionnaire[6]	87	81.2	87.9
Reading teacher questionnaire (child level)[6]	84	79.8	84.8
Mathematics teacher questionnaire (child level)[6]	41	81.0	83.7
Science teacher questionnaire (child level)[6]	40	75.2	80.0
Special education part A[6]	15	100.0	100.0
Special education part B[6]	14	93.4	93.3
Overall completion rates			
Child assessment[4]	86	51.9	68.0
Parent interview[5]	83	54.0	65.6
School administrator questionnaire[6]	86	49.1	66.0
Facilities check list[6]	89	51.3	68.3
Student records abstract[6]	75	44.2	57.6
Teacher-level questionnaire[6]	87	53.6	69.5
Reading teacher questionnaire (child level)[6]	84	52.7	67.0
Mathematics teacher questionnaire (child level)[6]	41	53.5	66.2
Science teacher questionnaire (child level)[6]	40	49.6	63.2
Special education part A[6]	15	66.0	79.1
Special education part B[6]	14	61.6	73.8

[1] In the first component, this is the completion rate for freshening. In the second component, this is the completion rate for the survey instruments. The product of the two components is the overall completion rate for the survey instruments.
[2] The freshening completes and completion rates for children in schools targeted for freshening.
[3] The freshening completes and completion rates for children in schools that agreed to the freshening procedure.
[4] Reading, mathematics, or science assessment was scorable, or child was disabled and could not be assessed.
[5] Family structure portion of parent interview was completed.
[6] A completed questionnaire was defined as one that was not completely left blank.
SOURCE: U.S. Department of Education, National Center for Education Statistics, Early Childhood Longitudinal Study, Kindergarten Class of 1998–99 (ECLS-K), spring 2004.

Table 5-23. Number of completed child-level cases and child-level completion rates, for children sampled in kindergarten and first grade, by survey instruments: School year 2003-04

Survey instrument	Children sampled in kindergarten			Children sampled in first grade			All children		
		Completion rates			Completion rates			Completion rates	
	Completes	Weighted	Unweighted	Completes	Weighted	Unweighted	Completes	Weighted	Unweighted
Child assessment[1]	11,260	84.7	93.6	86	51.9	68.0	11,346	83.9	93.4
Parent interview[2]	10,913	89.1	90.7	83	54.0	65.6	10,996	88.3	90.5
School administrator questionnaire[3]	10,937	77.1	89.6	86	49.1	66.0	11,023	76.4	89.4
Facilities check list[3]	11,154	78.8	91.4	89	51.3	68.3	11,243	78.1	91.2
Student records abstract[3]	10,015	69.9	82.1	75	44.2	57.6	10,090	69.3	81.9
Teacher-level questionnaire[3]	10,872	79.9	90.6	87	53.6	69.5	10,959	79.3	90.4
Reading teacher questionnaire (child level)[3]	10,793	79.3	90.0	84	52.7	67.0	10,877	78.7	89.8
Mathematics teacher questionnaire (child level)[3]	5,339	78.1	89.3	41	53.5	66.2	5,380	77.5	89.1
Science teacher questionnaire (child level)[3]	5,405	79.5	89.9	40	49.6	63.2	5,445	78.8	89.7
Special education part A[3]	960	92.2	93.8	15	66.0	79.1	975	91.6	93.7
Special education part B[3]	967	93.7	94.4	14	61.6	73.8	981	92.9	94.2

[1] Reading, mathematics, or science assessment was scorable, or child was disabled and could not be assessed.
[2] Family structure portion of parent interview was completed.
[3] A completed questionnaire was defined as one that was not completely left blank.
SOURCE: U.S. Department of Education, National Center for Education Statistics, Early Childhood Longitudinal Study, Kindergarten Class of 1998–99 (ECLS-K), spring 2004.

Table 5-24. Number of completed child-level cases and child-level completion rates, for children with scorable reading, mathematics or science assessment or children not assessed due to disabilities, by survey instruments: School year 2003-04

Survey instrument	Children sampled in kindergarten			Children sampled in first grade			All children		
		Completion rates			Completion rates			Completion rates	
	Completes	Weighted	Unweighted	Completes	Weighted	Unweighted	Completes	Weighted	Unweighted
Child assessment[1]	11,260	100.0	100.0	86	78.3	86.8	11,346	99.5	99.9
Parent interview[2]	10,445	92.4	92.8	77	71.5	77.6	10,522	91.9	92.7
School administrator questionnaire[3]	10,794	93.4	96.1	83	74.7	83.7	10,877	93.0	96.0
Facilities check list[3]	11,015	95.7	98.0	86	78.3	86.8	11,101	95.3	97.9
Student records abstract[3]	9,986	85.6	88.9	74	68.3	74.6	10,060	85.2	88.8
Teacher-level questionnaire[3]	10,799	93.4	96.1	84	77.3	84.8	10,883	93.0	96.0
Reading teacher questionnaire (child level)[3]	10,774	93.3	95.9	84	77.3	84.8	10,858	92.9	95.8
Mathematics teacher questionnaire (child level)[3]	5,331	92.7	95.5	41	76.2	82.7	5,372	92.3	95.4
Science teacher questionnaire (child level)[3]	5,394	92.8	95.4	40	74.9	80.7	5,434	92.4	95.3
Special education part A[3]	947	93.0	94.1	14	78.3	86.8	961	92.6	94.0
Special education part B[3]	959	95.1	95.3	14	78.3	86.8	973	94.7	95.2

[1] Reading, mathematics, or science assessment was scorable, or child was disabled and could not be assessed.
[2] Family structure portion of parent interview was completed.
[3] A completed questionnaire was defined as one that was not completely left blank.
SOURCE: U.S. Department of Education, National Center for Education Statistics, Early Childhood Longitudinal Study, Kindergarten Class of 1998–99 (ECLS-K), spring 2004.

When the completion rates are conditioned on the presence of the child weight, they are at least 13 points higher than the unconditional completion rates for all instruments but the parent interview and the special education questionnaires. For these last two instruments, the difference between the number of completes for the conditional and unconditional rates is very small; hence the conditional rates are not affected as much as for the other instruments. For the parent interview, the unconditional rate is fairly high for the reason explained earlier that is, movers in fifth grade have much larger weights than in previous years and there are more parent-responding movers than child-responding movers (the weighted completion rates are higher for the parent interview than for the child assessment). This results in the smaller difference between conditional and unconditional of about 4 percent. For all the other instruments, the conditional completion rates are higher by 13.6 points for child-level science teacher questionnaire and as high as 17.2 points for the facilities checklist.

As explained in section 4.5, four groups of children were excluded from the fifth-grade data collection. These are (1) children who became ineligible in an earlier round (because they had died or had moved out of the country), (2) children who were subsampled out in previous rounds because they had moved out of the original schools and were not subsampled to be followed, (3) children whose parents emphatically refused to cooperate (hard refusals) in any of the data collection rounds since spring-kindergarten, and (4) children eligible for the third-grade data collection for whom there are neither first-grade nor third-grade data. Table 5-25 shows the completion rates for all instruments had children in the last two exclusion groups been counted as nonrespondents. These are children who would have been eligible for the fifth-grade collection but past experience showed that they would be most likely nonrespondents. When compared to table 5-23, the completion rates for all instruments in table 5-25 are lower as expected, but only by about 2 percent, with the smallest difference for the student records abstract and the largest difference for the parent interview. Note that the rates for mathematics and science teacher appear to be unchanged. Recall that only about half of the children had mathematics teacher questionnaires and the other half had science teacher questionnaires. Since the mathematics/ science sampling flags were not assigned to children not included in the sample, we were not able to compute a correct completion rate for these two instruments for these children. But the pattern of completion rates would be the same as for the other instruments and we would expect a drop of about 2 percent for the mathematics and science teacher questionnaire.

Table 5-25. Number of completed child-level cases and child-level completion rates, if excluded children were fielded, by survey instruments: School year 2003-04

Survey instrument	Children sampled in kindergarten			Children sampled in first grade			All children		
		Completion rates			Completion rates			Completion rates	
	Completes	Weighted	Unweighted	Completes	Weighted	Unweighted	Completes	Weighted	Unweighted
Child assessment[1]	11,260	82.7	86.7	86	47.6	57.7	11,346	81.9	86.5
Parent interview[2]	10,913	87.0	84.0	83	49.7	55.8	10,996	86.1	83.8
School administrator questionnaire[3]	10,937	75.3	83.3	86	45.2	56.3	11,023	74.6	83.1
Facilities check list[3]	11,154	77.0	84.9	89	47.3	58.2	11,243	76.3	84.7
Student records abstract[3]	10,015	68.3	76.3	75	40.7	49.1	10,090	67.6	76.1
Teacher-level questionnaire[3]	10,872	78.0	84.1	87	49.2	58.9	10,959	77.3	83.9
Reading teacher questionnaire (child level)[3]	10,793	77.4	83.5	84	48.3	56.8	10,877	76.7	83.3
Mathematics teacher questionnaire (child level)[3]	5,339	78.1	89.3	41	52.1	65.2	5,380	77.5	89.1
Science teacher questionnaire (child level)[3]	5,405	79.5	89.9	40	48.3	62.3	5,445	78.8	89.7
Special education part A[3]	960	92.2	93.8	15	64.3	77.9	975	91.5	93.7
Special education part B[3]	967	93.7	94.4	14	60.0	72.7	981	92.9	94.2

[1] Reading, mathematics, or science assessment was scorable, or child was disabled and could not be assessed.
[2] Family structure portion of parent interview was completed.
[3] A completed questionnaire was defined as one that was not completely left blank.
SOURCE: U.S. Department of Education, National Center for Education Statistics, Early Childhood Longitudinal Study, Kindergarten Class of 1998–99 (ECLS-K), spring 2004.

This page is intentionally left blank.