Child Care & Early Education
RESEARCH CONNECTIONS

ICPSR 4075

Early Childhood Longitudinal Study [United States]: Kindergarten Class of 1998-1999, Third Grade

User Guide

United States Department of Education
National Center for Education Statistics

First ICPSR Version
December 2004

ICPSR
P.O. Box 1248
Ann Arbor, Michigan 48106
www.icpsr.umich.edu

About *Research Connections*

These data are made available by the Child Care and Early Education *Research Connections* (CCEERC), which promotes high quality research in child care and early education and the use of that research in policymaking. Our vision is that children are well cared for and have rich learning experiences, and their families are supported and able to work.

Research Connections is a partnership among the National Center for Children in Poverty at the Mailman School of Public Health, Columbia University; the Inter-university Consortium for Political and Social Research at the Institute for Social Research, University of Michigan; and the Child Care Bureau, at the Administration for Children and Families, U.S. Department of Health and Human Services.

Terms of Use

Bibliographic Citation: Publications based on ICPSR data collections should acknowledge those sources by means of bibliographic citations. To ensure that such source attributions are captured for social science bibliographic utilities, citations must appear in footnotes or in the reference section of publications. The bibliographic citation for this data collection is:

> U.S. Dept. of Education, National Center for Education Statistics. EARLY CHILDHOOD LONGITUDINAL STUDY [UNITED STATES]: KINDERGARTEN CLASS OF 1998-1999, THIRD GRADE [Computer file]. ICPSR version. Washington, DC: U.S. Dept of Education, Institute of Education Sciences [producer], 2004. Ann Arbor, MI: Inter-university Consortium for Political and Social Research [distributor], 2004.

Request for Information on Use of ICPSR Resources: To provide funding agencies with essential information about use of archival resources and to facilitate the exchange of information about ICPSR participants' research activities, users of ICPSR data are requested to send to ICPSR bibliographic citations for each completed manuscript or thesis abstract. Visit the ICPSR Web site for more information on submitting citations.

Data Disclaimer: The original collector of the data, ICPSR, and the relevant funding agency bear no responsibility for uses of this collection or for interpretations or inferences based upon such uses.

Responsible Use Statement: In preparing data for public release, ICPSR performs a number of procedures to ensure that the identity of research subjects cannot be disclosed. Any intentional identification or disclosure of a person or establishment violates the assurances of confidentiality given to the providers of the information. Therefore, users of data obtained from the ICPSR archive and/or any of its special topic archives agree:

- To use these datasets solely for statistical analysis and reporting of aggregated information, and not for investigation of specific individuals or organizations, except when identification is authorized in writing by ICPSR
- To make no use of the identity of any person or establishment discovered inadvertently, and to advise ICPSR of any such discovery
- To produce no links among ICPSR datasets or among ICPSR data and other datasets that could identify individuals or organizations

Redistribution: ICPSR data may not be redistributed or sold to other individuals, institutions, or organizations without the written agreement of ICPSR.

Bibliographic Description

ICPSR Study No.:	4075
Title:	Early Childhood Longitudinal Study [United States]: Kindergarten Class of 1998-1999, Third Grade
Principal Investigator(s):	United States Department of Education. National Center for Education Statistics
Series:	Early Childhood Longitudinal Study (ECLS) Series
Bibliographic Citation:	U.S. Dept. of Education, National Center for Education Statistics. EARLY CHILDHOOD LONGITUDINAL STUDY [UNITED STATES]: KINDERGARTEN CLASS OF 1998-1999, THIRD GRADE [Computer file]. ICPSR version. Washington, DC: U.S. Dept of Education, Institute of Education Sciences [producer], 2004. Ann Arbor, MI: Inter-university Consortium for Political and Social Research [distributor], 2004.

Scope of Study

Summary:	The Early Childhood Longitudinal Study, Kindergarten Class of 1998-1999 (ECLS-K) focuses on children's early school experiences beginning with kindergarten through fifth grade. It is a nationally representative sample that collects information from children, their families, their teachers, and their schools. ECLS-K provides data about the effects of a wide range of family, school, community, and individual variables on children's development, early learning, and early performance in school. This data collection contains the wave of data collected in the spring of third grade (2002). The third-grade data collection includes information about the diversity of the study children, the schools they attended, and their academic progress in the years following kindergarten. Other variables include child gender, child race, family background, childcare, childcare arrangements, food security, hours per week in child care, socioeconomic status, household income, highest level of education for parents and students, parents' employment status, teachers' evaluation practice, and usefulness of different activities in the classroom.
Subject Terms:	birth, child care, child development, early childhood education, elementary education, family life, infants, kindergarten, mathematics, minority groups, preschool children, reading skills, school, school age children, school readiness
Geographic Coverage:	United States
Time Period:	Spring 2001-2002
Date of Collection:	Spring 2002

Universe: Children and their families, teachers, and schools in the United States.

Data Type: survey data

Data Collection Notes: (1) Since the sample for the third-grade data is collected from all of the base-year respondents and children who were brought into the sample in the spring-first-grade wave, the documentation for the base-year and first-grade data are included as part of this collection. The data for the base year and first grade are available separately as EARLY CHILDHOOD LONGITUDINAL STUDY [UNITED STATES]: KINDERGARTEN CLASS OF 1998-1999 (ICPSR 3676). (2) The codebooks, data collection instruments, and user guide are provided by the data producer as Portable Document Format (PDF) files. The PDF file format was developed by Adobe Systems Incorporated and can be accessed using PDF reader software, such as the Adobe Acrobat Reader. Information on how to obtain a copy of the Acrobat Reader is provided on the ICPSR Web site.

Methodology

Sample: ECLS-K utilized a multistage probability sample design to select a nationally representative sample of children attending kindergarten in 1998-1999. The Third Grade sample consisted of all children who were base year respondents and children who were brought into the sample in spring-first-grade wave through the sample freshening procedure described in section 4.3 of the manual and their families, teachers, and schools. The first-grade data collection targeted base-year respondents, in which 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, the effort for 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 1998-1999 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. As in the first-grade data collection, only a subsample of students who had transferred from their kindergarten school was followed. 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) and who moved for the first time in 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 Third Grade User Manual is provided as part of the documentation for this collection.

Data Source: Computer-assisted telephone interviewing (CATI) for the parent interviews, computer-assisted personal interview (CAPI) and self-description questionnaires for child assessment, and self-administered questionnaires for the teachers and the school administrators

Extent of Processing: DDEF.PR/ CONCHK.PR/ UNDOCCHK.PR/ UNDOCCHK.ICPSR/ REFORM.DATA

Access and Availability

Extent of Collection: 3 data files + machine-readable documentation (PDF) + SAS data definition statements + SPSS data definition statements + Stata data definition statements + data collection instruments (PDF)

Data Format: Logical Record Length with SAS, SPSS, and Stata data definition statements, SAS transport file, and SPSS portable file

File Specifications

Part No.	*Part Name*	*File Structure*	*Case Count*	*Variable Count*	*LRECL*	*Records Per Case*
1	Third Grade Child Data	rectangular	15,305	3,225	5,798	2
2	Round 5 Base Weights and Adjustment Factors	rectangular	21,357	113	1,159	1
3	Cross-Round Weight Status	rectangular	22,978	16	37	1

Early Childhood Longitudinal Study

Kindergarten Class of 1998-99

USER'S MANUAL FOR THE ECLS-K BASE YEAR PUBLIC-USE DATA FILES AND ELECTRONIC CODEBOOK

NCES 2001-029 (revised)

Prepared by:

Westat
Rockville, Maryland
Karen Tourangeau
John Burke
Thanh Le
Siu Wan
Margaret Weant
Eugene Brown
Nancy Vaden-Kiernan
Elizabeth Rinker
Richard Dulaney
Kirsten Ellingsen
Brandon Barrett
Ismael Flores-Cervantes
Nicholas Zill

Educational Testing Service
Princeton, New Jersey
Judith Pollack
Donald Rock

University of Michigan
School of Education
Ann Arbor, Michigan
Sally Atkins-Burnett
Samuel Meisels

The National Center for Education Statistics
U.S. Department of Education
Washington, DC
Jonaki Bose
Jerry West

Education Statistics Services Institute
Washington, DC
Kristen Denton
Amy Rathbun
Jill Walston

TABLE OF CONTENTS

TABLE OF CONTENTS (continued)

TABLE OF CONTENTS (continued)

TABLE OF CONTENTS (continued)

TABLE OF CONTENTS (continued)

List of Tables

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)

TABLE OF CONTENTS (continued)

List of Tables (continued)

TABLE OF CONTENTS (continued)

List of Exhibits

Exhibit

1. INTRODUCTION

The Early Childhood Longitudinal Study-Kindergarten Class of 1998-99 (ECLS-K) is a multisource, multimethod study that focuses on children’s early school experiences beginning with kindergarten. 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 the Survey Research Center and the School of Education at the University of Michigan and Educational Testing Services (ETS).

The ECLS-K is following a nationally representative cohort of children from kindergarten through fifth grade. A total of 21,260 children throughout the country participated by having a child assessment or parent interview in the fall and/or spring of kindergarten. The children were in kindergarten when sampled. Base year data were collected in the fall of 1998 and spring of 1999. Four waves of data collection are planned beyond kindergarten: fall and spring first grade, and spring third and fifth grades. All data collection will be completed in the spring of 2004 when most of the children will be in fifth grade.

The ECLS-K has several major objectives and numerous potential applications. The ECLS-K combines elements of (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) a cross-sectional study 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) will follow a national sample of children, born in the year 2001, from birth through first grade. ECLS-B will focus on the characteristics of children and their families that influence children’s first experiences with the demands of formal school, as well as children’s early health care and in- and out-of-home experiences. Together these cohorts will provide the range and breadth of data required to more fully describe and understand children’s health, early learning, development, and education experiences.

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

1.1 Background

National policymakers and the public at large have increasingly recognized that the prosperity of the United States depends on the successful functioning of the American education system. There is also growing awareness that school reform efforts cannot focus solely on the secondary and postsecondary years but must pay attention to the elementary and preschool years as well. Increased policy interest in the early grades is reflected in an intensified recent national policy aimed at ensuring that children are capable of reading by the third grade, providing college student and adult volunteer tutors for children who are having difficulty learning to read, and increasing the number of children served by Head Start to 1 million by the year 2002.

Efforts to expand and improve early education will benefit from insights gained through analyses of data from the large-scale, nationally representative, longitudinal ECLS-K database. The ECLS-K database contains information about the types of preschool and elementary 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 affected by 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. This produces a more accurate measurement of these antecedent factors and enables stronger causal inferences to be made about their relationship to later academic progress.

The ECLS-K enables educational policy analysts to use an ecological perspective on early childhood education, using techniques such as multilevel modeling to study how school and classroom factors affect the progress of individual children. The data collected will enable analysts to examine how children's status at school entry and performance in school are jointly determined by an interaction of child characteristics and family and school environments.

Data collected during the kindergarten year can serve as baseline measures to examine how schooling shapes later individual development. 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 school experiences in kindergarten and the early grades.

A goal of the kindergarten data collection has been to describe accurately the diversity of kindergarten children and the programs they attend. For instance, national data are available for the first time on public and private kindergarten programs and the children who attend them. 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 and approaches 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 interaction 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, socio-emotional and intellectual development and incorporates factors from the child's family, community, and school-classroom environments. The conceptual model is presented in figure 1-1. 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 primary 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.

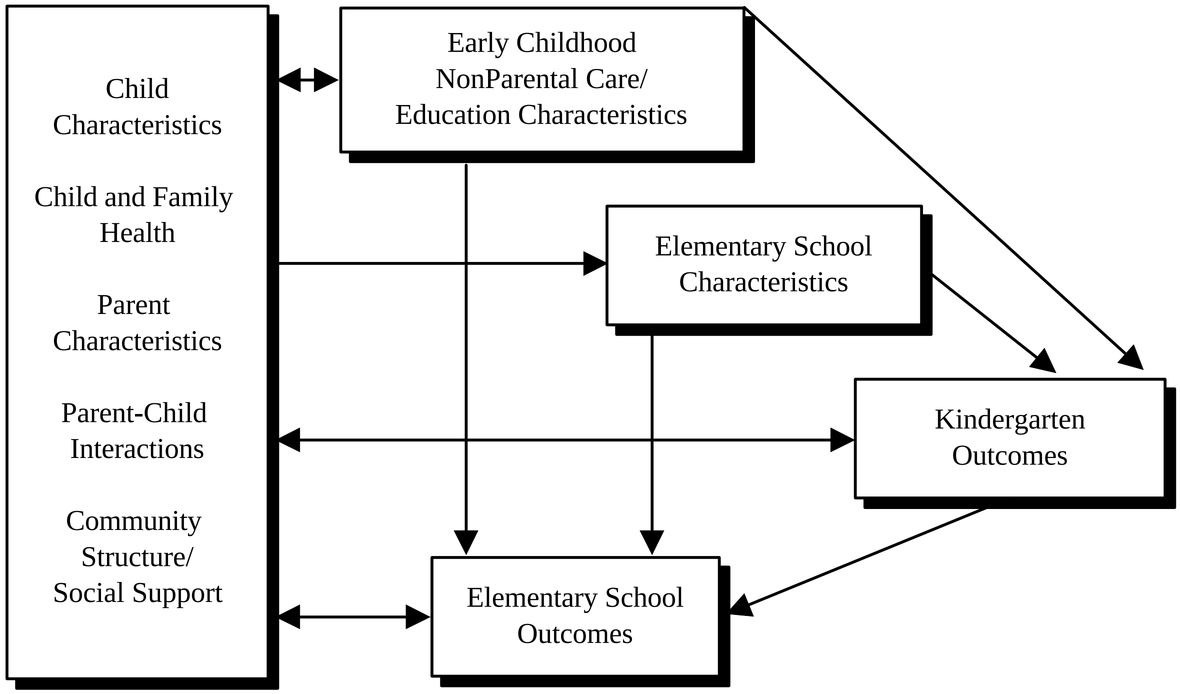


Figure 1-1. ECLS-K conceptual model

1.3 Study Components

The emphasis that is being placed on the whole of the child's environments and development has critical implications for the design of the ECLS-K. The design of the study allows for the collection of data from the child, the child's parents/guardians, teachers, and schools.

- **Children** are asked to 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 (e.g., general knowledge, literacy, and quantitative skills) and noncognitive (e.g., fine motor and gross motor coordination, socio-emotional) skills and knowledge. All measures of a child's cognitive skills are obtained through an untimed one-on-one assessment of the child. During later years of the ECLS-K, children will report on their own experiences in and out of school.

- **Parents/guardians** are an important source of information about the families of the children selected for the study and about themselves. They are asked to provide key information about their children, especially during the first years of the study. Parents are one of the important sources of 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 telephone interviewing (CATI) [or computer-assisted personal interviewing (CAPI) if they do not have a telephone].

- **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 upon 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. Teachers complete self-administered questionnaires each time children are assessed, with the exception of the fall first grade data collection.

- **School administrators** are asked to complete self-administered questionnaires during the spring data collection. They 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.

1.4 Contents of Manual

This manual provides documentation and guidance for users of the three public-use data files of the ECLS-K: the child file, teacher file, and the school file. The manual contains information about the data collection instruments (chapter 2) and the psychometric properties of these instruments (chapter 3). The manual describes the ECLS-K sample design (chapter 4); data collection procedures and response rates (chapter 5); and data processing procedures (chapter 6). In addition, the manual shows how the public-use data file is structured, provides definitions of composite variables (chapter 7), and explains how to use the Electronic Code Book (chapter 8). The Electronic Code Book contains unweighted frequencies for all variables.

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 employ 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.

1.5 Differences Between the ECLS-K Restricted-Use Base Year Files and the ECLS-K Base Year Public-Use Files

In preparing the ECLS-K Base Year Public-Use data files the National Center for Education Statistics (NCES) has taken steps to minimize the likelihood that an individual school, teacher, parent or child participating in the study can be identified. This is in compliance with the Privacy Act of 1974 and the National Education Statistics Act of 1994, both of which mandate the protection of the confidentiality of respondents. The process began with the ECLS-K Restricted-Use Base Year data files, which underwent a formal disclosure risk analysis. Variables identified as posing the greatest disclosure risk were altered, and in some instances entirely suppressed, and in this way the ECLS-K Base Year Public-Use data files were created. Every effort has been made to alter the files as little as possible, consistent with the requirement for confidentiality protection. After altering the variables the disclosure risk analysis was repeated to verify that the disclosure risk had been reduced to acceptable levels.

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

- Outlier values were top- or bottom- coded;
- Individual cases for which a particular variable posed an especially high risk for disclosure had the value of that variable altered (usually by no more than 5 to 10 percent) to reduce the risk;
- Some continuous variables were modified into categorical variables, and certain categorical variables had their categories collapsed; and
- Certain variables with too few cases and a sparse distribution were suppressed altogether, rather than modified.

A comprehensive list of the variables that have been altered or suppressed can be found in section 7.9.

Both the public- and restricted-use files provide data at the individual child, teacher, and school levels. The modifications that were implemented to avoid the identification of schools, teachers, 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 have been suppressed, there are a few users who might require the restricted files. Those researchers examining certain rare subpopulations such as the disabled, or children with specific non-English home languages or countries of birth and those interested in examining the type and number of hours of kindergarten programs offered in schools will find that the restricted-use files contain a few more variables. 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 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.

2. DESCRIPTION OF DATA COLLECTION INSTRUMENTS

This chapter describes the instruments used to collect base year data in the fall of 1998 and spring of 1999 for the Early Childhood Longitudinal Study-Kindergarten Class of 1998-99 (ECLS-K). In the fall, data were collected by teacher questionnaires, parent computer-assisted interviews (CAIs), and direct child assessments. Most of the fall instruments were repeated for the spring of 1999 data collection, although the parent and teacher measures varied by content between the two data collection points to ease respondent burden. This chapter also describes the Head Start verification study, where part of the data was collected in the fall parent interview and student record abstract. Table 2-1 below lists all of the instruments used in each of the two rounds of data collection.

Table 2-1. Instruments used in the fall and spring ECLS-K

Fall-Kindergarten	Spring-Kindergarten
Parent Interview	Parent Interview
Child Assessment	Child Assessment
Teacher Questionnaire-Part A	Teacher Questionnaire-Part A
Teacher Questionnaire-Part B	Teacher Questionnaire-Part B
Teacher Questionnaire-Part C	Teacher Questionnaire-Part C
	Special Education Teacher Questionnaire-Part A
	Special Education Teacher Questionnaire-Part B
	Adaptive Behavior Scale
	School Administrator Questionnaire
	Student Record Abstracts
	School Facilities Checklist
	Salary and Benefits Questionnaire

Appendix A contains a copy of the base year instruments, except for the direct child assessment, the social rating scale[1] in the parent interview and teacher questionnaire, and the adaptive

[1] Adapted with permission from Elementary Scale A ("How Often?"), Frank M. Gresham and Stephen N. Elliott ©1990, American Guidance Service, Inc.

behavior scale.[2] These latter measures contain copyright protected materials and agreements with the test publishers that restrict their distribution.

2.1 The Direct Child Assessments

Fall Child Assessments. In the fall of the base year, one-on-one child assessments were conducted with the sampled children. This assessment included cognitive, psychomotor, and physical components. The assessment took approximately 50-70 minutes and was designed to provide data on the developmental status of children in the United States at the start of their formal schooling. The ECLS-K cognitive assessment battery consisted of questions in three subject areas: language and literacy, mathematical thinking, and general knowledge. Psychomotor assessments were also included in the fall, along with assessments of the child’s height and weight.

The assessment began by verifying the child’s name and administering a set of warm-up exercises similar in form to the items used to administer the cognitive component. Prior to administering the cognitive assessment battery, a language-screening assessment was administered to those children identified from their school records (or by their teacher, if no school records were available) as coming from a language minority background (meaning that their primary home language was not English). This screening test was used to determine if a child was able to understand and respond to the cognitive assessment items in English.

The Language Screener. Efforts were made to include children who spoke a language other than English in the ECLS-K assessment. Field supervisors either checked the school records to determine children’s home language or, if records were not available, requested this information directly from children’s teachers. (See chapter 5 for a complete description of this process.) A brief language screener, the Oral Language Development Scale (OLDS), was given to those children who had a non-English language background. The screener determined if children understood English well enough to receive the direct child assessment in English. Children who passed the language screener received the full ECLS-K direct assessment battery. Children who did not pass an established cut score on the language screener received a reduced version of the ECLS-K assessments.

[2] Lambert, Nadine, Nihira, Kazuo, and Leland, Henry, Adaptive Behavior Scale-Second Edition, ©1993, The American Association on Mental Retardation.

The OLDS was also used to capture baseline information on children whose primary language was not English. The baseline data for these children can be used by researchers to examine English language acquisition over time.

The OLDS measured children's listening comprehension, vocabulary, and ability to understand and produce language. The OLDS consisted of three parts extrapolated from the preLAS 2000 (Duncan, S.E. and De Avila, E.A., 1998[3]). For the OLDS, children participated in Simon Says, Art Show, and Let's Tell Stories. Part one, Simon Says, measured listening comprehension of basic English instructions (i.e., asking a child to do things such as touch ear, pick up paper, or knock on table). Part two, Art Show, was a picture vocabulary assessment where children were asked to name pictures they were shown. The Art Show served as an assessment of a child's ability to produce language and measured the child's command of expressive language. The final part of the OLDS, Let's Tell Stories, was used to obtain a sample of a child's natural speech by asking a child to retell a story read by the assessor. The child was read two different stories and asked to tell what happened using pictures as prompts. The assessor recorded on paper exactly what the child said and scored the story using the established preLAS 2000 scoring rules. The scores assigned were based on the complexity of the child's sentence structure and vocabulary in his or her retelling of the story. These scores provide researchers with a direct measure of oral language performance.

Children who passed the language screener received the full English direct assessment. Certain components of the direct child assessment could also be conducted in Spanish. If a child did not pass the language screener but spoke Spanish, he or she was administered a Spanish translated form of the mathematics assessment and an alternate form of the language screener, the Spanish version of the Oral Language Development Scale (Spanish OLDS), as well as a psychomotor assessment that was conducted in Spanish. The Spanish OLDS that was administered is similar in content to the English OLDS and measures the same constructs.

A variety of steps were undertaken to confirm that the scores obtained from the Spanish mathematics assessment would be comparable to those for the English version. After the test items were translated into Spanish, a back-translation was carried out, followed by a review of the assessment instrument by two Spanish-speaking math experts. Psychometric analyses were also performed to compare the English and Spanish mathematics test results. Differential Item Functioning (DIF)

[3] Duncan, S.E. and De Avila, E. A., preLAS 2000 Cue Picture Book English Form C, CTB/McGraw-Hill Companies, Inc., 1998

procedures were carried out to determine whether the relative difficulty of each of the mathematics items was comparable in the English and Spanish version. Of the 31 mathematics items with sufficient numbers of observations for analysis, 16 appeared to be relatively easier for the Spanish speaking children and 15 for the English version. Of these 31 DIF statistics, 28 showed differences that were slight and/or not statistically significant. The three items marked as having "C" level DIF (sizeable difference in performance, and statistical significance) were split between the groups, with one item found to be relatively easier on the Spanish test and two relatively easier in English. A finding of differential item functioning does not necessarily indicate bias in the test; it may simply be an indication that achievement differences among groups may be more pronounced for certain skills than for others. A review of the "C" level test items by developmental and Spanish-language experts found no evidence of bias in either the content or the translation of the items. When Item Response Theory (IRT) procedures were carried out to calculate scaled scores, plots showing the fit of data to item parameters were examined for the Spanish compared to the English mathematics tests. The results were very similar to the DIF findings: there was an essentially identical fit for almost all of the test items; there were small differences in one direction or the other for a few items; and there was no evidence of systematic bias. All of these analyses support the conclusion that the language of administration had little or no impact on the scores obtained.

Children who did not pass the established cut score on the language screener and whose native language was not Spanish were excluded from the assessment; however, assessors collected physical measurements of these children's height and weight. Table 2-2 shows the paths of the direct child assessment by home language and scores on the English OLDS.

Table 2-2. Flow of the fall direct child assessment

Home Language	English OLDS	Spanish OLDS	Warm-up Booklet	Reading	Math	Psycho-motor	General Knowledge	Height/ Weight
English			✓	✓	✓	✓	✓	✓
Other	✓ Score **at or Above** Cutpoint		✓	✓	✓	✓	✓	✓
	✓ Score **Below** Cutpoint Speaks Spanish	✓	✓ **Spanish**		✓ **Spanish**	✓ **Spanish**		✓
	✓ Score **Below** Cutpoint Doesn't Speak Spanish							✓

Two-Stage Assessment Design

The direct cognitive assessment consisted of a set of two-stage assessments: a first-stage routing section for each of the three subject areas, followed by several alternative second-stage forms. The same reading, mathematics, and general knowledge routing sections, consisting of 12 to 20 items with a broad range of difficulty, were administered to all children. A child's performance on the routing section determined the second-stage form that was administered. The reading and mathematics assessments had low, middle, and high difficulty second-stage options, while the general knowledge assessment had two second-stage alternatives. The purpose of this adaptive assessment design was to maximize accuracy of measurement and minimize administration time.

The second-stage forms varied by level of difficulty so that a child would be administered questions appropriate to his or her current level of ability for each cognitive domain. Administering assessment items that are too hard for a particular child not only causes frustration and distress but also provides very little information on the precise level of the child's ability. Because most of the items are likely to be answered incorrectly, all that can be concluded is that the child's ability level is below the difficulty level of the questions, but there is no information on *how much* below. Similarly, giving a larger number of very *easy* items to a child of *high* ability may be boring and, again, they are not very useful in pinpointing the child's achievement level. The assessment items that provide the best information are those that are slightly too easy or slightly too hard for an individual. The pattern of right and wrong responses on such items makes it possible to estimate ability within a narrow range. The number of questions included in this assessment was limited in order to minimize the time and burden on the children. Consequently, it was important to match the difficulty of the questions to the ability level of the children, to the extent that this was possible with preselected sets of items. The routing section provided a rough estimate of each child's achievement level, so that a second-stage form with items of the appropriate difficulty for maximizing measurement accuracy could be selected.

The cognitive assessment included both multiple choice and open-ended items. For ease of administration, questions of similar format were grouped together in order of increasing difficulty within each group. When the question format changed, practice items were used to introduce children to the new format. Assessments were shortened or discontinued if the administrator perceived that the child was uncomfortable or distressed about responding to the assessment items. When a child did not respond to a question, the assessor repeated the question. If there was still no response, or the child did not know the answer, the assessor entered a code for “don’t know” and moved on to the next question, while periodically reminding the child to try. For a child who did not respond to several questions in a row, the assessor still waited ten seconds after reading each question and entered a code for “don’t know,” but did not continue to repeat each question. If no response was given to ten questions in a row, assessors entered a “refuse” code into the computer for the remainder of the items in that subject area, without reading the questions, until reaching the next subject area, where he or she resumed reading the questions. This procedure was used to give children who did not want to respond to questions in one subject area (e.g., reading) a chance to respond to questions in another subject area (e.g., math). Scores in each subject area were computed only if at least ten questions were answered in the combined first and second stages.

Cognitive Components. The cognitive assessment focused on three general areas of competence: (1) language use and literacy (reading); (2) mathematics; and (3) knowledge of the social and physical world, referred to as “general knowledge.” The assessment did not ask the children to write anything or to explain their reasoning; rather, they used pointing or verbal responses to complete the tasks. The data were collected using a computer-assisted interviewing methodology. The assessment included the use of a small easel with pictures, letters of the alphabet, words, short sentences, numbers, or number problems. A brief description of the three components of the cognitive assessment follows.

Language and Literacy. The language and literacy (reading) assessment included questions designed to measure basic skills (print familiarity, letter recognition, beginning and ending sounds, rhyming sounds, 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.

The reading assessment contained five proficiency 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 that comprised the earlier levels as well. These five levels were: (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 words by sight; and (5) reading word in context.

Mathematical Thinking. The mathematics assessment items were 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. Each of the mathematics assessment forms 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 for the appropriate parts of the assessment.

The items in the mathematics assessment could also be grouped into five-level proficiency levels or progression of skills, though the math clusters were less homogeneous in content than the reading clusters. The clusters of math items included: (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 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.

General Knowledge. The general knowledge assessment consisted of science and social studies material. The science items measure two broad classes of science competencies: a) conceptual understanding of scientific facts, and b) skills and abilities to form questions about the natural world, to try to answer them on the basis of the tools and the evidence collected, and to communicate answer and how the answers were obtained. Social studies material included questions relating to history/government, culture, geography, and economics. The assessment items drew on children's experiences with their environment, and many questions related to more than one of the categories.

The subject matter of the assessment of general knowledge was too diverse and insufficiently ranked or graded to permit formation of a set of proficiency levels. A score was calculated to represent each child's breadth and depth of understanding of the world around them. This assessment 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 also measured.

Physical Components. There were two parts to the physical component of the child assessment, psychomotor and antropometric. The assessor administered a psychomotor measure during the baseline data collection in the fall of kindergarten and recorded each child's height and weight. The physical component included a spiral bound psychomotor booklet, psychomotor instructions and score sheet, a height and weight recording form, the Shorr Board (for measuring height), and a digital bathroom scale. Materials used for measuring fine motor skills included ten wood blocks, a pencil, and two pieces of plain white paper. The psychomotor assessment was divided into measuring fine and gross motor skills. Fine motor skills were assessed by having each child use building blocks to replicate a model, copy forms (e.g., an asterisk, a square) on paper, and draw a person. The child was asked to skip, hop on one foot, walk backward, and stand on one foot to assess gross or large motor skills. To measure physical growth and development, children's height and weight were measured. Table 2-3 displays the major domains measured during the child direct assessments in the fall and spring of kindergarten.

Table 2-3. Direct child assessment

Direct child assessment	Fall-Kindergarten	Spring-Kindergarten
Language screener (OLDS)	X	/
Reading (language and literacy)	X	X
General knowledge (science and social studies)	X	X
Mathematical thinking	X	X
Psychomotor	X	
Height and weight	X	X

Note: The columns to the right of each construct correspond to the waves of questionnaire administration. Waves that included the construct are marked with an "X." A "/" indicates that the OLDS was given to students new in the spring, or who did not pass the cut score in the English version during the fall OLDS administration.

Child Assessment Spring

In spring of the base year, the children who participated in the fall data collection were assessed a second time. In addition, children who were not assessed in the fall were assessed for the first time. These children included those sampled in converted schools (see chapter 5, section 5.2 for more detail). Assessments began at least 12 weeks before the end of the school year. Table 2-4 displays the

Table 2-4. Flow of the spring direct child assessment

Home Language	English OLDS	Spanish OLDS	Warm-up Booklet	Reading	Math	General Knowledge	Height/ Weight
English			✓	✓	✓	✓	✓
Other	Score **at or Above** Cutpoint in fall-kindergarten		✓	✓	✓	✓	✓
	✓ Score at or **Above** Cutpoint		✓	✓	✓	✓	✓
	✓ Score **Below** Cutpoint Speaks Spanish	✓	✓ **Spanish**		✓ **Spanish**		✓
	✓ Score **Below** Cutpoint Doesn't Speak Spanish						✓

of the spring direct child assessments. The same assessment materials used in the fall were used with the following exceptions:

1. The psychomotor assessment was not administered in the spring-kindergarten assessment and beyond.

2. If the sampled child passed the OLDS in fall-kindergarten, he or she was automatically routed by the computer program to take the assessment in English and did not have to retake the OLDS. If the child did not score above the cut point for the OLDS in fall-kindergarten, he or she took the OLDS again in spring-kindergarten and was routed according to the new spring-kindergarten OLDS score.

3. If the child did not take the assessment at all in fall-kindergarten or was a newly sampled child in spring-kindergarten, he or she was routed according to his or her home language, just as the children were in fall-kindergarten.

2.2 Parent Interview

The majority of parents participating in the base year data collection were interviewed in the fall of 1998 and again in the spring of 1999. Because more children were added to build school response rates, a group of parents completed their first interview in the spring of 1999. (see chapter 5, section 5.2). Parents or guardians were asked to provide important information about the sampled child, the home environment, parent behavior (e.g., interactions with the child's teacher, activities with the sampled child), and family characteristics using a computer-assisted telephone interview (CATI) or computer-assisted personal interview (CAPI) for families without a telephone (see chapter 5, sections 5.4.4 and 5.6.4 for more details). Questions regarding family structure, child care use, household income, and child rearing practices were also included in the parent interview.

Fall Parent Interview

Typically the respondent for the parent interview was the mother of the child; however, the respondent could be a father, stepparent, adoptive parent, foster parent, grandparent, another relative, or nonrelative guardian. The respondent had to be knowledgeable about the child's care and education, be 18 years of age or older, and be living in the household with the child. In fall-kindergarten, respondents for the parent interview were selected according to the following order of preference:

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

The parent interview was conducted primarily in English, but provisions were made to interview parents who spoke other languages. The questionnaire was translated into Spanish, which was then printed on hardcopy. Bilingual interviewers were trained to conduct the parent interview in either English or Spanish. If the interview was conducted in Spanish, the interviewer used the hardcopy questionnaire and then entered respondents' answers into the CAI program. The parent interview was also translated into Chinese, Lakota, and Hmong languages and administered using the same data collection procedures as were used with Spanish speaking parents.

Topics addressed in the fall parent interview included a roster of current household members, family socio-demographic characteristics, languages spoken in the child's home, child care arrangements (currently and previously used), child's physical functioning, home activities, parent education and employment, and receipt of public assistance.

Parents were also asked to tell how often a student exhibited certain social skills and behaviors. The social rating scale (SRS) has five scales: approaches to learning, self-control, social interaction, impulsive/overactive, and sad/lonely.

See chapter 3, section 3.6.2 for variable names, ranges, means, and standard deviations for these scales.

- The **Approaches to Learning** scale (Parent SRS) includes six items that rate how often a child shows eagerness to learn, interest in a variety of things, creativity, persistence, concentration, and responsibility.

- The **Self-Control** scale (Parent SRS) has five items that indicate children's ability to control their behavior. It includes items that are worded positively as well as negative behaviors that are reverse coded (e.g., frequency with which a child fights, argues, throws tantrums, or gets angry).

- The **Social Interaction** scale (Parent SRS) asks about children's interactions with peers and adults. The three items address children's ease in joining in play, ability to make and keep friends, and positively interacting (comforting, helping) with peers.

- The **Impulsive/Overactive** scale (Parent SRS) has two items that ask about children's impulsivity and activity level.

- The **Sad/Lonely** scale (Parent SRS) has four items that ask parents about children's problems with being accepted and liked by others, sadness, loneliness, and low self-esteem.

Table 2-5 shows the broad content areas addressed in the parent interview and the point of data collection.

Spring Parent Interview

The content of the parent interview in spring-kindergarten differed from the fall interview. To avoid redundancy and increased respondent burden, many questionnaire topics were split between fall and spring data collection points. In spring-kindergarten, the parent interview included updating the household roster, parent's participation in activities in the child's school, parent's attitudes toward child-rearing, parent's psychological well-being and health, and the household's food situation. In addition, items from the SRS scales were repeated in the spring parent interview.

Table 2-5 shows the overall structure of the interview and distribution of topics across the two base year data collection points.

Table 2-5. ECLS-K parent interview by major content topics and point of data collection

Parent Questionnaire	Fall-Kindergarten	Spring-Kindergarten
Family structure	X	X
Demographics	X	X
Household roster	X	X
Marital status	X	X
Immigration status		X
Primary language(s) spoken in home	X	
Parent's involvement with child's school		X
Child care	X	
Arrangements with relatives	X	
Arrangements with nonrelatives	X	
Head Start attendance year before kindergarten	X	
Year before kindergarten child care arrangements	X	
Child's health and well-being	X	X
Birth weight	X	
Physical functioning	X	
Services for children with special needs	X	
Social skills rating	X	X
Home environment and cognitive activities	X	X
Frequency of literacy activities	X	X
Computer use		X
Television viewing		X
Parental educational expectations for child	X	
Neighborhood safety		X
Parent education	X	
Parent employment	X	
Parent income		X

Table 2-5. ECLS-K parent interview by major content topics and point of data collection (continued)

Parent Questionnaire	Fall-Kindergarten	Spring-Kindergarten
Welfare and other public assistance use	X	X
Parent/child interaction		X
Parent discipline		X
Parent health and emotional well-being		X
Relationships and social support	X	X
Marital satisfaction		X
Background data for fall-kindergarten nonresponding parents		X
Child's physical functioning/birth weight of child		/
Home language		/
WIC benefits		/
Parent education/mother's employment history		/
Prekindergarten Head Start attendance		/
Services for children with special needs		/

Note: The columns to the right of each construct correspond to the waves of questionnaire administration. Waves that included the construct are marked with an "X." Content areas asked in spring only to new parent participants are marked with a "/."

In the spring parent interview, households were routed to one of two questionnaire paths, either as a new household or as a continuing household that participated in fall data collection. Parents who completed the fall-kindergarten parent interview were classified as continuing households. Parents who were selected to participate in fall-kindergarten, but either refused, did not finish the interview, could not be located, or were unavailable during the field period were also classified as new households in the spring. Parents of children in schools converted in 1999 were classified as new households. The group of parents entering the study in spring 1999 were asked the same questions in the spring interview as those who participated in the fall. These respondents were also asked an additional set of questions to replicate some of the items included in the fall interview.

The rules for respondent selection varied by questionnaire path. Respondents in continuing households were selected according to the following order of preference:

1. The fall-kindergarten respondent;
2. The child's mother;

3. Another parent or guardian; and

4. Another household member.

Respondents in new households were selected according to the following order of preference:

1. The child's mother;

2. Another parent or guardian; and

3. Another household member.

2.3 Teacher Questionnaires

Each kindergarten teacher received a self-administered questionnaire consisting of three distinct parts. The first section, part A, asked about the teacher's class and classroom characteristics. It was designed to collect data about the composition and demographics of the children in the class. Part B addressed more specific questions on class organization, typical class activities, and evaluation methods, as well as teacher views on kindergarten readiness, school environment, and overall school climate. Background questions about the teacher were also included in this section.

Part C asked teachers to report about the sampled children in their classrooms. Teachers were asked to respond to 20 questions about the child's academic performance. The academic rating scale (ARS) gathered data on each sampled child's skills in areas of language and literacy, general knowledge, and mathematical thinking. For example, some questions asked if the child used complex sentence structure, demonstrated early writing behaviors, formed explanations based on observations, or solved problems involving numbers. Part C also included questions from the SRS that collected data on five areas of children's social skills.

2.3.1 Content of the ARS

There are three scales of the ARS: language and literacy, general knowledge, and mathematics. Each of these is described below. 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 based on the teacher's past observation and experience with the child. Unless otherwise noted, the fall and spring ARS have the same number of items for each concept.

- The **Language and Literacy** section of the ARS asks teachers to rate each child's proficiency in speaking (1 item), listening (1 item), early reading (3 items fall, 4 items spring), and writing (1 item fall, 2 items spring). In addition, teachers rate the child's computer literacy (1 item).

- The **General Knowledge** section of the ARS asks teachers to rate each child's proficiency in social studies (1 item fall, 2 items spring) and science (3 items).

- In the **Mathematics** section, teachers rate each child on one item on each of five skills: concept of numbers, solving number problems, use of math strategies, data analysis (graphing), and measurement.

In all sections, the teacher rated the child's skills, knowledge and behaviors on a scale from "Not Yet" to "Proficient" (see table 2-6). If a skill, knowledge, or behavior had not been introduced into the classroom, the teacher coded that item as N/A (Not applicable).

See chapter 3, section 3.3 for scale scores, value ranges, means, and standard deviations for the ARS.

Table 2-6. ARS response scale

Not yet:	Child has not yet demonstrated skill, knowledge, or behavior.
Beginning:	Child is just beginning to demonstrate skill, knowledge, or behavior but does so very inconsistently.
In progress:	Child demonstrates skill, knowledge, or behavior with some regularity but varies in level of competence.
Intermediate:	Child demonstrates skill, knowledge, or behavior with increasing regularity and average competence but is not completely proficient.
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.

2.3.2 Teacher SRS

Teachers rated individual students as part of a self-administered questionnaire. These items are intended to measure approaches to learning, self-control, and interpersonal skills. The items were rated on a scale of one (Never) to four (Very often). See chapter 3, section 3.6.1 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.

- The **Self-Control** (Teacher SRS) Scale 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 five **Interpersonal Skills** (Teacher SRS) items 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 two problem behavior scales reflect behaviors that may interfere with the learning process and with the child's ability to interact positively in the classroom.

- **Externalizing Problem Behaviors** (Teacher SRS) include acting out behaviors. Five items on this scale rate the frequency with which a child argues, fights, gets angry, acts impulsively, and disturbs ongoing activities.

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

This measure is adapted with permission from the instrument Elementary Scale A ("How Often?") (Gresham, F. M. and Elliott, S.N., 1990).

Fall-Kindergarten Teacher Questionnaires

Part A of the fall-kindergarten teacher questionnaire was given to all kindergarten teachers, regardless of whether they taught a sampled child. It included items characterizing the classroom and students. Part B included questions regarding class organization and evaluation, typical class activities, and views on kindergarten readiness. The part C questionnaire asked teachers to complete an ARS and SRS rating for each sampled child in his or her classroom.

Spring-Kindergarten Teacher Questionnaires

Similar to the fall instrument, the spring questionnaire was also divided into three sections. The content of the items varied between fall and spring instruments to ease respondent burden. Part A was again completed by all kindergarten teachers in the school, regardless of whether they taught a sampled child. This section asked about the characteristics of the kindergarten class and the children in the class. The spring part B questionnaire included some, but not all, of the questions used in the part B fall questionnaire. Part B of the teacher questionnaire was completed only by kindergarten teachers new to the study. This questionnaire was similar to the part B questionnaire completed in fall. It was used to gather information about teacher background, instructional practices, views on kindergarten readiness, and the teaching climate of the school. The teacher part C questionnaire was completed by kindergarten teachers who taught sampled children. This questionnaire, similar to the part C questionnaire completed in fall, asked teachers to provide ratings of each sampled child in their classroom in academic and social areas. The spring part C also included a section titled “student information” that asked for overall academic and physical activity comparisons of the sampled child with the other children in the class and student support service participation. Table 2-7 shows the overall structure of the teacher questionnaire and the distribution of topics among the fall and spring data collection points.

2.4 Special Education Teacher Questionnaires

The special education teacher questionnaires were new in the spring. ECLS-K supervisors reviewed accommodation and inclusion information for children who received special education services. During the preassessment visit, the field supervisors specified primary special education teachers of sampled children and listed special education staff working with each child (e.g., speech pathologists, reading instructors, and audiologists). These questionnaires were given to special education teachers who taught sampled children. If a child received special education services from more than one special education teacher, a field supervisor determined the child’s primary special education teacher.

Table 2-7. Teacher questionnaire

Teacher Questionnaire	Fall-Kindergarten	Spring-Kindergarten
Part A		
Description of class—age, race, and gender distribution	X	
Class organization	X[1]	X
Children with special needs		X
Classroom aides		X
Class assignment and grouping		X
Types of activity/ interest areas	X	
Class activities	X[1]	X
Instructional time in different subjects		X
Types of materials and activities		X
Child vs. teacher initiated activities	X[1]	
Parent involvement		X
Professional development		X
Part B		
Evaluation and grading practices of students	X	/
Sharing progress information with parents		/
Teachers' views on school readiness	X	/
Perceptions about school climate	X	/
Perception of personal influence on policies and classroom planning	X	
Teacher demographic information	X	/
Teacher experience and education	X	/
Job satisfaction	X	/
Transition to school activities	X	/

Table 2-7. Teacher questionnaire (continued)

Teacher Questionnaire	Fall-Kindergarten	Spring-Kindergarten
Part C		
Indirect child cognitive evaluation by teacher	X	X
Language and literacy, mathematics, general knowledge	X	X
Teacher evaluation of target child's social skills	X	X
Sampled child additional information		X
Participation in special services and programs		X
Target child's overall academic skills and physical activity		X
Reading group participation		X

[1] Collected in part B for fall-kindergarten.

Note: The columns to the right of each construct correspond to the waves of questionnaire administration. Waves that included the construct are marked with an "X." Content areas asked in spring only to new teacher participants are marked with a "/."

The primary special education teacher was defined as:

- The teacher who managed the child's Individualized Education Plan (IEP), or
- The teacher who spent the most amount of time providing special education services to the child, or
- The teacher who is most knowledgeable about the child's special needs and equipment.

Items in the special education teacher 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.

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. Table 2-8 provides a summary of the content areas addressed in the special education teacher questionnaire.

Table 2-8. Special education teacher questionnaire[1]

Teachers of Sampled Students with IEPs Questionnaire	Spring-Kindergarten
Part A	
Teacher's gender	X
Teacher's age	X
Teacher's race-ethnicity	X
Teaching experience	X
Educational background	X
Special education teacher background	X
Location of service provision	X
Student load/week	X
Part B	
Disability category	X
IEP goals for the school year	X
Extent of services	X
Types of services provided for the year	X
Primary placement	X
Teaching practices, methods, and materials	X
Assistive technologies used by child	X
General education goals, expectations and assessments	X
Collaboration/communication with child's general education teacher	X
Frequency of communicating with child's parents	X
Receipt of formal evaluations in the past year	X

[1] Data collected only in the spring.

2.5 Adaptive Behavior Scale

The Adaptive Behavior Scale was completed for all sampled children who were identified in spring-kindergarten as excluded from the direct child assessment. A child was excluded from the assessment if he or she needed the assessment administered in Braille, enlarged print, or sign language, or if the child's IEP specifically prohibited the child from taking standardized assessments. This questionnaire was completed by the child's primary special education teacher and asked the teacher to provide ratings of the sampled child in three domains: independent functioning, language development, and numbers and time.

2.6 School Administrator Questionnaire

The school principal, administrator, or headmaster was asked to complete the school administrator questionnaire in the spring of 1999 (shown in table 2-9). This self-administered questionnaire was intended to gather information about the school, student body, teachers, school policies, and administrator characteristics. The questionnaire was divided into nine sections. The first seven sections requested mainly factual information about each school and the programs offered at the school. These sections could be completed by either a principal or a designee who was able to provide the requested information. The school's principal was asked to complete the remaining two sections concerning his or her background and evaluations of the school climate. This questionnaire was administered only in the spring of the base year data collection.

2.7 School Facilities Checklist

ECLS-K supervisors completed the facilities checklist. The facilities checklist collects information about the (1) availability and condition of the selected schools, (2) presence and adequacy of security measures, (3) presence of environmental factors that may affect the learning environment, and (4) overall learning climate of the school. (See chapter 5, section 5.6.5 for more detail on the collection of these data.)

2.8 School Records Abstract Form

School staff completed the student records abstract form. This instrument was used to obtain information about the child's attendance record, report card, and use of an individualized education plan. Information about the type of language or English proficiency screening that the school used and whether the child participated in Head Start prior to kindergarten was also retrieved from the students' records. (See chapter 5, section 5.6.5 for more detail on the collection of these forms.)

Table 2-9. School administrator questionnaire[1]

School Questionnaire	Spring-Kindergarten
School characteristics	X
Type of school	X
Admission requirements	X
School size	X
Student characteristics	X
Race-ethnicity of students	X
Children eligible for special services	X
Types of kindergarten programs	X
School facilities and resources	X
Equipment	X
Community characteristics and school safety	X
Teaching and other school staff characteristics	X
Range of salary paid to teachers	X
Race-ethnicity of staff	X
School policies and programs	X
Assessments, testing, and retention	X
School-family-community connections	X
Programs and activities for families	X
Parent involvement and participation	X
Programs for special populations	X
ESL and bilingual education	X
Special education	X
Gifted and talented	X
Principal characteristics	X
Gender, race-ethnicity, age of principal	X
Experience and education	X
School governance and climate	X
Goals and objectives for kindergarten teachers	X
School functioning and decisionmaking	X

[1] Collected only in the spring.

2.9 Salary and Benefits Questionnaire

The salary and benefits questionnaire collects information on the base salary, merit pay, and benefit pay of teachers and principals. The salary and benefits questionnaire was completed at the school or district level, generally by the school or district business administrator or by a private school administrator or headmaster. The teacher salary and benefits questionnaire can be used to develop child-level school resource variables that can be linked to child outcomes. The interest in payroll information stems from the fact that salaries and benefits constitute approximately 80 percent of all current expenditures in school budgets. Although instructional expenditures are 61.8 percent of total current expenditures, salaries and benefits for instruction alone constitute 56 percent of total current expenditures, and 91.6 percent of all instructional expenditures.[4] These data provide an opportunity to learn more about how resources are allocated and used in schools and how those spending decisions impact children's achievement.

2.10 Head Start Verification

The goal of this part of the study is to verify that parent and school reports of children's Head Start participation. Respondents to the fall-kindergarten parent interview were asked in a series of questions about childcare outside the home whether or not the sampled child had ever attended Head Start. If the response was "Yes," the respondent was asked whether or not the sampled child attended Head Start in the year before kindergarten. Information on the name and location of the Head Start facility was matched by the interviewer against a database of Head Start centers. Similar data were collected in the spring-kindergarten student record abstract. In the student record abstract, one item asked whether or not the sampled child had attended a Head Start center before entering kindergarten. If the answer was "Yes," then the school staff person was asked to record the name, address, and telephone number of the Head Start center and the name of the Head Start director. The next step was to verify that the centers reported by the respondents were Head Start centers and that the child did attend the center in the 1997-98 school year. Chapter 5, section 5.6.7, describes how the data were collected.

---

[4] U.S. Department of Education, National Center for Education Statistics. *Digest of Education Statistics*, 1999, NCES 2000-031, Washington DC 2000. Pp. 185, table 168.

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

Several types of scores are used in the Early Childhood Longitudinal Study-Kindergarten Class of 1998-99 (ECLS-K) to describe children's cognitive and social development during their kindergarten year. These scores are for the direct cognitive assessment, the academic rating scale (ARS), the psychomotor assessment, and the social rating scale (SRS). Descriptions of the scores for each assessment or scale are provided below, along with variable names, variable descriptions, and value ranges found in the ECLS-K data files. Guidelines for when to use each cognitive test score are also provided in this chapter.

3.1 Direct Cognitive Assessment

The direct cognitive assessment contained items on reading, mathematics, and general knowledge in the fall and spring of kindergarten. In each subject area, children received a 12 to 20 item routing test. Performance on the routing items guided the selection of one of several alternative second-stage forms. The second-stage form contained items of appropriate difficulty for the level of ability indicated by the routing items. There are five different types of scores that can be used to describe children's performance on the direct cognitive assessment: (1) number right scores and (2) item response theory (IRT) scores, which measure children's performance on a set of test questions with a broad range of difficulty; (3) standardized scores, which report children's performance relative to their peers; (4) criterion-referenced proficiency level and (5) proficiency probability scores, which evaluate children's performance with respect to subsets of test items that mark specific skills. See chapter 2, section 2.1 for a description of the ECLS-K assessment battery.

Table 3-1 shows the types of scores, variable names, descriptions, ranges means, and standard deviations for the direct cognitive assessment.

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 tests that are the same for all children. However, when these scores are for tests that vary in average difficulty, they are not comparable to each other. For example, a student who took the middle difficulty mathematics second-stage form would probably have gotten more questions correct if he or she had taken the easier low form and fewer correct if the more difficult high form had been administered. For this reason, raw number right scores are reported in the database only for the first stage (routing) tests, which were the same for all children. Each routing test consisted of sets of items spanning a wide range of skills. For example, the reading routing test had four questions each on letter recognition, recognizing beginning sounds, recognizing ending sounds, reading simple sight words, and selecting words in the context of a sentence. An analyst might use the routing test number right scores to report actual performance on this particular set of tasks.

See table 3-1 for the variable names, ranges, means, and standard deviations for the number right scores.

3.1.2 IRT Scale Scores

Scores based on the full set of test items were calculated using IRT procedures. IRT made it possible to calculate scores that could be compared regardless of which second-stage form a child took. IRT uses the pattern of right, wrong, and omitted responses to the items actually administered in a test 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 test, plus a core set of items shared among the different second-stage forms, 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 test forms had been administered.

IRT has several other advantages over raw number-right scoring. By using the overall pattern of right and wrong responses 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, a correct difficult item is, in effect, assumed to have been guessed. 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, in effect, 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 test questions. Finally, IRT scoring makes possible longitudinal measurement of gain in achievement over time, even though the tests administered are not identical at each point. The common

Table 3-1. Direct cognitive assessment: types of scores, variable names, descriptions, ranges, means, and standard deviations[1]

Type of Score	Variable	Description	Range of Values	Mean	Standard Deviation
Number Right	C1RROUNR	C1Reading Routing Test, - Number Right	0 - 20	5.7	3.9
	C1MROUNR	C1 Mathematics Routing Test, - Number Right	0 - 16	4.4	2.9
	C1GROUNR	C1 Genl Knowledge Routing Test, - # Right	0 - 12	4.7	2.9
	C2RROUNR	C2Reading Routing Test, - Number Right	0 - 20	9.8	4.2
	C2MROUNR	C2 Mathematics Routing Test, - Number Right	0 - 16	7.1	3.3
	C2GROUNR	C2 Genl Knowledge Routing Test, - # Right	0 - 12	6.2	3.0
IRT	C1RSCALE	C1 Reading IRT Scale Score	0.0 - 72.0	22.0	8.3
	C1MSCALE	C1 Math IRT Scale Score	0.0 - 64.0	19.1	7.2
	C1GSCALE	C1 General Knowledge IRT Scale Score	0.0 - 51.0	22.1	7.4
	C2RSCALE	C2 Reading IRT Scale Score	0.0 - 72.0	31.6	10.3
	C2MSCALE	C2 Math IRT Scale Score	0.0 - 64.0	27.1	8.8
	C2GSCALE	C2 General Knowledge IRT Scale Score	0.0 - 51.0	26.8	7.8
Standardized Score	C1RTSCOR	C1 Reading T-Score	0.0 - 90.0	50.0	10.0
	C1MTSCOR	C1 Math T-Score	0.0 - 90.0	50.0	10.0
	C1GTSCOR	C1 General Knowledge T-Score	0.0 - 90.0	50.0	10.0
	C2RTSCOR	C2 Reading T-Score	0.0 - 90.0	50.0	10.0
	C2MTSCOR	C2 Math T-Score	0.0 - 90.0	50.0	10.0
	C2GTSCOR	C2 General Knowledge T-Score	0.0 - 90.0	50.0	10.0
Proficiency Level Score	C1RPROF1	C1 Prof 1 - Letter Recognition	0 - 1	0.65	0.8
	C1RPROF2	C1 Prof 2 - Beginning Sounds	0 - 1	0.30	0.46
	C1RPROF3	C1 Prof 3 - Ending Sounds	0 - 1	0.18	0.38
	C1RPROF4	C1 Prof 4 - Sight Words	0 - 1	0.04	0.20

[1] See chapter 7, section 7.3 for variable naming conventions.

Table 3-1. Direct cognitive assessment: types of scores, variable names, descriptions, ranges, means, and standard deviations (continued)

Type of Score	Variable	Description	Range of Values	Mean	Standard Deviation
	C1RPROF5	C1 Prof 5 - Word in Context	0 - 1	0.26	0.44
	C1MPROF1	C1 Prof 1 - Count, Number, Shape	0 - 1	0.90	0.31
	C1MPROF2	C1 Prof 2- Relative size	0 - 1	0.55	0.50
	C1MPROF3	C1 Prof 3 - Ordinality, sequence	0 - 1	0.20	0.402
	C1MPROF4	C1 Prof 4 -Add/Subtract	0 - 1	0.04	0.197
	C1MPROF5	C1 Prof 5 - Multiply/Divide	0 - 1	0.02	0.131
	C2RPROF1	C2 Prof 1 - Letter Recognition	0 - 1	0.92	0.277
	C2RPROF2	C2 Prof 2 - Beginning Sounds	0 - 1	0.70	0.46
	C2RPROF3	C2 Prof 3 - Ending Sounds	0 - 1	0.50	0.50
	C2RPROF4	C2 Prof 4 - Sight Words	0 - 1	0.14	0.35
	C2RPROF5	C2 Prof 5 - Word in Context	0 - 1	0.24	0.43
	C2MPROF1	C2 Prof 1- Count, Number, Shape	0 - 1	0.95	0.21
	C2MPROF2	C2 Prof 2 - Relative Size	0 - 1	0.82	0.38
	C2MPROF3	C2 Prof 3 - Ordinality, Sequence	0 - 1	0.54	0.50
	C2MPROF4	C2 Prof 4 - Add/Subtract	0 - 1	0.17	0.38
	C2MPROF5	C2 Prof 5 - Multiply/Divide	0 - 1	0.04	0.19
Proficiency Probability Score	C1RPROB1	C1 Prob 1 - Letter Recognition	0.0 - 1.0	0.66	0.43
	C1RPROB2	C1 Prob 2 - Beginning Sounds	0.0 - 1.0	0.29	0.36
	C1RPROB3	C1 Prob 3 - Ending Sounds	0.0 - 1.0	0.16	0.27
	C1RPROB4	C1 Prob 4 - Sight Words	0.0 - 1.0	0.02	0.13
	C1RPROB5	C1 Prob 5 - Word in Context	0.0 - 1.0	0.01	0.08
	C1MPROB1	C1 Prob 1 - Count, Number, Shape	0.0 - 1.0	0.92	0.20
	C1MPROB2	C1 Prob 2 - Relative Size	0.0 - 1.0	0.54	0.38
	C1MPROB3	C1 Prob 3 - Ordinality, Sequence	0.0 - 1.0	0.20	0.31
	C1MPROB4	C1 Prob 4 - Add/Subtract	0.0 - 1.0	0.04	0.12

Table 3-1. Direct cognitive assessment: types of scores, variable names, descriptions, ranges, means, and standard deviations (continued)

Type of Score	Variable	Description	Range of Values	Mean	Standard Deviation
	C1MPROB5	C1 Prob 5 - Multiply/Divide	0.0 - 1.0	0.00	0.04
	C2RPROB1	C2 Prob 1 - Letter Recognition	0.0 - 1.0	0.93	0.23
	C2RPROB2	C2 Prob 2 - Beginning Sounds	0.0 - 1.0	0.70	0.36
	C2RPROB3	C2 Prob 3 - Ending Sounds	0.0 - 1.0	0.50	0.36
	C2RPROB4	C2 Prob 4 - Sight Words	0.0 - 1.0	0.13	0.27
	C2RPROB5	C2 Prob 5 - Word in Context	0.0 - 1.0	0.04	0.16
	C2MPROB1	C2 Prob 1 - Count, Number, Shape	0.0 - 1.0	0.99	0.08
	C2MPROB2	C2 Prob 2 - Relative Size	0.0 - 1.0	0.84	0.26
	C2MPROB3	C2 Prob 3 - Ordinality, Sequence	0.0 - 1.0	0.53	0.40
	C2MPROB4	C2 Prob 4 - Add/Subtract	0.0 - 1.0	0.16	0.26
	C2MPROB5	C2 Prob 5 - Multiply/Divide	0.0 - 1.0	0.02	0.10
	C1RPRINT	C1 Print Familiarity	0 - 3	1.8	1.1
	C2RPRINT	C2 Print Familiarity	0 - 3	2.3	0.9

items present in the routing test and in overlapping second-stage forms allow the test scores to be placed on the same scale, even as the two-stage test design adapts to children's growth over time.

The IRT scale scores in the database represent estimates of the number of items students would have answered correctly if they had taken all of the 72 questions in the first- and second-stage reading forms, the 64 questions in all of the mathematics forms, and the 51 general knowledge items. These scores are not integers because they are probabilities of correct answers, summed over all items in the pool. Gain scores may be obtained by subtracting the estimated number right at time 1 from the estimated number right at time 2. (Note that scores for different subject areas are not comparable to each other because they are based on different numbers of test questions; that is, 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.)

See table 3-1 for variable names, ranges, means, and standard deviations for the IRT scale scores.

3.1.3 Standardized Scores (T-Scores)

T-scores provide norm-referenced measurements of achievement, that is, estimates of achievement level *relative to the population as a whole*. A high T-score mean for a particular subgroup indicates that the group's performance is high in comparison to other groups. It does not mean that group members have mastered a particular set of skills, only that their mastery level is greater than a comparison group. Similarly, a change in T-score means over time reflects a change in the group's status with respect to other groups. In other words, they provide information on *status compared to children's peers*, while the IRT scale scores and proficiency scores represent *status with respect to achievement on a particular criterion set of test items*. The T-scores can 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 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 fall- and spring-kindergarten. For example, a T-score of 55 (C1RTSCOR) represents a reading achievement level that is one-half of a standard deviation higher than the mean for the fall-kindergarten population represented by the tested sample of ECLS-K participants.

See table 3-1 for variable names, ranges, means, and standard deviations for the standardized (T) scores.

3.1.4 Proficiency Scores

Proficiency scores 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. Since the ECLS-K direct cognitive child assessment was a two-stage design (where not all children were administered all items), information on children's specific proficiencies are presented in two ways: proficiency scores (raw scores) and proficiency probability scores (IRT-based scores). In most situations, analysts use the proficiency probability scores in analyzing children's specific reading and mathematics knowledge and skills. Clusters of assessment questions having similar content and difficulty were included at several points along the score scale of the reading and mathematics assessments. No proficiency scores were computed for the general knowledge test, because the questions did not follow a hierarchical pattern. The following proficiencies were identified in the reading and mathematics assessments.

Reading:

- **Letter recognition:** identifying upper- and lowercase letters by name;
- **Beginning sounds:** associating letters with sounds at the beginning of words;
- **Ending sounds:** associating letters with sounds at the end of words;
- **Sight words:** recognizing common words by sight; and
- **Comprehension of words in context:** reading words in context.

Mathematics:

- **Number and Shape:** identifying some one-digit numerals, recognizing geometric shapes, and one-to-one counting of up to ten objects;
- **Relative Size:** reading all single-digit numerals, counting beyond ten, recognizing a sequence of patterns, and using nonstandard units of length to compare objects;

- **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;

- **Addition/Subtraction:** solving simple addition and subtraction problems; and

- **Multiplication/Division:** solving simple multiplication and division problems and recognizing more complex number patterns.

Clusters of items provide a more reliable test of proficiency than do single items because of the possibility of guessing; it is very unlikely that a student who has not mastered a particular skill would be able to guess enough answers correctly to pass a four-item cluster. 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 have indicated nonmastery at higher levels. Only a very small percentage of students in fall- and spring-kindergarten 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 both fall- and spring-kindergarten, 5.5 percent of the children did not follow the model for reading and 6.6 percent of the children did not follow the model for mathematics. This does not necessarily indicate a different order or learning for these children - since most of the proficiency-level items were multiple choice, many of these reversals are due to children guessing.

Proficiency level (dichotomous) scores and proficiency probability (continuous) level scores are two types of proficiency scores used in the ECLS-K. The following is a description of these scores.

Proficiency Level Scores (Dichotomous)

The proficiency level scores reflect the children’s raw ECLS-K direct cognitive assessment scores. These scores are intended for very distinct kinds of analysis. Most analysts should use the proficiency probability scores.

For each proficiency level, a score of one was assigned to children who correctly answered at least three of the four items in the cluster, and a score of zero was given if at least two items were incorrect or don’t know. If children did not answer enough items for pass or fail to be determined for a particular cluster, a pass/fail score was assigned only if the remaining proficiency scores indicated a level that was unambiguous. That is, a “fail” could be inferred if there were easier cluster(s) that had been failed and no higher cluster passed; and a “pass” was imputed if harder cluster(s) were passed and no easier one failed. In the case of ambiguous (e.g., pass, blank, fail, where the blank could be either a pass or a fail) or contradictory (e.g., fail, blank, pass) patterns, missing cluster scores were left blank.

Averaging a population subgroup's zero and one scores for a particular proficiency cluster results in an estimate of the proportion of children in the subgroup who had mastered the material at that level. The difference between this average at two points in time represents the proportion of children who attained mastery during that time period. These scores are not designed to extrapolate to the entire population of kindergarten children. These scores simply show - of the children who took the items represented by the level, how many passed the level. For example, these scores would be used in an analysis involving only the population of children that comprehend words in context. The analyst would analyze data on the subset of children that received a value of 1 on reading proficiency level 5 (comprehension of words in context), and would base the analysis only on this group of children.

See table 3-1 for variable names, ranges, means, and standard deviations for the proficiency level scores.

Proficiency Probability Scores (Continuous)

The proficiency probability scores are based on the same clusters of items as the proficiency level scores but differ from them in several ways. They are continuous rather than dichotomous and can take on any value between zero and one. They are estimates based on overall performance rather than counts of actual item responses. They are also estimates for all children with scorable test data, not only for the ones who were administered the test items in the cluster.

Due to the two-stage format of the cognitive assessment battery, not all children received all items. An IRT model was employed to produce proficiency probability scores, which reflect the probability that a child would have passed a proficiency level. 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. Because the proficiency probabilities were based on overall performance, they could be calculated for all children who had scorable test data, not just those with relatively complete sets of responses to the necessary item clusters.

The proficiency probability scores can be averaged to produce estimates of mastery rates within population subgroups. These continuous measures can provide a closer 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 are making their largest gains in achievement during a particular time interval. Thus, students' school experiences can be related to improvements in specific skills.

Proficiency level scores differ from proficiency probability scores. Proficiency level scores are based on the items administered to each child. Since not all children received the same items because of the two-stage assessment design, these scores only represent those children who were administered the items. The use of proficiency level scores to estimate the total population of children mastering a specific proficiency level is not recommended, because stopping rules within the test resulted in missing data for the lower-achieving children. The proficiency probability scores are more suited for estimating the total population of children mastering specific proficiency scores.

See table 3-1 for variable names, ranges, means, and standard deviations for the proficiency probability scores.

Familiarity with Conventions of Print

Some items from the child assessment measured children's familiarity with conventions of print but were not part of the set of proficiency scores because they did not fit the hierarchical pattern. The score for these questions was obtained by counting the number of correct answers (0-3) for the following three items, administered while the child was looking at an illustrated story.

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

These items were part of the reading score calculations in the direct cognitive assessment but were not part of the hierarchical set of proficiency and proficiency probability scores because they did not fit the proficiency scoring pattern. The proficiency levels assume that mastery of a higher level usually means that the child has mastered lower levels. This was not the case with conventions of print. Some children scored high on conventions of print but could not recognize letters, while others had the reverse pattern. Thus, the score for familiarity with conventions of print is reported separately.

See table 3-1 for variable names, ranges, means, and standard deviations for the conventions of print scores.

3.1.5 Choosing the Appropriate Score for Analysis

Each of the types of scores described above 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 vs. achievement in specific skills;
- An indicator of status at a single point in time versus growth over time; and
- A criterion-referenced vs. norm-referenced interpretation.

IRT-Based Scores

The scores derived from the IRT model (IRT scale scores, T-scores, proficiency probabilities) are 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, are used to estimate a point on an ability continuum, and this ability estimate, theta, then provides the basis for criterion-referenced and norm-referenced scores.

- **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 also be used as longitudinal measures of growth, but it is important to remember that gains made at different points on the score scale have qualitatively different interpretations. For example, children who make gains in recognizing letters and letter sounds are learning very different things from 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 gains in scale score points is most meaningful for groups that started with similar initial status. When initial status is very different, comparisons of scale score gains may be misleading because the skills being learned are qualitatively different, and comparisons of total number of points gained may be difficult to interpret.

- 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.

- **Proficiency probability scores**, although derived from the overall IRT model, are criterion-referenced measures of proficiency in **specific skills**. Because each proficiency score targets a particular narrow set of skills, they are ideal for studying the **details of achievement**, rather than the single summary measure provided by the 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 SES kindergarten children show 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 are making big gains in basic skills, but most have not yet made major gains in reading words and sentences. 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.

Non-IRT Based Scores

The routing test number right, proficiency level, and Conventions of Print scores do **not** depend on the assumptions of the IRT model. They are counts of actual number correct for specific sets of test items, rather than estimates based on patterns of overall performance.

- **Routing test number right scores** for the reading, math, and general knowledge assessments are based on 20, 16, and 12 items 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 test items, which was the same for all children who took the test.

- **Proficiency level scores** are based on the same sets of items as the proficiency probability scores, but are dichotomous, rather than continuous, measures of proficiency. They have a somewhat more intuitive interpretation than the probability scores, since they simply report whether children were able to answer correctly on at least three out of four actual test items in a cluster. Users of the proficiency level scores should be aware of possible bias due to missing data. Stopping rules employed in the administration of the tests to minimize stress on low-performing children results in substantial numbers of missing scores for the higher proficiency levels. Estimates based on variables with substantial amounts of missing data can be assumed to generalize to the whole sample only if "missing-ness" is unrelated to what the variable is measuring. This condition is called "MAR," or Missing-At-Random. The missing level four and five scores for low-achieving children are *not* missing-at-random, they were not administered based on performance. Interpretations of results based on these scores must take this into account. Similarly, missing data for the *lowest* math proficiency level are due to items in this cluster having been taken from the low second stage test, which was not taken by *high*-achieving children. Estimates based on proficiency level scores, without adjustments for missing data, would overstate the population performance at the high proficiency levels and understate performance at math level one.

- **Conventions of print scores**, like the proficiency level scores, are based on a count of the number correct for a particular set of items. Users may wish to relate this score to process variables to get a perspective that is somewhat different from that of the hierarchical levels of reading skills.

3.1.6 Reliabilities

Reliability statistics appropriate for each type of score were computed for each subject area, for fall- and spring-kindergarten assessments. For the IRT-based scores, the reliability of the overall ability estimate, theta, is based on the variance of repeated estimates of theta. These reliabilities, ranging from 0.88 to 0.95, 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 0.78 to 0.88. Split half reliabilities were computed for the item clusters that made up the dichotomous proficiency level scores and the conventions of print cluster. These reliabilities were higher for the reading clusters (0.60 to 0.83) than for the math levels (0.27 to 0.66). The difference in internal consistency statistics is due to the reading items being essentially replications of the same task, while the math items had a greater diversity of content.

Note that the split half reliabilities for the low level item clusters decreased from fall- to spring-kindergarten, while the reliabilities for the clusters at the upper end increased. This is a consequence of changes in the variance of the cluster scores as children progressed in their development of skills. By spring-kindergarten, the vast majority of children had mastered the lowest proficiency levels, so the sample variance was low, resulting in lower reliability than for the previous fall. Conversely, the sample variance for the difficult tasks was very low in the fall-kindergarten, when most children had *not* mastered these skills, and the reliability rose as some children attained high-level proficiency by the spring, increasing the total variance. This effect is more pronounced for the math than for the reading clusters for two reasons. First, the math item clusters were more heterogeneous than the reading, in terms of content and difficulty. Second, the reading item clusters were based entirely on items from the routing test, which was taken by all children, while the lowest math cluster employed items from the low level second stage test as well. By spring-kindergarten, fewer than half of the test takers were routed to the low form, and this constrained variance in ability resulted in a lower reliability for children who had this score. Tables 3-2 and 3-3 present the reliability statistics for all of the test scores.

Table 3-2. Reliability of IRT-based scores and routing test number correct

	IRT-based Scores (Reliability of Theta)		Routing Test Number Correct (Alpha Coefficient)	
	Fall-k	Spring-k	Fall-k	Spring-k
Reading	0.93	0.95	0.86	0.88
Math	0.92	0.94	0.78	0.81
General Knowledge	0.88	0.89	0.79	0.79

Table 3-3. Split half reliability of item-cluster-based scores (proficiency level scores)

	Fall-k	Spring-k
Reading Level 1	0.83	0.79
Reading Level 2	0.76	0.76
Reading Level 3	0.72	0.76
Reading Level 4	0.78	0.77
Reading Level 5	0.60	0.69
Conventions of Print	0.70	0.68
Math Level 1	0.41	0.27
Math Level 2	0.58	0.49
Math Level 3	0.63	0.66
Math Level 4	0.54	0.63
Math Level 5	0.46	0.53

3.2 Indirect Cognitive Assessment (ARS)

The ARS was developed for the ECLS-K to measure teachers' evaluations of students' academic achievement in the three domains that are also directly assessed in the cognitive battery: language and literacy (reading), general knowledge (science and social studies), and mathematical thinking. The difference between the direct and indirect cognitive assessments, and the scores available, are described below. For a discussion of the content areas and response scales of the ARS, see chapter 2, section 2.3.1.

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 measures only the products of children's achievement. 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 math problems, or investigate a scientific phenomenon.

Another major difference between the ARS and direct cognitive assessment is that the skills, knowledge, and behaviors on the ARS reflect a broader sampling of the most recent national curriculum standards and guidelines from early childhood professionals and researchers. The ARS items were not limited by the constraints of a standardized testing format as were the direct cognitive items. Therefore, the scope of curricular content represented in the indirect measures is broader than the content represented on the direct assessment battery.

3.3 IRT Scores Available for the ARS

IRT was employed to calculate scores for the ARS in order to compare performance of students from fall to spring and to be able to compare students who were not rated on all items. The Rasch

Rating Scale Model uses the pattern of ratings on the items actually administered to determine an estimate of the difficulty of each item and to place each student on a continuous ability scale.

A Rasch analysis was performed on the spring ARS data. In the fall-kindergarten, a large percentage of the teachers had not introduced at least some of the items to the classroom setting, resulting in a large number of missing ratings for all but four items.

By the spring-kindergarten, teachers had introduced the skills, knowledge, and behaviors represented in the items to their classrooms. The NA category was used in less than three percent of the cases on all items except "Composes simple stories" (7.52 percent NA); "Uses the computer" (15.83 percent NA); "Recognizes ways people rely on each other" (3.26 percent NA); and "Uses instruments for measuring" (9.82 percent NA). On the majority of the items, the use of the NA category was less than one percent.

The item difficulties from the spring analysis were applied to the fall items and ability estimates were computed for each of the children based on the difficulty estimates of the items and the pattern of ratings children received on those items. The Rasch analysis of the spring data showed that the reliability of the estimates of child ability was very high for all domains in both spring and fall (see table 3-4).

Table 3-4. Reliability for the Rasch-based score

	Fall-Kindergarten	Spring-Kindergarten
ARS Language and Literacy	0.87	0.91
ARS Mathematical Thinking	0.92	0.93
ARS General Knowledge	0.92	0.94

The ARS Scores were rescaled to have a low of one and a high of five to correspond to the five-point rating scale that teachers used in rating children on these items. The ARS scores in the database represent estimates of the rating students would have received in that domain if they had been rated on all of the items in the ARS. Gain scores may be obtained by subtracting the estimated rating at fall (time 1) from the estimated rating at spring (time 2).

The variable names, descriptions, value ranges, means, and standard deviations for the fall (T1) and spring (T2) kindergarten ARS scores are shown in table 3-5.

Table 3-5. ARS: variable names, descriptions, ranges, means, and standard deviations[1]

Variable Name	Description	Range of Values	Mean	Standard Deviation
T1ARSLIT	T1 Literacy ARS Score	1 - 5	2.6	0.8
T1ARSMAT	T1 Math ARS Score	1 - 5	2.5	0.8
T1ARSGEN	T1 General Knowledge ARS Score	1 - 5	2.6	1.0
T2ARSLIT	T2 Literacy ARS Score	1 - 5	3.3	0.8
T2ARSMAT	T2 Math ARS Score	1 - 5	3.5	0.9
T2ARSGEN	T2 General Knowledge ARS Score	1 - 5	3.6	1.0

[1] See chapter 7, section 7.3 for variable naming conventions.

The majority of teachers rated more than one student on the ARS. The number of students rated by each teacher ranged from 1 to more than 20.

3.4 The Oral Language Development Scale (OLDS)

The language assessment scores (OLDS scores) for language minority children are located with the other child scores on the file. There are a total of 16 scores – 4 English and 4 Spanish for each of the two rounds, i.e., fall-kindergarten and spring-kindergarten. Children in households speaking languages other than English were first given the English OLDS. Of that group, those scoring below the cut point of the English OLDS were administered the Spanish OLDS if the child's home language was noted as Spanish by the school. (See chapter 2, section 2.1 for more detail on the content of the OLDS items).

The variable names, descriptions, value ranges, means, and standard deviations for the OLDS are shown in table 3-6.

Table 3-6. OLDS: variable names, descriptions, ranges, means, and standard deviations[1]

Variable	Description	Range of Values	Mean	Standard Deviation
C1SCTOT	C1 AIQ400 Child's Total OLDS Score	1-60	34.0	16.0
C1SCORD	C1 AIQ400 Simon Says Child Score	1-10	8.0	2.3
C1SCART	C1 AIQ400 Art Show Child Score	1-10	7.3	2.7
C1SCSTO	C1 AIQ400 Tell Stories Child Score	4-40	22.4	9.0
C1SSCTOT	C1 SAIQ400 Spanish Total OLDS Scores	2-44	22.2	6.5
C1SSCORD	C1 SAIQ400 Spanish Simon Says Child Score	1-10	9.7	1.1
C1SSCART	C1 SAIQ400 Spanish Art Show Child Score	1-10	8.0	1.8
C1SSCSTO	C1 SAIQ400 Spanish Tell Stories Score	4-24	10.6	5.3
C2SCTOT	C2 AIQ400 Child's Total OLDS Score	1-60	30.9	15.2
C2SCORD	C2 AIQ400 Simon Says Child Score	1-10	7.6	2.4
C2SCART	C2 AIQ400 Art Show Child Score	1-10	6.8	2.6
C2SCSTO	C2 AIQ400 Tell Stories Child Score	4-40	20.7	8.7
C2SSCTOT	C2 SAIQ400 Spanish Total OLDS Scores	2-44	23.7	6.4
C2SSCORD	C2 SAIQ400 Spanish Simon Says Child Score	1-10	9.8	0.6
C2SSCART	C2 SAIQ400 Spanish Art Show Child Score	1-10	8.5	1.6
C2SSCSTO	C2 SAIQ400 Spanish Tell Stories Score	4-24	10.0	5.0

[1] See chapter 7, section 7.3 for variable naming conventions.

3.5 Psychomotor Assessment

The psychomotor assessment includes two scales, one measuring visual motor skills (eye-hand coordination) and the other measuring gross motor skills (balance and motor planning). The visual motor skills score is the sum of the points for seven tasks: build a gate, draw a person, and copy five simple figures. Children could receive up to two points for each of the first two tasks and one point for each of the figures. Gross motor skills consisted of balancing, hopping, skipping and walking backward—children could receive up to two points for each skill. Confirmatory factor analysis during the ECLS-K design phase (using LISREL) confirmed the two scales. The internal consistency of the scales was constrained by the limited number of items in each scale combined with the variety of motor skills measured and the limited variance in item scores (maximum score on items was 1-2). Alpha coefficients (reliabilities) were 0.57 for fine motor skills, 0.51 for gross motor skills, and 0.61 for the composite motor skills. Variable names, descriptions, value ranges, means, and standard deviations for the three scales are shown in table 3-7 below.

Table 3-7. Psychomotor scales: variable names, descriptions, ranges, means, and standard deviations[1]

Variable	Description	Range of Values	Mean	Standard Deviation
C1FMOTOR	C1 Fine Motor Skills	0 - 9	5.7	2.1
C1GMOTOR	C1 Gross Motor Skills	0 - 8	6.3	1.9
C1CMOTOR	C1 Composite Motor Skills	0 - 17	12.1	3.1

[1] See chapter 7, section 7.3 for variable naming conventions.

3.6 SRS

The SRS asked both teachers and parents to tell how often a student exhibited certain social skills and behaviors. Teachers and parents used a frequency scale (see table 3-8) to report on how often the student demonstrated the behavior described. See chapter 2, section 2.2 and 2.3.2 for additional information on the parent and teacher SRS instruments. The scale scores on all SRS scales are 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 same items were administered in the fall and in the spring so change scores may be computed by subtracting time 1 (fall) from time 2 (spring). The reliability for the teacher SRS scales is high (see table 3-9). The reliability is lower for the parent scales (see table 3-10).

Table 3-8. SRS response scale

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.

Table 3-9. Split half reliability for the teacher SRS scale scores

	Fall-Kindergarten	Spring-Kindergarten
Approaches to Learning	0.89	0.89
Self-Control	0.79	0.80
Interpersonal	0.89	0.89
Externalizing Problem Behaviors	0.90	0.90
Internalizing Problem Behaviors	0.80	0.78

Table 3-10. Split half reliability for the parent SRS scale scores

	Fall-Kindergarten	Spring-Kindergarten
Approaches to Learning	0.68	0.69
Self-Control	0.74	0.75
Social Interaction	0.70	0.68
Impulsive/Overactive	0.46	0.47
Sad/Lonely	0.60	0.61

3.6.1 Teacher SRS

Teachers rated individual students as part of a self-administered questionnaire. The five social skill teacher scales are: approaches to learning, self-control, interpersonal skills, externalizing problem behaviors, and internalizing problem behaviors. (See chapter 2, section 2.3.2 for a description of the teacher scales.)

Variable names for the teacher scores, descriptions, ranges, means, and standard deviations for these scales are shown in table 3-11.

Table 3-11. Teacher social rating scores: variable names, descriptions, ranges, means, and standard deviations[1]

Variable	Description	Range of Values	Mean	Standard Deviation
T1LEARN	T1 Approaches to Learning	1-4	3.0	0.7
T1CONTRO	T1 Self-Control	1-4	3.1	0.6
T1INTERP	T1 Interpersonal	1-4	3.0	0.6
T1EXTERN	T1 Externalizing Problem Behaviors	1-4	1.6	0.6
T1INTERN	T1 Internalizing Problem Behaviors	1-4	1.6	0.5
T2LEARN	T2 Approaches to Learning	1-4	3.1	0.7
T2CONTRO	T2 Self-Control	1-4	3.2	0.6
T2INTERP	T2 Interpersonal	1-4	3.1	0.6
T2EXTERN	T2 Externalizing Problem Behaviors	1-4	1.7	0.7
T2INTERN	T2 Internalizing Problem Behaviors	1-4	1.6	0.5

[1] See chapter 7, section 7.3 for variable naming conventions.

Care should be taken when entering these scales into the same analysis due to problems of multicollinearity. The factor intercorrelations among the scales for social skills are high. The factor intercorrelations with the internalizing problem behaviors are the lowest. The correlations between the teacher SRS factors range from 0.25 to 0.78 in fall-kindergarten and from 0.30 to 0.80 in spring-kindergarten (absolute values).

3.6.2 Parent SRS

The items on the parent SRS were administered as part of a longer telephone or in-person survey. (See chapter 2, section 2.2 for a description of the parent scales.) The factors on the parent SRS are similar to the teacher SRS; however, the items in the parent SRS are geared to the home environment and thus are not the same items. It is also important to keep in mind that parents and teachers observe the children in very different environments. The five social skill parent scales are: approaches to learning, self-control, social interaction, impulsive/overactive, and sad/lonely. The correlations between the parent SRS factors were not as high as the teacher SRS factors. They ranged from 0.05 to 0.45 in fall-kindergarten, and from 0.08 to 0.45 in spring-kindergarten (absolute values).

Variable names for the parent scores, descriptions, ranges, means, and standard deviations for these scales are shown in table 3-12.

Table 3-12. Parent social rating scores: variable names, descriptions, ranges, means, and standard deviations[1]

Variable	Description	Range of Values	Mean	Standard Deviation
P1LEARN	P1 Approaches to Learning	1-4	3.1	0.5
P1CONTRO	P1 Self-Control	1-4	2.8	0.5
P1SOCIAL	P1 Social Interaction	1-4	3.3	0.6
P1SADLON	P1 Sad/Lonely	1-4	1.5	0.4
P1IMPULS	P1 Impulsive/Overactive	1-4	2.0	0.7
P2LEARN	P2 Approaches to Learning	1-4	3.1	0.5
P2CONTRO	P2 Self-Control	1-4	2.9	0.5
P2SOCIAL	P2 Social Interaction	1-4	3.4	0.5
P2SADLON	P2 Sad/Lonely	1-4	1.6	0.4
P2IMPULS	P2 Impulsive/Overactive	1-4	2.0	0.7

[1] See chapter 7, section 7.3 for variable naming conventions.

4. SAMPLE DESIGN AND IMPLEMENTATION

The Early Childhood Longitudinal Study-Kindergarten Class of 1998-99 (ECLS-K) employed a multistage probability sample design to select a nationally representative sample of children attending kindergarten in 1998-99. 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.

4.1 Selecting the Area Sample

The point of departure for the ECLS-K area sample frame development was an existing multipurpose frame of PSUs created using 1990 county-level population data. This frame contained 1,404 PSUs that were counties or groups of contiguous counties. PSUs did not cut across census regional boundaries but were allowed to cross state boundaries. Each 1990 Metropolitan Statistical Area (MSA) constituted a single PSU except where an MSA crossed census regions, and it was split into two PSUs. The minimum size of a PSU in the multipurpose frame was 15,000 persons.

Since the focus of the ECLS-K is kindergarten students, the existing PSU frame was updated with 1994 population estimates of five-year-olds by race-ethnicity, the most up-to-date estimates available from the U.S. Census Bureau at the time. The counts of five-year-olds by race-ethnicity were used to revise PSU definitions relative to a different minimum PSU size and to construct a measure of size that facilitated the oversampling of Asians and Pacific Islanders (APIs).

Each PSU in the frame that did not have at least 320 five-year-olds was collapsed with an adjacent PSU. This minimum PSU size was developed based on assumptions concerning anticipated school response rates, the average number of schools that would be selected per PSU, and the target number of students to be sampled per school. After this collapsing, the final ECLS-K PSU frame contained 1,335 records.

The measure of size used for selecting PSUs took into account the amount of oversampling of APIs required to meet the ECLS-K precision goals. The weighted measure of size was calculated as follows:

$$MOS = 2.5 \times n_{API} + n_{other}$$

where 2.5 is the oversampling rate for APIs and n_{API} and n_{other} are the counts of five-year-old APIs and all others, respectively. The oversampling rate for APIs was calculated as the target number of completed API cases divided by the expected number of completed API cases without oversampling.

In all, 100 PSUs were selected for the ECLS-K. The 24 PSUs with the largest measures of size were designated as certainty selections or self-representing (SR) and were set aside. Once the SR PSUs were removed, the remaining PSUs were partitioned into 38 strata of roughly equal measure of size. The frame of non-SR PSUs was first sorted into eight superstrata by MSA/nonMSA status and by Census region. Within the four MSA superstrata, the variables used for further stratification were race-ethnicity (high concentration of API, Black, or Hispanic), size class (MOS >= 13,000 and MOS < 13,000) and 1988 per capita income. Within the four non-MSA superstrata, the stratification variables were race-ethnicity and per capita income. Details of the stratification of non-self-representing strata are presented in table 4-1.

Two PSUs were selected from each non-SR stratum using Durbin's Method.[1] This method selects two first-stage units per stratum without replacement, with probability proportional to size and a known joint probability of inclusion. The Durbin method was used because it allows variances to be estimated as if the units were selected with replacement. Table 4-2 summarizes the characteristics of the ECLS-K PSU sample.

4.2 Selecting the School Sample

In the second stage of sampling, public and private schools offering kindergarten programs were selected. For each ECLS-K PSU, a frame of public and private schools offering kindergarten

---

[1] Durbin, J. (1967). Design of Multi-Stage Surveys for the Estimation of Sampling Errors. *Journal of the Royal Statistical Society* C, 16, 152-164.

Table 4-1. Stratum definitions for the 38 non-self-representing strata

Metro status	Census region	Race-ethnicity (%)	PSU size (MOS)	Per capita income Range ($)	
MSA	Northeast	-	≥13,000	22,062	25,424
				16,342	22,030
			<13,000	18,128	29,084
				16,697	18,032
				12,279	16,616
	Midwest	-	≥13,000	17,277	18,150
				16,103	17,092
			<13,000	16,552	24,009
				15,732	16,475
				14,450	15,693
				10,185	14,433
	South	Hispanic ≥30	-	-	-
		Black ≥40	-	-	-
		26≤ Black <40	-	14,743	18,731
			-	10,892	14,573
		Black <26	≥13,000	16,435	16,601
				14,586	16,337
			<13,000	15,572	22,824
				14,194	15,432
				11,262	13,979
	West	API ≥15	-	-	-
		API <15	-	-	-
		Hispanic ≥30	-	-	-
		12≤ Hispanic <30	-	-	-
		Hispanic <12	-	15,048	21,840
			-	9,993	14,839
NonMSA	Northeast	-	-	-	-
	Midwest	-	-	14,124	17,446
				13,277	14,121
				12,169	13,272
				6,992	12,147
	South	Black ≥42	-	-	-
		25≤ Black <42	-	-	-
		Black <25	-	12,727	20,059
				11,165	12,676
				6,018	11,142
	West	-	-	12,887	23,286
				6,959	12,884

Table 4-2. Distribution of the ECLS-K PSU sample by SR status, MSA status, and census region

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

programs was constructed using existing school universe files: the 1995-96 Common Core of Data[2] (CCD) and the 1995-96 Private School Universe Survey[3] (PSS). The school frame was freshened in the spring of 1998 to include newly opened schools that were not included in the CCD and PSS and schools that were in the CCD and PSS but did not offer kindergarten according to those sources. A school sample supplement was selected from the freshened frame.

4.2.1 Frame Construction

The 1995-96 CCD Public School Universe File was the primary source for the ECLS-K public school sampling frame. Most schools run by the Bureau of Indian Affairs (BIA) and the schools run by the Department of Defense (DOD) are not included on the CCD. The 1995-96 Office of Indian Education Programs Education Directory was consulted in order to complete the list of BIA schools in the CCD file. For the DOD schools, a 1996 list of schools obtained directly from the DOD was used. The 1995-96 PSS Universe File was used as the primary source of the private school sampling frame.

The first step in frame construction involved subsetting the file to schools located in counties that constituted the ECLS-K PSU sample. Further subsetting retained only those schools that offered transitional kindergarten, kindergarten, or transitional first grade, or which were strictly ungraded, as indicated by the school's gradespan.

[2] U.S. Department of Education, National Center for Education Statistics, Common Core of Data, Public School Universe Survey, 1995-96.

[3] U.S. Department of Education, National Center for Education Statistics. *Private School Universe Survey, 1995-96*, NCES 98-229, by Stephen P. Broughman and Lenore A. Colaciello. Washington, DC: 1998.

The constructed ECLS-K school frame included 18,911 public-school records and 12,412 private-school records. The school frame was freshened in the spring of 1998 to include schools that would be operational in fall 1998 but that were not included in the frame just described. The freshening procedures are given later in this section.

4.2.2 School Measure of Size, Stratification, and Sample Selection

Within each PSU, schools with fewer than a predetermined minimum number of kindergarten students were clustered together before sampling in order to obtain a sample that is closer to self-weighting. The minimum was 24 kindergartners for public schools and 12 for private schools. For simplicity's sake, the term "schools" will be used in reference to both individual schools and clusters of schools in the remainder of this discussion.

Schools were selected with probability proportional to size. As with the PSU sample, a weighted measure of size was constructed taking into account the oversampling of APIs.

The target number of sampled schools per PSU was calculated separately for public and for private schools and was adjusted upward to offset anticipated school response and eligibility rates. The number of schools allocated to each PSU was set proportional to the weighted measure of size of the PSU. A minimum of one school per PSU was imposed for any PSU so small that it would not otherwise have been allocated a school.

Public and private schools constituted distinct sampling strata. Within each of these strata, schools were sorted to ensure good sample representation across other characteristics.

The public school sample was selected using a traditional nested two-stage design of schools within the 100 PSUs. Within each PSU, public schools were ranked by measure of size and partitioned into three classes of roughly equal aggregate measure of size. Within each size class, schools were sorted in a serpentine manner by the proportion of APIs.

The private school sample was nested within PSUs only for the 76 non-SR PSUs, where schools were grouped within PSU by affiliation (religious vs. nonsectarian) and then sorted in a serpentine manner by the measure of size. To better control the sample distribution of religious/

nonsectarian schools, schools in the 24 SR PSUs were treated as if they were sampled from a single PSU and the sort was by affiliation and measure of size alone.

The selection of both public and private schools was systematic, with probability proportional to the measure of size. A total of 1,280 schools were selected for the ECLS-K, of which 934 were public and 346 were private schools.

4.2.3 Freshening the School Sample

Each public school district having one or more schools sampled was sent a sampling frame-based list of all schools offering kindergarten and was asked whether any school expected to offer kindergarten in academic year 1998-1999 was missing from the list. For each such school identified, school name, address, telephone number, grade span, and kindergarten enrollment were obtained. Also contacted were districts that fell within the boundaries of the ECLS-K PSUs but for which the CCD file listed no schools offering kindergarten, unless it was clear from their name that they were strictly secondary school districts (e.g., Middlebury Union High School District). The information obtained from the school districts was checked against the ECLS-K public school frame to confirm that these schools were truly new or newly eligible. Bona fide new schools were given an appropriate chance of being sampled. A new school's chance of selection was conditioned on the school district's probability of selection, which had been calculated exactly, based on the sampling intervals used during the systematic selection of the main school sample, and the positions in the frame and measures of size of all schools in the frame. Overall 252 new public schools were identified, and 19 were selected using systematic sampling with probability proportional to size.

The procedure for obtaining new school information from Catholic dioceses was exactly the same as for public schools. Since a diocese may cut across county or even state lines, each school identified by a diocese had to be associated with the correct county, and hence the correct PSU, before checking to see whether it was truly new. Since dioceses may cross PSU boundaries, a new Catholic school's chance of being sampled had to be conditioned on the diocese probability of selection within the PSU where the new school was located. There were 126 new Catholic schools identified, and 6 were selected using systematic sampling with probability proportional to size.

The search for non-Catholic private schools was considerably more complicated. Three classes of PSS schools that had previously not been given a chance of selection were reconsidered. Those were schools that had an unknown grade span because they had not responded to the 1995-96 PSS, those that responded but did not report offering kindergarten, and those that appeared for the first time on the 1997-98 PSS file. All told these accounted for 2,544 potential new non-Catholic private schools. Beyond these additions from PSS, procedures similar to those used by the Census Bureau in the PSS area frame search[4] were followed. These procedures included collecting lists of schools from different sources, matching them against the PSS list frame to remove duplicates, and further screening by telephone to verify new school status. The majority of new schools found by the Census Bureau for PSS came from Yellow Pages listings. The Yellow Pages search was the main source of new non-Catholic private schools in the ECLS-K as well, yielding an additional 8,861 possible new private schools. Since the number of kindergartners enrolled in these schools was unknown, a minimum kindergarten enrollment was assumed for sampling purposes (typically 24, unless the name was suggestive of daycare in which case 12 was assumed). From the 11,405 schools, a sample of 279 schools was selected using systematic sampling with a probability proportional to these imputed enrollments. The sampled schools were screened, during which process it was ascertained whether each school was public or private; if it was private, whether it would be open in academic year 1998-1999; and if it would be open in 1999-2000, whether it would offer kindergarten. If the answer to the last question was yes and the school was not Catholic, the school was released for data collection.

Local Education Agencies (LEAs) and local government offices were contacted for information on non-Catholic private schools, only in the smallest ECLS-K PSUs, on the theory that if these PSUs had coverage problems their large weights were likely to introduce a larger bias in the estimates. All LEAs within these PSUs were contacted by telephone. For each city/town within the PSU, a list of local government offices was compiled using the Blue Pages. Successive government offices were called within a city or town until one was found that could provide information on private schools. As with the Yellow Pages, new schools identified by LEAs and local government offices were unduplicated against the PSS file before being added to the new school frame. Since kindergarten enrollment was unknown, it was imputed as described in the previous paragraph and sampling was performed using systematic sampling with probability proportional to size. The LEA search resulted in the identification of 30 new private schools after unduplication, of which 14 were sampled. The local government search yielded 19 new schools, of which 8 were sampled. Finally, three additional new

[4] ibid.

private schools were reported by field staff based on personal knowledge. Of these, two schools were sampled. All told, there were 109 new non-Catholic private schools selected.

The characteristics of the ECLS-K original school sample are presented in table 4-3. Schools that were discovered to be ineligible during recruitment have been omitted from the tabulation. For counts of responding schools and a discussion of the limited use of school substitution within the ECLS-K, see chapter 5, sections 5.5, Fall-Kindergarten Response Rates, and 5.8, Spring-Kindergarten Response Rates.

4.2.4 Sampling Children, Parents, and Teachers within Schools

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 each targeted subpopulation. As mentioned earlier, APIs were the only subgroup that needed to be oversampled to meet the sample size goals. For each sampled school, the field staff obtained a complete list of kindergartners enrolled, taking special care that no child was excluded from the list because of disability or language problems.

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 with twins being sampled as a unit rather than as individuals, i.e. if one twin was sampled, both were included. In general, the target number of children sampled at any one school (not including the second twin) was 24. In some schools the oversampling goal for API students could not be met. For example, in a school with 24 kindergartners enrolled, all students would be sampled, which meant that API students could not be sampled at 2.5 times the rate of non-API students in that school. To offset shortfalls of this kind, the oversampling rate for APIs was increased to three at this stage of sampling.

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. Characteristics of the ECLS-K original school sample

	Public	Private	Total
Total	914	363	1,277
Region			
Northeast	161	82	243
Midwest	210	88	298
South	306	112	418
West	237	81	318
Urbanicity			
Large central city	168	77	245
Mid-size central city	172	76	248
Urban fringe of large city	265	117	382
Urban fringe of mid-size city	78	21	99
Large town	24	9	33
Small town	76	36	112
Rural	131	27	158
KG enrollment			
1 – 24	51	187	238
25 – 49	95	110	205
50 – 99	402	59	461
100 – 149	226	7	233
150 – 169	49	0	49
170 +	91	0	91
Religious affiliation			
Catholic	-	120	120
Other religious	-	149	149
Nonreligious, private	-	94	94
Free Lunch Program			
Low (<=25% eligible students)	268	-	268
Medium low (>25% and <=50%)	157	-	157
Medium high (>50% and <=75%)	129	-	129
High (>75%)	114	-	114
Missing	246	-	246

During the fall 1998 data collection, a census of kindergarten teachers was taken at each school. Each sampled child was linked to his or her kindergarten teacher. A child could be linked to only one teacher. If a child was taught by more than one teacher, a "primary" teacher was identified for the child. For the specifics of how this was done, see section 5.4.2, Fall-Kindergarten Data Collection, Distribution of Teacher Questionnaires. In spring 1999, 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 of kindergarten data collection. If a sampled child received special education services from such a teacher, the teacher was linked to that child. As with regular teachers, a child could be linked to only one special education teacher.

4.3 Calculation and Use of Sample Weights

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 general, there are three types of weights: child, teacher, and school-level weights. The use of these weights is essential to produce estimates that are representative of the population of kindergarten children, kindergarten teachers, and schools offering kindergarten programs. Data collected from different sources can be used to produce estimates at these three levels. For example, data collected from parents are used to produce estimates of characteristics of children as reported by parents.

Several sets of weights were computed for each of the two rounds of data collection (fall- and spring-kindergarten). Longitudinal weights were also computed for children with complete data from both rounds of the study. Unlike surveys that have only one type of survey instrument aimed at one type of sampling unit, the ECLS-K is a complex study with multiple types of sampling units, each having its own survey instrument. Each type of unit was selected into the sample through a different mechanism: children were sampled directly through a sample of schools; parents of the sampled children were automatically included in the survey; all kindergarten teachers in the sampled schools were included; special education teachers were in the sample if they taught any of the sampled children. Each sampled unit had its own survey instrument: children were assessed directly using a series of cognitive and physical assessments; parents were interviewed with a parent instrument; teachers filled out at least two different types of questionnaires depending on the round of data collection and on whether they were regular or special education teachers; school principals reported their school characteristics using the school administrator questionnaire. The stages of sampling in conjunction with the different nonresponse level at each stage and the diversity of survey instruments require that multiple sampling weights be computed for use in analyzing the ECLS-K data.

This section describes the different types of sample weights computed for the ECLS-K, how they were calculated, how they should be used, and their statistical characteristics.

4.3.1 Types of Sample Weights

Weighting was carried out in stages to produce child, teacher, and school weights. Several sets of child-level weights were computed for each round of data collection and for children with complete data from both rounds. While it is straightforward to use school- and teacher-level weights to produce school- and teacher-level estimates, careful consideration should be given to the choice of a child-level weight since it depends on the type of data analyzed. Each set of child-level weights is appropriate for a different set of data or combination of sets of data. Teacher-level weights were computed for each round of data collection, but there are no longitudinal teacher-level weights. School level weights were computed for use with data collected in spring-kindergarten through the school administrator questionnaires. These weights can also be used with any school-level data such as data from the school facilities checklists.

Tables 4-4 and 4-5 summarize the different types of cross-sectional weights and how they should be used. Cross-sectional weights provide an accurate estimate for the specific round of data collection. Table 4-4 describes weights for fall-kindergarten estimates, and table 4-5 describes weights for spring-kindergarten estimates, and table 4-6 describes weights for base-year or longitudinal estimates.

These tables are designed to help users choose appropriate weights for their analysis. Answers to the following three questions can help in the selection of the correct weight.

1. Is the analysis concerned with one point in time or two?
2. What is the population of interest or unit of analysis (i.e. child, teacher or school)?
3. What instruments do the data to be used in the analysis come from?

1. Is the analysis concerned with one point in time or two? If the analysis pertains only to fall kindergarten (single point in time) then table 4-4 guides the selection of weights, spring kindergarten (single point in time) then go to table 4-5, and both fall- and spring-kindergarten (two points in time) then go to table 4-6.

Table 4-4. The ECLS-K: fall-kindergarten cross-sectional weights

	Fall-kindergarten cross-sectional weights
Child-level weight	to be used for analysis of ...
C1CW0	fall-kindergarten 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) any fall-kindergarten teacher questionnaire A, B or C data, and c) data from the school administrator questionnaire or facilities checklist
C1PW0	fall-kindergarten parent interview data alone or in combination with a) fall child assessment data, b) fall-kindergarten teacher questionnaire A, B, or C data, and c) data from the school administrator questionnaire or facilities checklist. *Exception*: If data from the parent AND child assessment AND teacher questionnaire A or B (not C) are used then C1CPTW0 should be used.
C1CPTW0	fall-kindergarten direct child assessment data combined with fall-kindergarten parent interview data AND fall-kindergarten teacher questionnaire A or B (not C) data alone or conjunction with data from the school administrator questionnaire or facilities checklist
Teacher-level weight	to be used for analysis of …
B1TW0	fall-kindergarten teacher data, questionnaire part A or B alone or with data from the school administrator questionnaire or facilities checklist. *Exception*: When using items that were in the spring-kindergarten teacher questionnaire B (i.e. questions asked of teachers who were not present during fall-kindergarten data collection) B2TW0 weight should be used.

Table 4-5. The ECLS-K: spring-kindergarten cross-sectional weights

Spring-kindergarten cross-sectional weights	
Child-level weight	to be used for analysis of ...
C2CW0	spring-kindergarten 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) any spring-kindergarten teacher questionnaire A, B or C data, and c) data from the school administrator questionnaire or facilities checklist
C2PW0	spring-kindergarten parent interview data alone or in combination with a) spring child assessment data, b) spring-kindergarten teacher questionnaire A, B, or C data, and c) data from the school administrator questionnaire or facilities checklist. *Exception*: If data from the parent AND child assessment AND teacher questionnaire A or B (not C) are used then C2CPTW0 should be used
C2CPTW0	spring-kindergarten direct child assessment data combined with spring-kindergarten parent interview data AND spring-kindergarten teacher data alone or in conjunction with data from the school administrator or facilities checklist
Teacher-level weight	to be used for analysis of ...
B2TW0	spring-kindergarten data from questionnaire part A; fall- or spring-kindergarten data from questionnaire part B; or combination of data from fall- and spring-kindergarten teacher questionnaires A and/or B (there is no longitudinal teacher weight) alone or in conjunction with data from the school administrator questionnaire or facilities checklist
School administrator weight	to be used for analysis of ...
S2SAQW0	school administrator data or facility checklist data

2. What is the population of interest or unit of analysis (i.e. child, teacher or school)? After identifying the appropriate table based on the response to question 1, the next step involves whether the analysis requires a child-, teacher- or school-level weight. If the population of inference is kindergarten children, then the child-level weights will be appropriate. If generalizing to kindergarten teachers or classrooms, then the teacher-level weights should be used, and if generalizing to the population of schools with kindergartens, the school-level weight will be appropriate.

3. What instruments do the data to be used in the analysis come from? There are several options when deciding on which child-level weights to use, and the source of the data affect which weight to use. In each of the tables details under “to be used in the analysis of . . .” provide guidance based on whether the data were collected through the child assessments, parent interviews or teacher questionnaires.

Weight C1CW0 is used to estimate child-level characteristics or assessment scores for fall-kindergarten, and C2CW0 is for spring-kindergarten. Examples of such estimates are the percent of kindergarten children who are male, the percent of children who are API, the percent of children who are 6 when they enter kindergarten, the mean reading score of children, and the mean math score of children. These weights are also used for estimates of child characteristics for language minority (LM)/not Spanish children and children with disabilities. Some of these children were not assessed but their background characteristics such as age, gender, race-ethnicity, and characteristics of parents, teachers, and classrooms are available from the parent interview and the teacher questionnaires. The social rating scores (see chapter 3, section 3.5) from parents and teachers are also available for LM/not Spanish children and children with disabilities, regardless of whether they were assessed. In this chapter the terminology “LM/ not Spanish” refers to those children who spoke a non-English and non-Spanish language at home and did not pass the cut score in the OLDS. For these children, only height and weight measurements were taken. Similarly, in this chapter, the terminology “children with disabilities” refer to those children who were not able to participate in the assessment due to reasons of disability as specified in their IEP. Both these groups of children were assigned child weights even though they did not have cognitive assessment data.

When analyzing child data in conjunction with teacher data (e.g., teacher characteristics from teacher questionnaires A or B or social rating scores reported by teachers from teacher questionnaire C) and classroom data from teacher questionnaire A, weights C1CW0 (for fall-kindergarten) and C2CW0 (for spring-kindergarten) should be used. An example for the use of C1CW0 is in the analysis of the relationship between children’s approaches to learning as rated by their teachers, the teacher’s type of teaching certification, and the children's cognitive skills and knowledge. 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. However, different weights should be used for analysis of child data in conjunction with both parent and teacher data (CPTW0).

C1PW0 (for fall-kindergarten) and C2PW0 (for spring-kindergarten) are 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 are read to at least every day. 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, when exploring the relationship between home literacy behaviors and children's reading skills.

C1CPTW0 (for fall-kindergarten) and C2CPTW0 (for spring-kindergarten) are used when child direct assessment and teacher and parent data 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. These weights should be not be used for estimates using only direct child assessment data or only parent interview data. Also, any analysis of the subgroup of children who moved to a different school between fall- and spring-kindergarten should be done using C2CW0 or C2PW0 because these movers were treated separately in the nonresponse adjustment of C2CW0 and C2PW0.

B1TW0 (for fall-kindergarten) and B2TW0 (for spring-kindergarten) are used for teacher- and classroom-level estimates. For example, these weights would be used to estimate the percent of kindergarten teachers who teach a particular type of kindergarten program, the percent of teachers who use a language other than English in their classroom, or the percent of teachers who do not have any teacher certification. These weights would also be used in the estimation of classroom characteristics such as the percent of kindergarten classrooms with computer areas. Weights for the corresponding round should be used to produce round specific estimates; for example, B1TW0 should be used to estimate the number of teachers teaching kindergarten programs in fall 1998; B2TW0 should be used to estimate the number of kindergarten teachers in spring 1999. Teachers who were new to the study in the spring were asked a subset of questions from the fall-kindergarten teacher questionnaires A and B (spring-kindergarten teacher questionnaire B). When analyzing data from these items B2TW0 should be used (even though the variables start with the B1 preface, see chapter 7 for details). A panel weight was not created as there were very few repeated measurements between fall- and spring-kindergarten. Most of the data collected from the teachers about themselves and their classrooms were meant to pertain to the school year.

S2SAQW0 is used in the estimation of school characteristics such as the percent of schools that offer programs for children with special needs, or the percent of schools that use a standardized achievement tests as a requirement for admission, or the percent of school administrators who believe that their school computer resources are inadequate.

The longitudinal or panel weights (table 4-6) are used for estimates of differences at two points in time. Examples of analysis using longitudinal weights include:

- Kindergarten fall-spring difference in mean child assessment scores (BYCW0);
- Kindergarten fall-spring difference in mean social skills as rated by children's teachers (BYCW0) (Data collected using the teacher questionnaire C are at a child-level and are considered as part of the child assessment data.);
- Kindergarten fall-spring difference in mean social skills as rated by children's parents (BYPW0);
- The relationship between the gains children make in their reading knowledge and skills, how often their parents read to them, how often their parents take them to the library, teacher certification, and how much class time teachers spend on reading (BYCPTW0) (This weight is used when the analysis includes data from all six components- fall and spring child assessment, teacher and parent data.); and
- The relationship between the gains children make in their reading knowledge and skills and parent and teacher beliefs on kindergarten readiness (BYCOMW0).

The difference between BYCPTW0 and BYCOMW0 is that BYCPTW0 is used for analysis of both rounds of child direct assessment data in conjunction with both rounds of parent interview data and both rounds of teacher data, while BYCOMW0 is used when analyzing a *single round* of parent interview or teacher data in conjunction with *both rounds* of child direct assessment data.

Table 4-6. The ECLS-K: base year longitudinal weights

Base year longitudinal (panel) weights	
<u>Child-level weight</u>	<u>to be used for analysis of ...</u>
BYCW0	child direct assessment data and child characteristics from both fall- and spring-kindergarten, alone or in conjunction with any combination of a) a limited set of child characteristics (e.g. age, sex, race-ethnicity), b) fall- and/or spring- kindergarten teacher questionnaires A, B or C data, and c) data from the school administrator questionnaire or facilities checklist
BYPW0	parent interview data from BOTH fall- and spring-kindergarten alone or in combination with a) fall- and/or spring-kindergarten child assessment data, b) fall- and/or spring-kindergarten teacher questionnaire A, B, or C data, and c) data from the school administrator questionnaire or facilities checklist. *Exception*: If data from the fall- AND spring-kindergarten parent, child assessment, AND teacher questionnaire A or B (not C) are used then BYCPTW0 should be used
BYCPTW0	fall- AND spring-kindergarten parent, child assessment, AND teacher questionnaire A or B (not C) data
BYCOMW0	both rounds of child assessment data in conjunction with at least one or more rounds (fall- and/or spring-kindergarten) of parent and/or teacher questionnaires A and B (not C) data. This may or may not be in conjunction with the school administrator questionnaire and facilities checklist data. *Exception*: Whenever BOTH rounds of parent data are used in the analysis either BYPW0 or BYCPTW0 is used.

Minimal teacher data were collected repeatedly in both rounds that would allow analysis of change over time for those data. In the absence of the teacher panel weights, the spring-kindergarten teacher weights can be used to cross-tabulate the fall- and spring-kindergarten teacher data from questionnaire part A (TQA).

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 C1CPTW0 were used in an analysis of child and teacher/classroom data only, then the resulting estimates will be inefficient compared to estimates using C1CW0. The lower parent response causes C1CPTW0 to result in lower sample size with positive weights. There may be combinations of data from a different source for which no weights were developed, but most analyses are possible from the weights provided. For example, no parent-level weights were calculated but parents are linked one-to-one to children so that the child-level weights can be used for parent-level analysis, e.g., education levels of parents of kindergartners. No child-teacher weights were computed for analyzing child data in conjunction with teacher data because the response rates for the teachers are high; for the analysis of child assessment data in conjunction with teacher data, the child-level weights should be used. For further advice on which weights to use when analyzing a complex combination of data, contact NCES at ECLS@ed.gov.

For each type of weight, table 4-7 gives the distribution of schools by number of sampled students with nonzero weights and the mean number of sampled students with nonzero weights per school. This is useful in analysis using hierarchical linear modeling. For spring-kindergarten, the increase in the count of schools with one to five sampled students is due to students transferring to other schools from the original sampled schools. For the longitudinal weights, schools are classified on the basis of the number of students who did not transfer schools between rounds of data collection.

Table 4-7. Distribution of schools by number of cases (children) with nonzero weights

	1 – 5	6 – 10	11 – 15	16 – 20	21 - 27	Mean cases per school
Fall-kindergarten						
C1CW0	32	42	52	187	636	20
C1PW0	39	45	78	301	486	19
C1CPTW0	41	53	109	328	401	18
Spring-kindergarten						
C2CW0	398	45	63	348	520	15
C2PW0	439	51	114	421	387	13
C2CPTW0	85	59	163	428	300	17
Longitudinal for base year						
BYCW0	36	40	74	373	423	19
BYPW0	41	57	149	431	267	17
BYCPTW0	46	80	190	414	198	17
BYCOMW0	42	48	112	404	324	18

4.3.2 Weighting Procedures

In general, weights were computed in two stages. In the first stage, base weights were computed. They are the inverse of the probability of selecting the unit—if units were sampled at a rate of 1 in 100, sampled units must be weighted by 100 to represent the entire population. In the second stage, base weights were adjusted for nonresponse. Nonresponse leads to bias in the survey estimates when the characteristics of the nonrespondents are very different from those of the respondents. Adjusting for nonresponse is intended to reduce the bias.

Nonresponse adjustment cells were generated using variables with known values for both respondents and nonrespondents. Analyses using the Chi-squared Automatic Interaction Detector (CHAID) were conducted to identify variables most highly related to nonresponse. At the school level, school characteristics used for constructing nonresponse cells were the type of school (public, Catholic private, non-Catholic private, or nonsectarian private), the school locale (large city, mid-size city, suburb of large city, suburb of mid-size 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. At the child level, the variables used for constructing nonresponse cells were the type of school, the locale and the geographic region where the school is located; the size classification of the school; and child characteristics such as age group, gender and race-ethnicity, and whether the child moved from the original sampled school (spring-kindergarten only). For the teachers, nonresponse cells were constructed using the type of school, the school locale, region, and the school size classification.

Once the nonresponse cells were determined, the nonresponse adjustment factors are the reciprocals of the response rates within the selected nonresponse cells. A detailed technical description of the nonresponse adjustment procedure can be found in the ECLS-K methodology report (forthcoming).

Response rates are presented in chapter 5 for the different populations and different types of instruments. A detailed analysis of response rates is available that includes a study of nonresponse bias. In this study, the ECLS-K survey estimates are compared with estimates from the sampling frames; they are also compared with estimates from other surveys such as the Current Population Survey (CPS) and the National Household Education Survey (NHES). The study also includes comparison of estimates using the nonresponse adjusted weights with estimates using unadjusted weights. Finally, a nonresponse simulation study is also provided to estimate the potential nonresponse bias.

4.3.3 Computation of School and Teacher Weights

School Base Weights

School base weights were used in calculating teacher and child weights for teacher- and child-level estimates and school administrator weights for school-level estimates. 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.

If schools were selected through the freshening procedure, as described in section 3.2.3, an additional factor equal to the inverse of the selection probability of the district or diocese was included in the base weight. This factor is necessary because new public and Catholic schools were identified through the freshening procedure with the district/diocese, and their selection probability must be conditioned on the probability of selecting that district/diocese within the stratum. This additional factor did not apply to non-Catholic private schools; these were selected directly from lists, and the school base weights were the simple inverse of the school selection probability.

School Administrator Weights (S2SAQW0)

School administrator weights were computed for schools sampled at the beginning of the study (fall-kindergarten) that completed the school administrator questionnaire in spring-kindergarten. The school administrator weight is the school base weight adjusted for school administrator nonresponse.

Teacher Weights (B1TW0 and B2TW0)

At each school sampled at the beginning of the study, all kindergarten teachers were included in the study. Each of these schools was considered a respondent if it had at least one completed teacher questionnaire, part B. The teacher weights were computed in two stages. First, the school base weights were adjusted for school nonresponse. Then, the teacher weights were computed as the school nonresponse adjusted weights adjusted for teacher nonresponse.

4.3.4 Computation of Child Weights

Child Weights

In general, child weights were computed in two stages. In the first stage, school base weights were adjusted for school nonresponse and then multiplied by the poststratified within-school child weights. In the second stage, the resulting weights were adjusted for child nonresponse. The poststratified, within-school child weight is equal to the total number of children in the school divided by the number of children sampled in the school. This is calculated separately for API and non-API children because different sampling rates were used for these two groups of children. Within a school, all API children have the same base weights and all non-API children have the same base weights.

A school was classified as responding using different criteria for fall-kindergarten, spring-kindergarten, and the base year. In fall-kindergarten, responding schools were eligible schools that agreed to cooperate. A school was considered cooperating if it agreed to provide lists of students (for sampling) and teachers (for distributing the teacher questionnaires) and certain information on students that would be used to plan for the assessment. In spring-kindergarten, schools that satisfied at least one of the following conditions were considered respondents: (1) have at least one child assessed in spring-kindergarten, or (2) have at least one sampled LM/not Spanish child who did not pass the Oral Language Development Scale (OLDS) cut score, or (3) have at least one sampled child with disabilities who could not be assessed according to the child's IEP, or (4) have at least one parent interviewed in spring-kindergarten. For the base year, a responding school was one that satisfied at least one of the following conditions: (1) have at least one child assessed in either round, or (2) have at least one sampled LM/not Spanish child in either round, or (3) have at least one sampled child with disabilities in either round, or (4) have at least one parent interviewed in either round. For each set of first-stage child weights, the appropriate school base weights were adjusted for school nonresponse, and then used in the computation of the final child weights.

Child Weights To Be Used with Direct Child Assessment Data (C1CW0, C2CW0, BYCW0)

In fall-kindergarten, responding children for this type of weight were eligible children who had fall-kindergarten scorable cognitive assessment data, or LM/not Spanish children who did not score at or above the OLDS cut score, but height and weight measurements were collected from them, or children with disabilities who according to specifications in their IEP could not participate in the assessments. A child was eligible if he or she was in kindergarten during fall 1998. A child who transferred to kindergarten in another school between sampling and assessment was considered to be a nonrespondent. In contrast, children who moved to first grade between sampling and assessment were considered ineligible. The fall-kindergarten child weights C1CW0 are the fall-kindergarten first-stage child weights adjusted for fall-kindergarten child nonresponse.

In spring-kindergarten, responding children were classified using rules similar to those used in fall-kindergarten. A child who transferred to another school between rounds and was not followed was considered a nonrespondent; children who moved outside the country were considered ineligible. The spring-kindergarten child weights C2CW0 are the spring-kindergarten first-stage child weights adjusted for spring-kindergarten child nonresponse. The child longitudinal weights BYCW0 were computed as the base year first-stage child weights adjusted for nonresponse. A respondent is defined as a child for whom both C1CW0 and C2CW0 were nonzero.

Table 4-8 shows the number of children who were not assessed due to the following special situations: children who were LM/not Spanish, children with disabilities, children who moved to another school between fall- and spring-kindergarten and who could not be located or because the new school was in a nonsampled county, and children who moved outside of the country or who were deceased. Only the LM/not Spanish and children with disabilities had child weights.

Table 4-8. Number of children who were not assessed due to special situations

	Number of children	
	Unweighted	Weighted
Fall-kindergarten		
With disabilities	88	18,106
LM/Not Spanish	415	39,148
Spring-kindergarten		
With disabilities	70	13,693
LM/Not Spanish	229	20,211
Moved schools in spring	606	129,562
Became ineligible in spring	67	13,340

Child Weights To Be Used with Parent Data (C1PW0, C2PW0, BYPW0)

The child weights C1PW0 (fall-kindergarten) and C2PW0 (spring-kindergarten) to be used with parent interview data are the corresponding first-stage child weights adjusted for nonresponse to the parent interview. In both fall- and spring-kindergarten, 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. The child longitudinal weights BYPW0 were computed as the base year first-stage child weights adjusted for nonresponse. A respondent is defined as a child for whom both C1PW0 and C2PW0 are nonzero. Note that these weights are at the child level even though the data were collected from the parents; they sum to all kindergarten children.

Child Weights To Be Used for Any Cross-Round Combination of Child Direct Assessment Data and Parent Interview Data and Teacher Data (C1CPTW0, C2CPTW0, BYCPTW0)

The child weights C1CPTW0 (fall-kindergarten) and C2CPTW0 (spring-kindergarten) to be used for analysis involving child, parent, and teacher data are the corresponding first-stage child weights adjusted for nonresponse. In both fall- and spring-kindergarten, a respondent for this type of weight was defined as a child who had scorable cognitive assessment data for the corresponding round (or LM/not Spanish children or children with disabilities), whose parent completed the FSQ section of the parent interview for the corresponding round, and whose teacher completed part B of the teacher questionnaire. The child longitudinal weights BYCPTW0 are the first-stage child weights for the base year adjusted for nonresponse. A respondent is defined as a child for whom both C1CPTW0 and C2CPTW0 are nonzero. Again, these weights are used to produce estimates of children even though the source of the data may be parent or teacher.

Child Weights To Be Used With a Single Round of Parent Interview or Teacher Data in Conjunction with Both Rounds of Child Direct Assessment Data (BYCOMW0)

These child longitudinal weights are the base year first-stage child weights adjusted for nonresponse. For this type of weight, a respondent is defined as a child whose (a) fall- and spring-kindergarten cross-sectional weights C1CW0 and C2CW0 are nonzero, and (b) either fall- or spring-kindergarten cross-sectional weight C1PW0 or C2PW0 is nonzero, and (c) either fall- or spring-kindergarten cross-sectional weight B1TW0 or B2TW0 is nonzero.

4.3.5 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. Any adjustments done 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. For example, the replicate weights for C1CW0 are C1CW1 through C1CW90. The method used to compute the replicate weights and how they can be used to compute the sampling errors of the estimates are described in the section on variance estimation.

4.3.6 Characteristics of Sample Weights

The statistical characteristics of the sample weights are presented in table 4-9 (teacher- and school-level weights) and in table 4-10 (child-level weights). For each type of weight, the number of cases with nonzero weights is presented together with the mean weight, the standard deviation, the coefficient variation (i.e., the standard deviation as a percentage of the mean weight), the minimum weight, the maximum weight, the skewness, the kurtosis, and the sum of weights.

Table 4-9. Characteristics of teacher- and school-level weights

	Number of cases	Mean	Standard deviation	CV (× 100)	Minimum	Maximum	Skewness	Kurtosis	Sum
Teacher									
Fall-kindergarten									
B1TW0	3,047	62.47	44.04	70.50	1.61	506.40	2.59	11.13	190,337
Spring-kindergarten									
B2TW0	3,243	58.64	39.67	67.64	1.60	453.44	2.43	10.09	190,166
School									
Spring-kindergarten									
S2SAQW0	866	83.44	53.07	63.60	6.42	484.64	2.24	8.32	72,260

Table 4-10. Characteristics of child-level weights

	Number of cases	Mean	Standard deviation	CV (× 100)	Minimum	Maximum	Skewness	Kurtosis	Sum
Fall-kindergarten									
C1CW0	19,173	201.63	91.94	45.60	1.64	755.65	1.35	4.85	3,865,946
C1PW0	18,097	213.62	96.19	45.03	2.03	832.40	1.47	5.71	3,865,946
C1CPTW0	17,124	225.76	104.57	46.32	2.17	1,018.25	1.45	5.49	3,865,946
Spring-kindergarten									
C2CW0	19,967	193.49	104.72	54.12	1.60	900.00	2.16	8.20	3,863,512
C2PW0	18,950	203.88	98.75	48.44	1.98	900.00	1.62	5.91	3,863,512
C2CPTW0	17,454	221.35	107.58	48.60	2.17	918.89	1.47	5.43	3,863,512
Longitudinal for base year									
BYCW0	18,211	212.14	119.54	56.35	1.59	900.00	2.45	9.82	3,863,204
BYPW0	16,906	228.51	109.75	48.03	2.22	900.00	1.62	5.63	3,863,204
BYCPTW0	15,420	250.53	121.33	48.43	2.54	1,146.11	1.58	6.07	3,863,204
BYCOMW0	17,060	226.45	126.48	55.85	1.59	900.00	2.33	8.82	3,863,204

The difference in the estimate of the population of teachers or students (sum of weights) between rounds of data collection is due to a combination of factors, among them: (1) the increase in the number of responding schools in spring-kindergarten that resulted from the refusal conversion efforts, and (2) the number of teachers and students who became ineligible after fall-kindergarten. The population of inference for all child-level weights is always the population of kindergartners in the school year 1998-99.

4.4 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 were selected with certainty and the first-stage non-self-representing sampling units were selected with two units per stratum, the paired jackknife replication method (JK2) is recommended. This section describes the JK2 and the Taylor Series estimation methods.

4.4.1 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:

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

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

The variance estimates of selected survey items presented in section 3.5 were produced using WesVar and the paired jackknife replication method.

Replicate Weights

Replicate weights were created to be used in the calculation of replicate 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 ECLS-K, replicate weights were created taking into account the Durbin method of PSU selection.[5] As mentioned in section 4.1, the Durbin method selects two first-stage units per stratum without replacement, with probability proportional to size and a known joint probability of inclusion.

[5] Durbin, J. (1967). Design of Multi-Stage Surveys for the Estimation of Sampling Errors. *Journal of the Royal Statistical Society C*, 16, 152-164.

In the ECLS-K PSU sample design, there were 24 SR strata and 38 non-self-representing (NSR) strata. Among the 38 NSR strata, 11 strata were identified as Durbin strata 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 above, 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, type of school (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 also numbered sequentially from 1 to 63, and so on until the list was exhausted, thus forming the desired 63 variance strata.

In strata with two units, a unit being a PSU in the case of NSR PSUs and a school in the case of SR PSUs, the base weight of the first unit was doubled to form the replicate weight, while the base weight of the second unit was multiplied by zero. In strata with three units, two variance strata were created: in the first variance stratum, the base weight of two of the three units was multiplied by 1.5 to form the replicate weight and the base weight of the last unit was multiplied by zero; in the second variance stratum, the base weight of a different group of two units was multiplied by 1.5, and the base weight of the third unit was multiplied by zero. Any adjustments done 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. For example, the replicate weights for child-level weights C1CW0 are C1CW1 through C1CW90.

4.4.2 Taylor Series Method

The Taylor Series method produces a linear approximation to the survey estimate of interest; then the variance of the linear approximation can be estimated by standard variance formulas. The stratum and first-stage unit 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 weight in each round of data collection and in the panel, while the combining of strata due to inadequate sample size and the sequential numbering of strata and first-stage units were done separately. 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 and STATA. They are described in table 4-11.

4.4.3 Specifications for Computing Standard Errors

Specifications for computing standard errors are given in table 4-12. For each type of analysis described in the table, users can choose between the replication method or the Taylor Series method for computing standard errors.

For the replication method using WesVar, the case identification (ID), the full sample weight, the replicate weights, and the method of replication are required parameters. All analysis of the ECLS-K data should be done using the paired jackknife method (JK2). As an example, to compute child level estimates (e.g., mean reading scores) and their standard errors, users need to specify CHILDID in the ID box of the WesVar data file screen, C1CW0 as the full sample weight, C1CW1 to C1CW90 as the replicate weights, and JK2 as the method of replication.

Table 4-11. ECLS-K Taylor Series stratum and first-stage unit identifiers

	Variable name	Description
Child level	C1TCWSTR	Sampling stratum – fall-kindergarten C-weights
	C1TCWPSU	First-stage sampling unit within stratum – fall-kindergarten C-weights
	C1TPWSTR	Sampling stratum – fall-kindergarten P-weights
	C1TPWPSU	First-stage sampling unit within stratum – fall-kindergarten P-weights
	CICPTSTR	Sampling stratum – fall-kindergarten CPT-weights
	C1CPTPSU	First-stage sampling unit within stratum – fall-kindergarten CPT-weights
	C2TCWSTR	Sampling stratum – spring-kindergarten C-weights
	C2TCWPSU	First-stage sampling unit within stratum – spring-kindergarten C-weights
	C2TPWSTR	Sampling stratum – spring-kindergarten P-weights
	C2TPWPSU	First-stage sampling unit within stratum – spring-kindergarten P-weights
	C2CPTSTR	Sampling stratum – spring-kindergarten CPT-weights
	C2CPTPSU	First-stage sampling unit within stratum – spring-kindergarten CPT-weights
	BYCWSTR	Sampling stratum – base year panel C-weights
	BYCWPSU	First-stage sampling unit within stratum – base year panel C-weights
	BYPWSTR	Sampling stratum – base year panel P-weights
	BYPWPSU	First-stage sampling unit within stratum – base year P-weights
	BYCPTSTR	Sampling stratum – base year panel CPT-weights
	BYCPTPSU	First-stage sampling unit within stratum – base year CPT-weights
	BYCOMSTR	Sampling stratum – base year panel COM-weights
	BYCOMPSU	First-stage sampling unit within stratum – base year COM-weights
Teacher level	B1TTWSTR	Sampling stratum – fall-kindergarten weights
	B1TTWPSU	First-stage sampling unit within stratum – fall-kindergarten
	B2TTWSTR	Sampling stratum – spring-kindergarten
	B2TTWPSU	First-stage sampling unit within stratum – spring-kindergarten
School level	S2SAQSTR	Sampling stratum
	S2SAQPSU	First-stage sampling unit within stratum

Table 4-12. Specifications for computing standard errors

Type of Analysis	Full Sample Weight	Computing Standard Errors: Replication Method (WesVarPC): ID	Replicate Weights	Jacknife Method	Taylor Series Method (SUDAAN & STATA): Sample Design	Nesting Variables		Approximating Sampling Errors: DEFT (Average Root Design Effect)
Fall-Kindergarten Cross-sectional								
■ Child-level								2.154
	C1CW0	CHILDID	C1CW1 - C1CW90	JK2	WR	C1TCWSTR	C1TCWPSU	
	C1PW0	PARENTID	C1PW1 - C1PW90	JK2	WR	C1TPWSTR	C1TPWPSU	
	C1CPTW0	CHILDID	C1CPTW1 - C1CPTW90	JK2	WR	C1CPTSTR	C1CPTPSU	
■ Teacher-level	B1TW0	T_ID	B1TW1 - B1TW90	JK2	WR	B1TTWSTR	B1TTWPSU	1.629
Spring-Kindergarten Cross-sectional								
■ Child-level								2.096
	C2CW0	CHILDID	C2CW1 - C2CW90	JK2	WR	C2TCWSTR	C2TCWPSU	
	C2PW0	PARENTID	C2PW1 - C2PW90	JK2	WR	C2TPWSTR	C2TPWPSU	
	C2CPTW0	CHILDID	C2CPTW1 - C2CPTW90	JK2	WR	C2CPTSTR	C2CPTPSU	
■ Teacher-level	B2TW0	T_ID	B2TW1 - B2TW90	JK2	WR	B2TTWSTR	B2TTWPSU	1.612
■ School-level	S2SAQW0	S_ID	S2SAQW1 - S2SAQW90	JK2	WR	S2SAQSTR	S2SAQPSU	1.279
Base Year Panel								
■ Child-level								1.884
	BYCW0	CHILDID	BYCW1 - BYCW90	JK2	WR	BYCWSTR	BYCWPSU	
	BYPW0	PARENTID	BYPW1 - BYPW90	JK2	WR	BYPWSTR	BYPWPSU	
	BYCPTW0	CHILDID	BYCPTW1 - BYCPTW90	JK2	WR	BYCPTSTR	BYCPTPSU	
	BYCOMW0	CHILDID	BYCOMW1 - BYCOMW90	JK2	WR	BYCOMSTR	BYCOMPSU	

For the Taylor Series method using either SUDAAN or STATA, the full sample weight, the sample design, the nesting stratum and PSU variables are required. For the same example above, the full sample weight (C1CW0), the without replacement sample design (WR), the stratum variable (C1TCWSTR) and the PSU variable (C1TCWPSU) must be specified.

The last column in table 4-12 gives the average root design effect that can be used to approximate the standard errors for each type of analysis. For a discussion of the use of design effects, see section 4.5.

4.5 Design Effects

An important analytic device is to compare the statistical efficiency of survey estimates with what would have been obtained in a hypothetical and usually impractical simple random sample (SRS) of the same size. 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 called 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:

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

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

4.5.1 Use of Design Effects

One method of computing standard errors for the ECLS-K is the paired jackknife method, as described in section 4.4, using programs designed specifically for analyzing complex survey data such as WesVar. Another approach, Taylor Series linearization (and software designed for it), is also discussed in the same section. If a statistical analysis software package such as SPSS (Statistical Program for the Social Sciences) and SAS (Statistical Analysis System) is used, the standard errors should be corrected using *DEFT*, since these programs calculate standard errors, assuming the data were collected with a simple random sample. The standard error of an estimate under the actual sample design can be calculated as follows:

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

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

Means

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

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.

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

4.5.2 Average 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. One way to produce design effects for analysts' use is to calculate them for a number of variables and average them. The averaging can be done overall and for selected subgroups. The tables that follow show estimates, standard errors, and design effects for selected means and proportions based on the ECLS-K child, parent, teacher, and school data. For each survey item, the tables present the number of cases, the estimate, the standard error taking into account the actual sample design (Design SE), the standard error assuming SRS (SRS SE), the root design effect (DEFT), and the design effect (DEFF). Standard errors (Design SE) were produced using the paired jackknife replication method (JK2).

For each survey estimate, the variable name as it appears in the ECLS-K Base Year Public-Use Electronic Code Book (ECB) is also provided in the table. If multiple variables were combined to arrive at the estimate, then the names of all the variables used are provided. For example, the estimate of the proportion of fall-kindergarten children whose home language was not English was computed using two different survey items, P1ANYLNG (parent questionnaire item PLQ020, whether another language was used at home) and P1PRMLNG (parent questionnaire item PLQ060, what was the primary language used at home). The first letter of the variable name indicates the source of the item: C – child assessment, P – parent instrument, A – teacher instrument part A, B – teacher instrument part B, T – teacher instrument part C, and S – school administrator questionnaire. The second letter of the variable name indicates when the data were collected: 1 – round 1, fall-kindergarten; 2 – round 2, spring-kindergarten. 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.

Child-Level Design Effects

Standard errors and design effects for the child-level items are presented in tables 4-13 to 4-15 for fall-kindergarten, spring-kindergarten, and for the base year. The survey items were selected so that there was a mix of items common to both fall- and spring-kindergarten and items that were specific to each round of data collection. For fall- and spring-kindergarten, the student-level items include the different scores from the assessment data, the social rating scores as provided by the parents and teachers, some characteristics of the parents, and some characteristics of the students as reported by the parents. For a small number of estimates, the data were subset to cases where the estimate is applicable; for example, the score for the Spanish math assessment applies only to students who were assessed in Spanish; the type of primary child care is only for children who had regular scheduled child care; the number of hours that the mothers work is only for women in the labor force; and the question on whether the parents have a happy relationship is only for parents in current relationships. For the base year student panel, design effects were calculated for the difference in scores between the two rounds of data collection, and also for some spring-kindergarten items.

The median design effect is 4.7 for fall-kindergarten (compared with 2.2 for the NELS:88 base year student questionnaire data) and 4.1 for spring-kindergarten (compared with 3.4 for the NELS:88 first followup). The size of the ECLS-K design effects is largely a function of the number of children sampled per school. With about 20 children sampled per school, an intraclass correlation of 0.2 might result in a design effect of about 5. The median design effect is 3.4 for the panel of students common to both fall- and spring-kindergarten, and the lower median design effect is due to the smaller cluster size in the panel. The ECLS-K design effects are slightly higher than the average of 3.8 that was anticipated during the design phase of the study, both for estimates for proportions and for score estimates.

Table 4-16 presents the median design effects for subgroups based on school type, child's gender and race-ethnicity, geographic region, and level of urbanicity. For fall-kindergarten, the median design effects vary from 2.0 (Hispanic) to 5.9 (children in small towns and rural areas). For spring-kindergarten and the panel, the range of variability of the median design effects is similar to that for fall-kindergarten; that is, 2.0 for Hispanic children to 6.7 for children in small towns and rural areas for spring-kindergarten and 1.6 for children of other race-ethnicity to 5.3 for children in small towns and rural areas for the panel. Once again, the variation in the design effects is largely a function of the sample size as well as the homogeneity of the children within schools.

Table 4-13. ECLS-K, fall-kindergarten: standard errors and design effects for the full sample – child level

Survey item	Variable name	Number of cases	Estimate	Design SE	SRS SE	DEFT	DEFF
Child scores (Mean)							
Reading score (English)	C1RSCALE	17,625	22.02	0.156	0.063	2.480	6.148
Math score (English)	C1MSCALE	17,615	19.52	0.141	0.054	2.606	6.789
Math score (Spanish)	C1MSCALE	1,021	12.88	0.279	0.132	2.115	4.472
General knowledge (English)	C1GSCALE	17,566	22.09	0.171	0.056	3.077	9.471
Composite motor skills	C1CMOTOR	18,422	12.08	0.050	0.023	2.176	4.734
Approaches to learning-Parent	P1LEARN	18,029	3.11	0.006	0.004	1.678	2.816
Self-control-Parent	P1CONTRO	18,023	2.83	0.006	0.004	1.430	2.045
Social interaction-Parent	P1SOCIAL	18,026	3.32	0.007	0.004	1.702	2.896
Withdrawn-Parent	P1SADLON	18,010	1.54	0.006	0.003	2.131	4.542
Impulsive/overactive-Parent	P1IMPULS	17,902	1.98	0.011	0.005	2.156	4.647
Approaches to learning-Teacher	T1LEARN	18,839	2.96	0.010	0.005	2.110	4.452
Self-control-Teacher	T1CONTRO	18,135	3.07	0.011	0.005	2.371	5.623
Interpersonal-Teacher	T1INTERP	17,923	2.96	0.010	0.005	2.125	4.516
Externalizing problems-Teacher	T1EXTERN	18,609	1.64	0.010	0.005	2.006	4.025
Internalizing problems-Teacher	T1INTERN	18,356	1.55	0.007	0.004	1.867	3.484
Child characteristics (Percent)							
Lived in single-parent family	P1HFAMIL	18,097	23.57	0.728	0.315	2.309	5.331
Lived in two-parent family	P1HFAMIL	18,097	74.43	0.790	0.324	2.438	5.942
Home language is non-English	P1ANYLNG, P1PRMLNG	18,059	12.09	0.724	0.243	2.983	8.896
Primary care is center-based	P1PRIMNW	8,173	36.82	1.195	0.533	2.240	5.019
Primary care is home-based	P1PRIMNW	8,173	63.18	1.195	0.533	2.240	5.019
Expected to graduate from college	P1EXPECT	17,968	74.74	0.703	0.324	2.168	4.702
Being read to every day	P1READBO	18,068	44.44	0.656	0.370	1.774	3.146
Was in excellent/very good/good health	P1HSCALE	18,055	96.92	0.164	0.129	1.274	1.624
Parents had high school or less	WKPARED	17,754	37.99	0.912	0.364	2.503	6.267
Mom worked 35 hours+/week	P1HMEMP	12,519	64.17	0.701	0.429	1.635	2.672
Child characteristics (Mean)							
Age of child in months	R1_KAGE	19,073	68.50	0.077	0.032	2.378	5.656
Child’s household size	P1HTOTAL	18,097	4.52	0.022	0.010	2.106	4.436
Number of children <18 in child’s HH	P1LESS18	18,097	2.50	0.020	0.009	2.247	5.047
Median						2.162	4.675
Mean						2.154	4.801
Standard deviation						0.406	1.775
Coefficient of variation						0.188	0.370
Minimum						1.274	1.624
Maximum						3.077	9.471

Table 4-14. ECLS-K, spring-kindergarten: standard errors and design effects for the full sample – child level

Survey item	Variable name	Number of cases	Estimate	Design SE	SRS SE	DEFT	DEFF
Child scores (Mean)							
Reading score (English)	C2RSCALE	18,937	31.64	0.213	0.075	2.855	8.151
Math score (English)	C2MSCALE	18,925	27.43	0.173	0.063	2.728	7.443
Math score (Spanish)	C2MSCALE	724	18.76	0.454	0.242	1.874	3.513
General knowledge (English)	C2GSCALE	18,903	26.81	0.158	0.056	2.801	7.844
Approaches to learning-Parent	P2LEARN	18,252	3.12	0.005	0.004	1.279	1.637
Self-control-Parent	P2CONTRO	18,251	2.87	0.006	0.004	1.685	2.839
Social interaction-Parent	P2SOCIAL	18,270	3.42	0.006	0.004	1.579	2.494
Withdrawn-Parent	P2SADLON	18,232	1.55	0.006	0.003	1.966	3.867
Impulsive/overactive-Parent	P2IMPULS	18,091	1.96	0.010	0.005	1.856	3.445
Approaches to learning-Teacher	T2LEARN	18,979	3.08	0.010	0.005	1.992	3.967
Self-control-Teacher	T2CONTRO	18,847	3.15	0.011	0.005	2.294	5.261
Interpersonal-Teacher	T2INTERP	18,767	3.09	0.009	0.005	1.999	3.997
Externalizing problems-Teacher	T2EXTERN	18,907	1.69	0.010	0.005	2.173	4.723
Internalizing problems-Teacher	T2INTERN	18,806	1.59	0.008	0.004	2.062	4.251
Child characteristics (Percent)							
Lived in single-parent family	P2HFAMIL	18,906	23.51	0.584	0.308	1.895	3.591
Lived in two-parent family	P2HFAMIL	18,906	74.33	0.645	0.318	2.030	4.120
Home language is non-English	P2ANYLNG, P2PRMLNG	18,862	12.53	0.739	0.241	3.064	9.391
Child used home computer	P2HOMECM	18,910	52.61	0.818	0.363	2.253	5.078
Child read outside school everyday	P2CHREAD	18,877	39.82	0.649	0.356	1.823	3.322
Parents had high school or less	WKPARED	17,607	37.14	0.910	0.364	2.499	6.244
Parent attended PTA	P2ATTENP	18,914	33.12	0.902	0.342	2.635	6.941
Parent thinks not safe for child to play outside	P2SAFEPL	18,898	96.17	0.220	0.140	1.574	2.479
Parents had happy relationship	P2MARRIG	14,291	97.77	0.164	0.123	1.333	1.776
Parent too busy to play with child	P2TOOBUS	18,600	45.55	0.591	0.365	1.618	2.617
Child characteristics (Mean)							
Age of child in months	R2_KAGE	19,890	74.76	0.070	0.032	2.209	4.881
Child's household size	P2HTOTAL	18,906	4.54	0.023	0.010	2.218	4.919
Number of children <18 in child's HH	P2LESS18	18,906	2.52	0.020	0.009	2.302	5.298
Median						2.030	4.120
Mean						2.096	4.596
Standard deviation						0.458	1.993
Coefficient of variation						0.218	0.434
Minimum						1.279	1.637
Maximum						3.064	9.391

Table 4-15. ECLS-K, panel: standard errors and design effects for the full sample – child level

Survey item	Variable name	Number of cases	Estimate	Design SE	SRS SE	DEFT	DEFF
Child scores							
(Difference between fall- and spring-kindergarten scores)							
Reading score (English)	C2RSCALE – C1RSCALE	16,751	9.88	0.123	0.047	2.603	6.778
Math score (English)	C2MSCALE – C1MSCALE	16,748	8.18	0.084	0.040	2.126	4.520
Math score (Spanish)	C2MSCALE – C1MSCALE	637	6.28	0.165	0.175	0.944	0.892
General knowledge (English)	C2GSCALE – C1GSCALE	16,697	5.16	0.070	0.031	2.255	5.084
Approaches to learning-Parent	P2LEARN - P1LEARN	16,326	0.01	0.005	0.003	1.467	2.151
Self-control-Parent	P2CONTRO - P1CONTRO	16,326	0.05	0.005	0.003	1.461	2.135
Social interaction-Parent	P2SOCIAL - P1SOCIAL	16,331	0.10	0.006	0.005	1.315	1.729
Withdrawn-Parent	P2SADLON – PISADLON	16,298	0.01	0.005	0.003	1.652	2.730
Impulsive/overactive-Parent	P2IMPULS - P1IMPULS	16,109	-0.01	0.009	0.005	1.670	2.790
Approaches to learning-Teacher	T2LEARN - T1LEARN	17,208	0.11	0.007	0.004	1.861	3.462
Self-control-Teacher	T2CONTRO - T1CONTRO	16,538	0.07	0.008	0.004	1.840	3.387
Interpersonal-Teacher	T2INTERP - T1INTERP	16,296	0.12	0.008	0.004	1.919	3.681
Externalizing problems-Teacher	T2EXTERN - T1EXTERN	16,944	0.06	0.005	0.003	1.481	2.193
Internalizing problems-Teacher	T2INTERN - T1INTERN	16,681	0.04	0.007	0.004	1.707	2.915
Child characteristics (Percent)							
Lived in single-parent family	P2HFAMIL	16,870	23.11	0.603	0.325	1.857	3.447
Lived in two-parent family	P2HFAMIL	16,870	74.80	0.671	0.334	2.007	4.029
Home language is non-English	P2ANYLNG, P2PRMLNG	16,906	11.65	0.706	0.247	2.860	8.182
Child used home computer	P2HOMECM	16,881	53.54	0.860	0.384	2.242	5.025
Child read outside school everyday	P2CHREAD	16,855	39.73	0.663	0.377	1.760	3.099
Parents had high school or less	WKPARED	15,733	36.19	0.939	0.383	2.451	6.006
Parent attended PTA	P2ATTENP	16,880	33.24	0.940	0.362	2.594	6.727
Parent thinks not safe for child to play outside	P2SAFEPL	16,866	96.19	0.219	0.147	1.487	2.212
Parents had happy relationship	P2MARRIG	12,823	97.83	0.168	0.129	1.305	1.702
Parent too busy to play with child	P2TOOBUS	16,642	45.73	0.590	0.386	1.528	2.335
Child characteristics (Mean)							
Age of child in months	R2_KAGE	18,146	74.76	0.072	0.033	2.171	4.713
Child's household size	P2HTOTAL	16,870	4.55	0.023	0.011	2.141	4.585
Number of children <18 in child's HH	P2LESS18	16,870	2.52	0.020	0.009	2.167	4.695
Median						1.857	3.447
Mean						1.884	3.748
Standard deviation						0.454	1.767
Coefficient of variation						0.241	0.471
Minimum						0.944	0.892
Maximum						2.860	8.182

Table 4-16. ECLS-K: median design effects for subgroups – child level

Subgroups	Fall-kindergarten[a] DEFT	Fall-kindergarten[a] DEFF	Spring-kindergarten[b] DEFT	Spring-kindergarten[b] DEFF	Panel[c] DEFT	Panel[c] DEFF
All students	2.162	4.675	2.030	4.120	1.857	3.447
Type of school						
Public	2.064	4.258	1.932	3.734	1.781	3.171
Private	1.995	3.979	1.954	3.817	1.782	3.174
Catholic private	1.771	3.136	1.738	3.022	1.654	2.736
Other private	1.937	3.754	1.706	2.910	1.709	2.920
Gender						
Male	1.771	3.135	1.735	3.011	1.533	2.349
Female	1.645	2.704	1.656	2.741	1.572	2.471
Race-ethnicity						
White	1.777	3.159	1.802	3.246	1.654	2.736
Black	1.594	2.546	1.462	2.137	1.417	2.009
Hispanic	1.397	1.952	1.406	1.977	1.366	1.865
Asian/Pacific Islander	1.971	3.883	2.107	4.438	1.422	2.021
Other	1.629	2.654	1.503	2.260	1.279	1.635
Region						
Northeast	1.760	3.099	1.824	3.328	1.541	2.374
Midwest	2.366	5.599	2.306	5.319	2.102	4.418
South	2.122	4.502	1.969	3.876	1.945	3.784
West	1.647	2.712	1.666	2.775	1.532	2.347
Urbanicity						
Central city	2.136	4.563	1.952	3.812	1.752	3.068
Urban fringe and large town	1.814	3.291	1.775	3.151	1.586	2.516
Small town and rural area	2.421	5.861	2.594	6.727	2.306	5.319

[a]Each median is based on 28 items.
[b]Each median is based on 27 items.
[c]Each median is based on 27 items.

Items with the highest design effects are those related to teacher data. For example, in spring-kindergarten, the estimate of the percent of children whose teachers have a master's degree or a higher degree is 35.6 percent with a design effect of 16.814; the estimate of the percent of children whose teachers spoke only English in class is 89.5 percent with a design effect of 17.871; the estimate of the mean number of years that these children's teachers taught in schools is nine years with a design effect of 12.157. The median design effect for these three items is about 14.8 for fall-kindergarten and 16.8 for spring-kindergarten. The high design effects are reasonable for this type of data because children in the same class have the same teacher, and the intraclass correlation is thus high.

Teacher-Level Design Effects

Standard errors and design effects for the teacher-level items are presented in table 4-17 for fall-kindergarten and table 4-18 for spring-kindergarten. Survey items were selected from both teacher instruments, part A and part B. In part A of the teacher instrument, teachers were asked to report about their children's and class's characteristics, for classes that they taught, whether they were morning, afternoon, or all day classes. In part B, they were asked about class organization, class activities, evaluation methods and also about their views on kindergarten readiness, school environment, and overall school climate. The topics covered in part B are not class-specific. Based on data collected in part A, teachers were classified as full-day (if they taught all day classes) or part-day teachers (if they taught morning, or afternoon, or both morning and afternoon classes). For both fall- and spring-kindergarten, a small number of teachers who filled out information for morning and all day classes or afternoon and all day classes could not be classified as part-day or full-day teachers and were excluded from the computation of design effects. This affects items such as the language that the teachers spoke in class and the class size since these estimates were computed over all classes taught, whether they were morning, afternoon, or all day classes.

The median design effect is 2.5 for both fall- and spring-kindergarten. These are lower than the child-level design effects because the number of responding teachers per school is relatively small. The design effect for teachers is largely a result of selecting a sample using a design most effective for child-level statistics.

Table 4-19 presents the median design effects for subgroups based on school type, geographic region, level of urbanicity, teacher type, and percent of minority enrollment in the school. For fall-kindergarten, the median design effect varies from 1.4 (teachers in Catholic private schools) to 3.2 (teachers in schools with 75 percent minority enrollment or more). The median design effects are generally lower for spring-kindergarten, but the range of variability is similar to that for fall-kindergarten (from 1.3 for teachers in Catholic private schools to 3.0 for teachers in schools with 25 to 49 percent minority enrollment).

Table 4-17. ECLS-K, fall-kindergarten: standard errors and design effects for the full sample – teacher level

Survey item	Variable name	Number of cases	Estimate	Design SE	SRS SE	DEFT	DEFF
Teacher characteristics (Percent)							
Used only English in class	A1ATNOOT, A1PTNOOT, A1DTNOOT	2,828	84.97	1.008	0.672	1.501	2.252
Used Spanish in class	A1ATSPNH, A1PTSPNH, A1DTSPNH	2,828	13.83	0.925	0.649	1.425	2.031
Had math area in class	B1MATHAR	3,037	94.71	0.712	0.406	1.753	3.072
Had computer area in class	B1COMPAR	3,031	83.31	1.095	0.677	1.617	2.615
Used 5-9 unpaid prep hours	B1NOPAYP	3,032	37.14	1.022	0.877	1.165	1.357
Had preschoolers in kindergarten	B1INKNDR	2,975	43.00	1.629	0.908	1.795	3.221
Teacher is Hispanic	B1HISP	2,973	6.53	0.766	0.453	1.692	2.862
Had at least a bachelor’s degree	B1HGHSTD	2,919	95.20	0.627	0.395	1.587	2.517
Had no teaching certification	B1TYPCER	2,923	3.20	0.401	0.325	1.233	1.520
Had highest teaching certification	B1TYPCER	2,923	62.05	1.409	0.897	1.570	2.464
Certified in early childhood	B1ERLYCT	2,941	52.95	1.607	0.920	1.746	3.048
Taught all day class only	A1ACLASS, A1PCLASS, A1DCLASS	2,860	62.38	2.231	0.906	2.463	6.068
Teacher characteristics (Mean)							
Age of teacher	B1AGE	2,923	41.29	0.249	0.191	1.301	1.693
Class size	A1ATOTAG, A1PTOTAG, A1DTOTAG	2,398	19.70	0.305	0.121	2.511	6.303
Had control of teaching techniques and discipline (scale 1 to 5)	B1CNTRLC	3,023	4.43	0.021	0.015	1.423	2.026
Number of years teaching kindergarten	B1YRSKIN	3,024	8.11	0.173	0.135	1.281	1.642
Median						1.579	2.491
Mean						1.629	2.793
Standard deviation						0.386	1.444
Coefficient of variation						0.237	0.517
Minimum						1.165	1.357
Maximum						2.511	6.303

Table 4-18. ECLS-K, spring-kindergarten: standard errors and design effects for the full sample – teacher level

Survey item	Variable name	Number of cases	Estimate	Design SE	SRS SE	DEFT	DEFF
Teacher characteristics (Percent)							
Used only English in class	A2AENGLS A2PENGLS A2DENGLS	3,037	89.22	0.925	0.563	1.643	2.700
Used Spanish in class	A2ACSPNH A2PCSPNH A2DCSPNH	3,037	11.68	0.935	0.583	1.604	2.573
Directed 2 hours of whole class activities	A2WHLCLS	3,032	33.80	1.132	0.859	1.318	1.737
Had daily reading and language arts	A2OFTRDL	3,063	94.82	0.720	0.400	1.798	3.234
Had math 3-4 times per week	A2OFTMTH	3,051	15.30	0.926	0.652	1.421	2.018
Had physical education 1-2 times per week	A2TXPE	3,060	54.61	2.301	0.900	2.557	6.539
Had adequate child size furniture	A2FURNIT	3,061	72.30	1.137	0.809	1.406	1.977
Attended 3 or more in-service training days	A2INSRVC	3,044	87.32	0.975	0.603	1.617	2.615
Parents see child's work 15 times or more	A2SHARED	3,046	22.36	1.226	0.755	1.624	2.637
Had math area in class	B1MATHAR	3,225	94.67	0.749	0.395	1.894	3.588
Had computer area in class	B1COMPAR	3,220	82.85	1.039	0.664	1.565	2.448
Used 5-9 unpaid prep hours	B1NOPAYP	2,970	37.13	1.054	0.887	1.188	1.412
Had preschoolers in kindergarten	B1INKNDR	3,170	43.42	1.582	0.880	1.797	3.230
Teacher is Hispanic	B1HISP	3,167	6.43	0.754	0.436	1.729	2.991
Had at least a bachelor's degree	B1HGHSTD	3,113	94.76	0.606	0.399	1.517	2.302
Had no teaching certification	B1TYPCER	3,114	3.44	0.445	0.326	1.363	1.858
Had highest teaching certification	B1TYPCER	3,114	61.50	1.290	0.872	1.479	2.187
Certified in early childhood	B1ERLYCT	3,137	52.50	1.565	0.892	1.755	3.079
Teacher characteristics (Mean)							
Age of teacher	B1AGE	3,111	41.11	0.253	0.187	1.354	1.834
Number of paid aides	A2PDAIDE	3,053	0.90	0.030	0.014	2.148	4.615
Had control on teaching techniques and discipline (scale 1 to 5)	B1CNTRLC	2,962	4.43	0.022	0.015	1.461	2.134
Number of years teaching kindergarten	B1YRSKIN	3,219	8.09	0.162	0.131	1.234	1.522
Median						1.585	2.511
Mean						1.612	2.692
Standard deviation						0.311	1.139
Coefficient of variation						0.193	0.423
Minimum						1.188	1.412
Maximum						2.557	6.539

Table 4-19. ECLS-K: median design effects for subgroups – teacher level

Subgroups	Fall-kindergarten[a]		Spring-kindergarten[b]	
	DEFT	DEFF	DEFT	DEFF
All teachers	1.579	2.491	1.585	2.511
Type of school				
Public	1.542	2.379	1.548	2.397
Private	1.319	1.738	1.243	1.544
Catholic private	1.162	1.360	1.147	1.315
Other private	1.226	1.503	1.212	1.467
Region				
Northeast	1.410	1.987	1.415	2.005
Midwest	1.518	2.305	1.512	2.284
South	1.561	2.437	1.559	2.429
West	1.532	2.349	1.549	2.398
Urbanicity				
Central city	1.682	2.830	1.547	2.393
Urban fringe and large town	1.534	2.356	1.484	2.202
Small town and rural area	1.617	2.616	1.700	2.893
Type of teacher				
Full day	1.513	2.290	1.638	2.683
Part day	1.411	1.990	1.339	1.793
Minority enrollment				
0 – 25%	1.368	1.871	1.367	1.869
25 – 49%	1.765	3.113	1.738	3.022
50 – 74%	1.424	2.027	1.406	1.977
75 – 100%	1.776	3.166	1.453	2.110

[a] Each median is based on 16 items.
[b] Each median is based on 22 items.

School-Level Design Effects

Standard errors and design effects for the school-level items are presented in table 4-20. Survey items are selected from the school administrator questionnaire. For items having to do with children with limited English proficiency (LEP), the data were subset to schools with LEP children. The median design effect is 1.6. Table 4-21 presents the median design effects for subgroups based on school type, geographic region, level of urbanicity, and percent of minority enrollment in the school. They vary from 1.1 for schools in the Northeast region to 2.1 for schools in small towns and rural areas.

Table 4-20. ECLS-K, spring-kindergarten: standard errors and design effects for the full sample – school level

Survey item	Variable name	Number of cases	Estimate	Design SE	SRS SE	DEFT	DEFF
School characteristics (Percent)							
Had a particular focus or emphasis	S2FOCUS	859	19.94	1.736	1.363	1.274	1.622
Used standardized achievement assessment as requirement for admission	S2STNDTE	856	10.84	1.392	1.063	1.310	1.715
Funding levels decreased significantly	S2FUNDLV	854	12.50	1.382	1.132	1.221	1.491
Received federal Title 1 funds this year	S2TT1	862	51.51	1.450	1.704	0.851	0.725
Required kindergartners to wear uniform	S2UNIFRM	858	18.92	1.936	1.337	1.448	2.096
Gave children readiness or placement assessment	S2RDITST	860	61.74	2.391	1.657	1.443	2.082
Tested kindergartners with standardized assessments	S2TESTK	838	33.40	2.030	1.629	1.246	1.552
Offered after- school child care	S2AFTSCH	856	56.61	2.187	1.694	1.291	1.666
Offered pre-kindergarten	S2PREKIN	856	45.47	2.149	1.702	1.263	1.594
Had LEP children	S2LIMENG	857	39.07	1.980	1.667	1.188	1.412
Had translators for LM-LEP families	S2TRANSL	414	77.69	2.928	2.046	1.431	2.047
Offered IEP to disabled children	S2ONIEP	853	80.77	1.714	1.350	1.270	1.613
Principal is male	S2GNDER	848	39.16	2.186	1.676	1.304	1.701
Principal is black	S2RACE3	820	6.94	0.991	0.888	1.116	1.246
Principal has master’s degree	S2EDLVL	806	87.49	1.665	1.166	1.428	2.040
School characteristics (Mean)							
Percent LEP students	S2LEPSCH	387	11.99	1.542	1.002	1.539	2.370
Total years as principal	S2TOTPRI	840	10.38	0.300	0.268	1.118	1.249
Median						1.274	1.622
Mean						1.279	1.660
Standard deviation						0.162	0.395
Coefficient of variation						0.127	0.238
Minimum						0.851	0.725
Maximum						1.539	2.370

Table 4-21. ECLS-K: median design effects for subgroups – school level

Subgroups	Spring-kindergarten[a]	
	DEFT	DEFF
All schools	1.274	1.622
Type of school		
Public	1.312	1.721
Private	1.150	1.323
Catholic private	1.105	1.220
Other private	1.079	1.165
Region		
Northeast	1.045	1.092
Midwest	1.202	1.445
South	1.374	1.888
West	1.414	2.000
Urbanicity		
Central city	1.279	1.635
Urban fringe and large town	1.223	1.496
Small town and rural area	1.445	2.088
Minority enrollment		
0 – 25%	1.213	1.471
25 – 49%	1.246	1.552
50 – 74%	1.182	1.396
75 – 100%	1.422	2.023

[a] Each median is based on 17 items.

5. DATA COLLECTION METHODS AND RESPONSE RATES

5.1 Data Collection Methods

The following sections discuss the procedures used in the data collection phase of the Early Childhood Longitudinal Study-Kindergarten Class of 1998-99 (ECLS-K) base year.

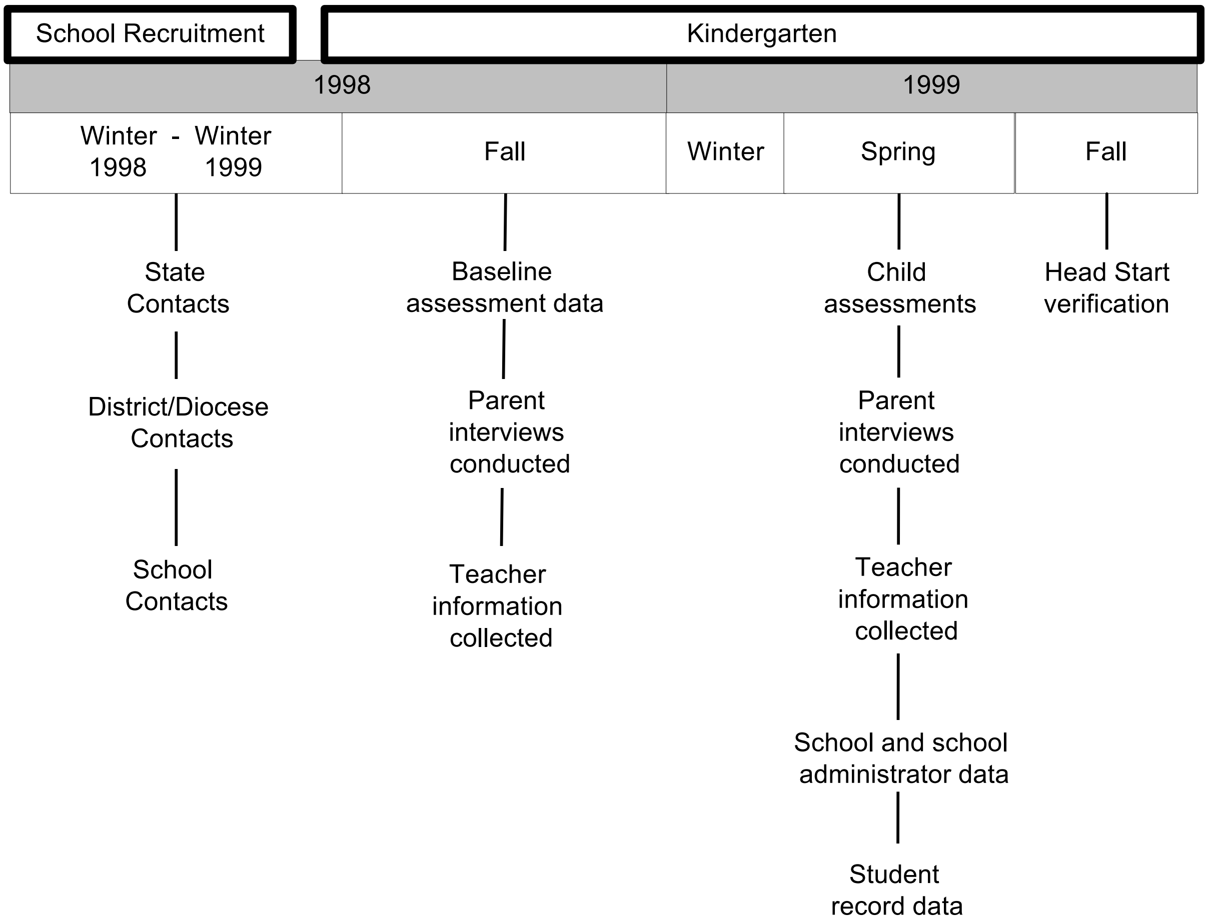
The ECLS-K base year data collection was conducted in the fall and spring of the 1998-99 school year. The timeline for base year data collection is shown in exhibit 5-1. Computer-assisted personal interviewing (CAPI) was the mode of data collection for the child assessment, and computer-assisted telephone/personal interviewing (CATI/CAPI) was the mode of data collection for the parent interview; self-administered questionnaires were used for gathering information from teachers, school administrators, student records, and about teacher salary and benefits. The field staff was organized into 100 work areas, each with a data collection team consisting of one field supervisor and three 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, distributed and collected all school and teacher questionnaires, and completed a school facilities checklist.

5.2 Study Endorsements and School Recruitment

Prior to contacting state agencies or dioceses for a commitment to participate, key educational organizations were contacted for endorsement. The ECLS-K study received the endorsement of many national associations and organizations representing parents, school administrators, teachers, and private religious and nonreligious schools.

Before the data collection effort could begin, it was first necessary to secure a commitment to participate in the study from the administrator of each sampled school. Several levels of cooperation were sought before school administrators were approached. The process of notifying states, districts, dioceses, and schools about the study began in January 1998. The contact procedures varied for public, Catholic, and non-Catholic private schools. Public schools had three levels of contact—state, school district, and school; Catholic schools had two levels—diocese and school; and non-Catholic private schools had one—the school.

Exhibit 5-1. Timeline of base year data collection



State Contacts for Public Schools

Letters were sent to the chief state school officer of each of the 41 states that contained ECLS-K sampled schools to explain the objectives of the study and the data collection procedures, especially those for protecting individual and institutional confidentiality.[1] All states agreed to participate. The letter indicated the state test director would be called to discuss the study. Information packets that included a list of the public school districts and sampled public schools within the state were then sent to the state test directors. The state test directors were called to identify the procedures to be followed in contacting public school districts. Once approval was obtained at the state level, contact was made with the district superintendents.

[1] The ECLS-K area sample was not stratified by state and as a result not all states contain sampled schools. The smallest geographic area for which direct representation was guaranteed is Census region. For more information concerning the sample design see chapter 4.

District and Catholic Diocese Contacts

For the public schools, a letter that included a list of the sampled schools within the district was sent to the district superintendent. A similar package of materials was sent to the Catholic dioceses and Archdioceses in the sample to obtain permission to contact Catholic schools about the study. Beginning in late January, calls were placed to district superintendents and Catholic dioceses to explain the study, answer questions, and obtain permission to contact sampled schools within the district or diocese. Once approval was obtained at the district or diocesan level, contact was made with each school administrator. A total of 584 public school districts contained sampled schools, and 75 percent of them cooperated with the study. (See table 5-4 for school response rates.) A total of 55 Catholic dioceses or Archdioceses contained sampled schools, and 93 percent of them cooperated with the study (see table 5-4 for school response rates).

School Contacts

A letter and other study materials were mailed to school administrators in all the sampled schools. The school administrators in non-Catholic private schools were also contacted at this time. Telephone contacts began in March 1998 and continued through the end of the school year in June 1998; school recruitment continued in fall 1998 with the start of the new school year. To increase the number of participating public schools, substitute schools were recruited in primary sampling units (PSUs) where the public school response rate was less than 65 percent. Once the school administrator agreed to participate, he or she was asked to set an appointment for two visits by the ECLS-K field staff to the school in the fall of the 1998-99 school year. The first visit, the preassessment visit, was to select the sample of children (see section 5.4.2 for more detail on this visit), and the second visit was to conduct the child assessments (see section 5.4.3 for more detail on this visit). The school administrator was also asked to identify an individual, referred to as the school coordinator, to act as the school liaison with the ECLS-K.

Increasing Response Rates for Originally Sampled Schools

Recruitment of originally sampled schools continued into winter 1999 to improve school response rates before school substitution. Beginning in January 1999, a small group of highly experienced field supervisors converted originally sampled schools that had initially refused to participate in the fall-kindergarten assessments. Refusal conversion materials were developed that included endorsements from school principals, kindergarten teachers, and parents. Also, a fact sheet presented general information about the direct child assessment and the response to it by children, teachers, and parents who participated in the fall ECLS-K. Field supervisors made telephone or in-person contacts to refusal schools to encourage them to participate in the study. The refusal conversion effort continued through March 1999 and improved original school response rates by five percentage points.

5.3 Field Staff Training

In-person training sessions were conducted to prepare field staff for fall data collection tasks. Three different training sessions for fall-kindergarten were conducted: one for staff recruiting schools into the study, one for field supervisors, and one for assessors. Two different training sessions for spring-kindergarten were conducted: one for continuing staff and one for new staff.

School Recruitment Training

School recruitment staff—primarily field supervisors with experience on other education studies such as the National Assessment of Educational Progress (NAEP) and the Third International Mathematics and Science Study (TIMSS)—were trained for three days in March 1998 to recruit public school districts, Catholic dioceses, and private schools into the study. The topics covered in the training included an introduction to the study, practice exercises in recruiting schools, refusal avoidance techniques, and exercises on scheduling schools efficiently within a work area.

Fall-Kindergarten Field Staff Training

Field supervisor and assessor training was held during two time periods between late August and late September 1998. Field staff were selected who had experience on NAEP and TIMSS, as well as the Panel Study of Income Dynamics—Child Supplement and the Monitoring the Future Study. The majority of selected field staff were retired teachers, former educators, people experienced in conducting assessments, or people experienced in working in schools or with school-age children. Prior to the in-person training session, all field staff received eight hours of home study training on the study design, including a training video demonstrating the conduct of the direct child assessment, the field procedures, and computer keyboard skills, and eight hours of general interviewing techniques training. Field supervisors were trained to conduct supervisor activities in groups of 15. Field supervisors and assessors were trained for assessment activities and parent interviewing in groups of 20 to 24. In total, 112 field supervisors and 343 assessors completed training.

Field Supervisor Training. Field supervisor training preceded the assessor training and lasted for two days. Field supervisor 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, providing them with hands-on experience with the FMS and all field procedures prior to beginning data collection. The topics covered in the field supervisor training session included reviewing materials for prior contacts with schools; role plays to practice contacting school coordinators; conducting the preassessment activities such as selecting the student samples, identifying student home language, obtaining parent consent, and parent contacting procedures; meeting with teachers; and quality control observation procedures.

Assessor Training. One major goal of the assessor training was to train field staff in the proper procedures to conduct the direct child assessment. This included following standardized procedures for administration of all assessment items as well as maintaining a neutral rapport with the sampled children. Another major goal was to engage parents in the study to obtain consent to conduct the direct child assessment and complete the parent interview. The topics covered in the assessor training session included an introduction to the study, interactive lectures based on the direct child assessment and parent interview, practicing parent interviews in pairs using role-play scripts, practicing the direct child assessment using role-play scripts, techniques for parent refusal avoidance, and strategies for building rapport with kindergarten children. The sessions provided trainees with hands-on experience with all the direct child assessment materials and procedures and the CAPI programs prior to beginning data collection. Assessor training lasted for five days; field supervisors were also trained to perform all assessor activities. Trainees practiced entering information into the CAPI system on laptop computers during training presentations on conducting the direct child assessment and parent interview. The majority of training time (about 30 hours) was spent on conducting the direct child assessment, establishing rapport with the child, and maintaining a neutral approach. The trainees practiced the direct child assessment with kindergarten children brought onsite to the training session. These practice sessions and direct interactions with children provided useful experience prior to assessing sample children. The rest of the time was spent on conducting the parent interview. Spanish bilingual interviewers underwent an additional six hours of training during which they practiced the Spanish version of the direct child assessment [a Spanish Oral Language Development Scale (OLDS), a Spanish warm-up booklet, and translated versions of the mathematics and the psychomotor assessment] and the parent interview, and practiced gaining cooperation and answering parents' questions in Spanish (see section 5.7 on data collection quality control.)

Spring-Kindergarten Training

Field supervisors and assessors were trained over a four-week period in March 1999. The majority of field staff was trained primarily using a home study training program. The purpose of the home study training program was to introduce changes to the instruments between the fall and spring data collection and provide sufficient practice with the instruments and provide review and practice of the procedures to conduct child assessments and parent interviews. An in-person attrition training session was conducted for staff new to the study in the spring. This training session replicated the comprehensive fall training program using the revised instruments for the spring data collection. In total, 101 field supervisors and 321 assessors completed training.

Home Study Training Program. Field supervisor training was again conducted using the FMS. The field supervisor home study program was 24 hours long. The topics covered in the field supervisor training session included conducting the preassessment activities such as selecting the student samples in converted schools, identifying children who moved from the originally sampled school, identifying regular and special education teachers, linking them to children and distributing teacher questionnaires, distributing school administrator questionnaires and completing the facility checklist, and following quality control observation procedures. The field supervisors again entered training information into the FMS during the home study training exercises. The topics covered in the field supervisor training session included contacts with schools, conducting the preassessment activities, identifying children no longer in their originally sampled school, and following quality control observation procedures. A total of 89 field supervisors completed the home study training.

Child assessment home study materials included written exercises, self-study role-plays, scripted role-plays with partners, a training video focusing on improving assessment skills, and a final practice direct child assessment observed by the field supervisor. Parent interview home study materials included written exercises and trainees practicing the parent interviews in pairs on the telephone and completing a final role-play on the parent interview with their field supervisor. Assessors again entered training information into the CAPI system on laptop computers using role-play scripts to complete the home study exercises. The assessor home study program was 16 hours long. Field supervisors ensured that the assessors on their teams were proficient in conducting the direct child assessment by observing a practice assessment with a child and the parent interview by conducting the final role play. A total of 223 assessors completed the home study training (see section 5.7 on data collection quality control) .

Attrition Training Session. The in-person attrition training for new interviewers was held in March 1999 for nine days. The fall-kindergarten training materials were slightly modified for this attrition training to incorporate some spring-kindergarten changes, such as the elimination of the psychomotor assessment in the direct child assessment. The new spring-kindergarten parent and child home study materials were also incorporated into the attrition training sessions. A total of eight field supervisors and 98 assessors completed the attrition training.

5.4 Fall-Kindergarten Data Collection

5.4.1 Advance Mailings

Beginning in late summer 1998, letters were mailed to school administrators to confirm scheduled visits for the schools. A packet of material was also mailed to the school coordinators asking them to prepare for the preassessment visit to the school. Beginning in September, field supervisors called school coordinators to confirm the dates of the preassessment and assessment visits, answer any questions, and prepare for the preassessment visits. The school coordinators were asked to prepare a list of kindergartners for selecting the sample and to distribute materials such as the study brochure, summary sheets describing the role of teachers and parents in the study, and a letter to teachers to the kindergarten teachers.

5.4.2 Preassessment Visits

Most preassessment visits were scheduled two weeks before the assessment visit to allow time to contact parents to obtain their consent for the sampled children to participate in the study. During the preassessment visit, the field supervisor met with the school coordinator to obtain the lists of kindergarten children and teachers, draw the student sample (see chapter 4, section 4.2.4), meet with kindergarten teachers to secure their cooperation, distribute teacher questionnaires, and identify home language and accommodations information about the sampled children. In year-round schools, which comprised four percent of cooperating schools, field supervisors selected the child sample across the multiple tracks within the school. Field supervisors also prepared and distributed parent information packets that included consent forms.

Distribution of Teacher Questionnaires

During the meeting with the kindergarten teachers, the self-administered teacher questionnaires were distributed. All kindergarten teachers were asked to complete teacher questionnaire A and teacher questionnaire B; teacher questionnaire C was child-specific and completed by the teacher about each of the sampled children. If more than one teacher was in the kindergarten classroom, supervisors determined the primary teacher for completing teacher questionnaire C. The primary teacher was defined as the teacher who had primary academic responsibility for the children and could best report on the information being sought about the sampled children. Field supervisors made arrangements with the kindergarten teachers to collect completed questionnaires during the assessment visit.

Identifying Home Language and Accommodations for the Direct Assessment of Sampled Children

Field supervisors obtained information about each child who was scheduled to be assessed to determine whether the measure of oral English language (OLDS) would be administered and whether any administrative accommodations had to be made in the direct assessment. The reading and general knowledge assessment batteries were designed to be administered only in English. The warm-up booklet, mathematics, and psychomotor assessments were, however, translated into Spanish for children who could not be assessed in English but were able to take the assessment in Spanish. To determine whether language minority children could be validly and reliably assessed in English using the core direct assessment battery, the OLDS, a measure of basic oral English proficiency, was administered (see chapter 2, section 2.1 for details on the OLDS).

To determine whether languages other than English were used in the home, the school coordinator was asked to consult school records of the sampled children. If the information was not in the school records, the field supervisor asked the teachers to report on home language during the group teacher meeting. A series of four questions was used to ask the teacher about the children's home language:

- What are the names of the children in your classroom who speak a language other than English?
- What other children in the classroom have families who speak a language other than English in the home?
- What other children in your classroom have you observed participating in a conversation with peers or adults where a language other than English was used?
- What language other than English does the child speak?

Children identified with a home language other than English were first assessed using the English OLDS. Depending on the child's English OLDS score, the laptop was programmed to administer the appropriate version of the assessment (see section 5.4.3 for details on data collection for children with a home language other than English).

Field supervisors identified what accommodations, if any, needed to be made for children with disabilities to administer the direct assessment battery appropriately. For the purpose of the ECLS-K, children with disabilities are those who meet the federal eligibility requirements for participation in special education programs or services. All children with disabilities are expected to have an Individualized Education Plan (IEP), an Individualized Family Service Plan (IFSP), or a 504 Plan on file with the school district as it is a required component of the eligibility process.

The ECLS-K permitted the following accommodations for children with disabilities:

- **Setting.** Any environmental modification typically used by the student to do his or her schoolwork, such as special lighting, a quiet room, an adaptive chair or table was used for the assessment.

- **Scheduling/timing.** Assessments of children with disabilities were scheduled at particular times of the day that were best suited for the child or for longer or shorter (split) periods.

- **Health care aides (personal attendants).** Individuals who were actually written into the student's IEP to perform a variety of services for both the student and his or her teacher were allowed to be present during the assessment, but they could not provide answers or hints to the student. A form entitled "Guidelines for Health Care Aides" was reviewed with the aide prior to the assessment.

- **Assistive devices.** Assistive devices such as a brace, a hearing aid, a cane, or a voice synthesizer were legitimate to use during the ECLS-K to improve access to the assessment without providing answers.

Braille administration, enlarged print, and sign language administration were not valid accommodations for the ECLS-K. Children who required these accommodations were excluded from the direct assessment portion of the ECLS-K. These children, however, remain in the sample and all other data are captured for them.

A series of questions was asked of the teacher to determine if the student would be:

- Included in the assessment without accommodations;

- Included in the assessment with accommodations; or

- Excluded from a part or all of the assessment.

The questions were:

1. What are the names of the children in your class who have an IEP, IFSP, or 504 Plan on file?

2. Does the child's IEP or equivalent state that the child cannot participate in standardized assessments?

3. Does the child normally use and require any of the accommodations that are NOT offered in the ECLS-K? (Braille, enlarged print, sign language)?

4. Does the child normally use any of the accommodations that ARE offered in the ECLS-K (setting, scheduling/timing, health care aide, assistive device)?

5. What are the names of the children in your class who cannot perform the following fine and gross motor tasks?

   a. Hold a pencil or crayon to make marks on a piece of paper?

b. Grasp small objects and release them at will?

c. Demonstrate his or her skill at hopping, skipping, and/or walking backward due to impaired mobility or use of assistive devices such as wheelchairs, crutches, and braces?

6. What are the names of the children in your class who normally wear glasses or hearing aids?

7. Are there any children with other kinds of special needs that should be noted for the assessment (e.g., medications, allergies, shyness, etc.)?

Question 5 was asked to determine if the child could perform the fine and gross motor activities for the psychomotor assessment. If the child had an IEP or equivalent that required the use of an accommodation not offered by the ECLS-K, or could not perform the question 5 activities, the laptop computer was programmed to exclude the child from those sections of the assessment battery in which the child could not participate.

Obtaining Parent Contact Information and Consent

Obtaining parental contact information and consent was another important procedure of the preassessment period. Parent contact information was necessary to obtain consent and to conduct the parent interview. Two types of consent forms—implicit and explicit—were developed for the ECLS-K. Implicit consent did not require parent/guardian signature for the child to participate. Approximately one-half of the cooperating schools required implicit parent consent forms. If the consent form was not returned to the school with a signature and indicating that the child was not to participate by the date recorded, permission for the child to participate was assumed. Explicit consent required a parent/guardian's signature to allow his or her child to participate in the study. The consent form was included in parent information packets, which were mailed directly to the parents of sampled children. The study followed the consent procedure required by the school or district.

Preassessment visits to schools requiring explicit consent were scheduled as early as possible in the field period to allow for sufficient time to obtain consent before the assessment date. If a parent indicated that he or she did not want the child to participate in the assessment or did not return the consent form, refusal conversion attempts were made. Generally, a team member would place up to four calls to the home to gain the parent's cooperation. If there was no positive response to the telephone contacts, a visit to the home was made to encourage participation. The assessment team also attempted to contact parents to obtain parent consent while they were dropping off their children before school or picking them up after school. If the parent refusal could not be converted, the child was not assessed.[2]

5.4.3 Conducting the Direct Child Assessment

For the fall-kindergarten wave, the direct child assessment was administered during a 14-week field period that began in September and ended in early December. In year-round schools, assessment teams made multiple visits to the school to conduct direct child assessments. The assessment team visited the school when each track was in session to assess the sampled children. The direct child assessment was normally conducted in a school classroom or library. Before conducting the assessment, field supervisors and assessors set up the room for the assessment. They followed procedures for meeting children that were agreed upon during the preassessment visit at the school. Each child was signed out of his or her classroom prior to the assessment and signed back into the classroom upon the conclusion of the assessment. When scheduling schools in the fall and the spring, an attempt was made to conduct the direct child assessments at about the same point in time from the beginning of school year and at the end of the year to increase the chances that children’s exposure to instruction was about the same for all children.

The direct child assessment took approximately 50 to 70 minutes per child. Children identified as having a home language other than English were administered the OLDS prior to the assessment. Children passing the cut score for the OLDS were administered the English direct child assessment. Children who fell below the cut score for the OLDS and whose language was Spanish were administered the Spanish language version of the OLDS and parts of the direct child assessment that were translated into Spanish (the warm-up booklet, math, and psychomotor). These children also had their height and weight measured. Children who fell below the cut score for the OLDS and whose language was other than Spanish had only their height and weight measured. Table 5-1 presents the percentage of children who were routed into the various assessment alternatives in fall-kindergarten. Overall 15 percent of the sampled children were screened using the OLDS in the fall-kindergarten. Of the children whose home language was Spanish, 42 percent were at or above the cut score, and of the children whose home language was a language other than English or Spanish, 61 percent were at or above the cut score.

[2] Schools requiring signed explicit parent consent forms accounted for 69 percent of the final refusals.

Table 5-1. OLDS routing results in fall-kindergarten

	Total screened (%)	At or above cut score on OLDS (%)	Below cut score on OLDS (%)
Total screened	15	49	51
Spanish language	62	42	58
Other language	38	61	39

Fewer than 1 percent of participating children were excluded from the direct child assessment or required an accommodation offered in the assessment. Table 5-2 presents the number of children excluded from or requiring an accommodation to the direct child assessment.

Table 5-2. Number of children excluded from or accommodated in the fall-kindergarten assessment

	Number of children
Excluded for disability	88
Setting	48
Scheduling/timing	93
Health care aide	37
Assistive device	4

A very small number of children who participated in the ECLS-K assessment were excluded from different parts of the psychomotor assessment because they could not physically perform the task. Table 5-3 presents the number of children excluded from different parts of the psychomotor assessment.

Table 5-3. Number of children excluded from parts of the psychomotor assessment

	Number of children
Excluded from gross motor assessment	70
Excluded from fine motor assessment	5
Excluded from both gross and fine motor assessment	7

5.4.4 Conducting the Parent Interview

For the fall-kindergarten wave, the parent interview was administered from September 1998 through January 1999. The parent interview averaged 50 minutes. To administer the parent interview, assessors began by locating parents using the contact information obtained from the school. The contact information contained parents' names, addresses, and telephone numbers. Upon completing the interview, parents were mailed a thank-you letter and a copy of "Learning Partners – A Guide to Educational Activities for Families," May 1997, U.S. Department of Education, Office of Educational Research and Improvement.

The ECLS-K parent interview was conducted primarily by telephone, by field staff using a CATI. The parent interview was conducted in person if the respondent did not have a telephone. Three percent of all completed parent interviews were conducted in person. The parent interview was conducted primarily in English, but provisions were made to interview parents who spoke only Spanish, Lakota, Hmong, or Chinese. The questionnaire was translated into Spanish, Lakota, Hmong, and Chinese and printed on a hard copy parent interview form. Bilingual interviewers were trained to conduct the parent interview in either English or the other language they spoke. When the person answering the telephone was not able to speak English, and the field staff were not bilingual and were unable to identify an English-speaking household member, the case was coded as a "language problem." The field supervisor reviewed the case and assigned it to a bilingual field staff person if the language was offered by the ECLS-K. Slightly more than seven percent of the parent interviews were conducted in a language other than English; 94 percent of the non-English parent interviews were conducted in Spanish. One percent of parent interviews could not be conducted because of language problems.

In order to build response rates for the parent interview, a special refusal conversion effort was implemented. A refusal conversion workshop was developed from three focus groups with field staff to determine the types of resistance being expressed by parents. The focus group information was used to generate a framework for the workshop materials and procedures. A written conference guide was distributed to field staff identified as effective at converting refusals, and they completed written exercises on refusal conversion. Conference calls were held with the identified field staff to review the refusal conversion exercises and discuss strategies for converting refusals.

In addition to the specialized field staff training, a parent refusal conversion mailing was prepared that included a refusal conversion letter; copies of letters from participating school principals and kindergarten teachers endorsing the study; and the parent thank-you incentive, “Learning Partners – A Guide to Educational Activities for Families.” These endorsement letters from principals and teachers were addressed to the parents in the school with very positive statements about the study, encouraging the parents to cooperate and participate in the ECLS-K. The school principals and teachers agreed to allow the use of their letters and quotes about the study. The refusal conversion cases were assigned to a group of specially trained field staff. All refusal cases were reviewed by the field managers, and a refusal conversion strategy was discussed with the refusal converter. The refusing parents were all called, and a conversion attempt was made. The refusal conversion effort was implemented for a month, from December 20, 1998, to January 19, 1999, and response rates improved by five percentage points.

5.4.5 Teacher Data Collection

Data were collected from teachers between September and December 1998. Field supervisors distributed the teacher questionnaires to kindergarten teachers in the sampled schools during the preassessment visit. The field supervisors collected the teacher questionnaires while at school during the assessment visits. Field supervisors contacted teachers by telephone or by visiting the school to collect completed teacher questionnaires that were not available during the assessment visits. Teachers were asked to complete individual ratings for the sampled children in their classrooms, and they were reimbursed five dollars for each child rating (Teacher Questionnaire C) they completed.

5.5 Fall-Kindergarten Response Rates

Table 5-4 presents the number of schools cooperating during the fall-kindergarten data collection, as well as weighted and unweighted before-substitution school response rates. Substitute schools were recruited in PSUs where the public school response rate was low, i.e., where the unweighted school response rate was less than 65 percent and fewer than six originally selected schools were cooperating. Substitutes were purposively selected. For each refusing school, the frame was searched for a suitable replacement school, if any, based on characteristics such as size, percent minority, and type of locale. The number of cooperating schools was increased from 879 to 953 by recruiting substitutes. Substitutes were not selected for private schools. The ECLS-K design called for 20 percent of the base year student sample to come from private schools. Projections indicated that 20 percent would be attained

Table 5-4. The ECLS-K, fall-kindergarten: number of cooperating schools and before-substitution school-level response rates

School characteristic[a]	Number cooperating[b] Before substitution	After substitution	Before-substitution response rates School weight × KG enrollment	Unweighted
All schools	879	953	69.4	68.8
School type				
Public	641	715	70.1	70.1
Private	238	238	65.9	65.6
Catholic	100	100	83.0	83.3
Other private	138	138	56.2	56.8
Type of locale				
Large city	191	198	78.1	78.0
Mid-size city	180	197	73.6	72.9
Urban fringe of large city	255	271	67.8	66.8
Urban fringe of mid-size city	59	66	62.1	60.8
Large town	20	20	57.5	60.6
Small town	71	82	65.3	63.4
Rural	103	119	64.2	64.0
School size (Kindergartners)				
1 to 12	55	55	74.5	69.6
13 to 24	96	98	63.2	64.0
25 to 36	75	79	71.5	74.3
37 or more	653	721	69.7	69.0
Percent minority enrolled				
0 – 10%	225	249	66.1	64.8
11 – 49%	285	324	68.4	67.5
50 – 89%	180	188	70.9	72.0
90 – 100%	135	138	78.5	78.5
Unknown	54	54	62.9	62.8
Region				
Northeast	160	181	65.1	65.8
Midwest	231	238	77.4	77.5
South	268	294	67.0	64.1
West	220	240	68.7	69.2
High grade				
Trans K, K or trans 1	40	40	78.4	78.4
01-12	800	874	69.4	69.0
Ungraded	4	4	100.0	100.0
Unknown	35	35	57.5	56.5

[a]Based on frame data.

[b]One or more children or parents participated, or school is one of two schools in the Archdiocese of New Orleans where student sampling was performed too late in fall to field child or parent data collection.

through the private school freshening process (see section 4.2.3), without resorting to substitution, and those projections were later borne out by experience. The response rates presented in table 5-4 reflect levels of school cooperation prior to substitution. The presence of substitutes increases the sample size and, if chosen skillfully, may also reduce nonresponse bias. However, the bedrock upon which the study's quality rests is the originally selected sample of schools.

Since the primary analytic focus of the ECLS-K is the child, weighted school response rates are reported in terms of eligible children. For example, overall, the kindergartners enrolled in the ECLS-K fall-kindergarten cooperating schools are weighted up to 69.4 percent of the fall 1998 kindergarten population. The unweighted response rate is 68.8 percent. The weighted response rate for public schools (70.1 percent) was about four points higher than for private (65.9 percent). However, within the private school domain there was a much greater disparity between Catholic schools (83.0 percent) and other private schools (56.2 percent). During the computation of sampling weights, nonresponse adjustment cells were constructed with an eye toward reducing the potential biases associated with these rates.

Table 5-5 presents weighted and unweighted child-level completion rates for fall-kindergarten data collection, broken out by school characteristics. A completion rate is a response rate conditioned on the results of an earlier stage of data collection. For example, a weighted 89.9 percent of all children sampled in cooperating schools completed the child assessment. Of course as noted earlier, not all sampled schools cooperated. In this table and the ones that follow, data from cooperating substitute schools have been included along with those from original schools. Since parent and teacher data appear on the child data file (see chapter 8), rates have been presented at the child-level for these instruments as well. If one looks at subpopulations defined by school characteristics, it is striking how characteristics that were important during school recruitment are no longer so for schools that agreed to participate. For example, comparing public and private schools, the completion rate for the child assessment differs only by 0.3 percent, with private schools having the higher rate (90.2 vs. 89.9), and Catholic and other private schools differ by 2.2 percent (91.3 vs. 89.1).

Table 5-6 continues presenting child-level fall-kindergarten weighted and unweighted completion rates, this time broken out by child characteristics. Again, completion rates do not differ greatly except possibly for the parent interview where Asian and Pacific Islander parents responded at a rate of 74.1 percent, largely because of language problems, in contrast with a rate of 85.3 percent for parents overall.

Table 5-5. The ECLS-K fall-kindergarten: number of completed child-level cases and child-level completion rates, by school characteristic

	Child assessment			Parent interview		
		Completion rates			Completion rates	
School characteristic[a]	Completes[b]	Weighted	Unweighted	Completes[c]	Weighted	Unweighted
All children	19,173	89.9	89.8	18,097	85.3	84.7
School type						
Public	15,229	89.9	89.7	14,283	85.0	84.1
Private	3,944	90.2	90.1	3,814	87.5	87.2
Catholic	2,007	91.3	91.3	1,947	88.6	88.6
Other private	1,937	89.1	88.9	1,867	86.5	85.7
Type of locale						
Large city	3,954	88.2	88.1	3,611	80.8	80.5
Mid-size city	4,135	91.1	91.1	3,977	87.8	87.6
Urban fringe of large city	5,452	89.5	89.0	5,092	84.4	83.1
Urban fringe of mid-size city	1,359	89.5	90.1	1,302	85.6	86.3
Large town	416	91.8	92.0	415	91.9	91.8
Small town	1,601	91.0	91.1	1,532	86.9	87.2
Rural	2,256	90.8	90.6	2,168	87.6	87.1
School size (Kindergartners)						
1 to 12	366	86.1	85.7	358	85.7	83.8
13 to 24	1,550	90.5	91.7	1,478	86.9	87.5
25 to 36	1,640	91.8	91.6	1,577	88.0	88.1
37 or more	15,617	89.8	89.5	14,684	84.9	84.2
Percent minority enrolled						
0 – 10%	5,134	90.0	89.9	5,027	88.2	88.1
11 – 49%	6,704	90.0	89.6	6,405	85.8	85.6
50 – 89%	3,637	90.0	89.8	3,381	84.6	83.5
90 – 100%	2,905	89.9	89.7	2,514	78.8	77.7
Unknown	793	88.9	90.3	770	87.9	87.7
Region						
Northeast	3,605	89.0	88.9	3,373	83.4	83.2
Midwest	4,769	89.8	90.1	4,575	86.1	86.5
South	6,126	91.5	91.1	5,853	87.7	87.1
West	4,673	88.4	88.3	4,296	82.3	81.2
High grade						
Trans K, K or trans 1	552	90.4	88.9	526	85.8	84.7
01-12	18,174	90.0	89.9	17,135	85.3	84.7
Ungraded	40	74.6	71.4	35	56.7	62.5
Unknown	407	90.0	89.8	401	88.7	88.5

Table 5-5. The ECLS-K, fall-kindergarten: number of completed child-level cases and child-level completion rates, by school characteristic (continued)

	Teacher questionnaire A			Teacher questionnaire B			Teacher questionnaire C		
		Completion rates			Completion rates			Completion rates	
School characteristic[a]	Completes[d]	Weighted	Unweighted	Completes[d]	Weighted	Unweighted	Completes[d]	Weighted	Unweighted
All children	19,444	90.8	91.0	20,647	96.6	96.7	19,492	91.4	91.3
School type									
Public	15,468	90.7	91.1	16,382	96.6	96.5	15,489	91.4	91.2
Private	3,976	91.2	90.9	4,265	97.1	97.5	4,003	91.5	91.5
Catholic	2,015	91.1	91.7	2,186	99.4	99.5	2,036	92.2	92.6
Other private	1,961	91.4	90.0	2,079	95.0	95.5	1,967	90.9	90.3
Type of locale									
Large city	4,022	88.8	89.7	4,215	93.1	94.0	3,996	88.8	89.1
Mid-size city	4,314	95.7	95.0	4,496	99.4	99.1	4,238	93.2	93.4
Urban fringe of large city	5,372	89.0	87.7	5,836	95.6	95.3	5,522	90.6	90.2
Urban fringe of mid-size city	1,452	97.0	96.3	1,497	99.4	99.3	1,388	91.5	92.0
Large town	425	96.3	94.0	452	100.0	100.0	423	93.5	93.6
Small town	1,589	91.3	90.4	1,746	99.7	99.4	1,631	92.9	92.8
Rural	2,270	85.3	91.2	2,405	95.7	96.6	2,294	92.4	92.1
School size (Kindergartners)									
1 to 12	415	97.8	97.2	420	98.4	98.4	371	87.7	86.9
13 to 24	1,561	82.3	92.4	1,604	92.0	94.9	1,564	91.0	92.5
25 to 36	1,660	92.3	92.7	1,766	98.5	98.6	1,668	93.5	93.1
37 or more	15,808	91.3	90.6	16,857	96.9	96.6	15,889	91.3	91.1
Percent minority enrolled									
0 – 10%	5,190	91.3	90.9	5,601	98.4	98.1	5,196	91.3	91.0
11 – 49%	6,808	91.8	91.0	7,254	97.1	97.0	6,804	91.2	91.0
50 – 89%	3,856	95.2	95.2	3,948	97.8	97.5	3,744	92.7	92.4
90 – 100%	2,839	87.7	87.7	2,972	91.1	91.8	2,943	90.6	90.9
Unknown	751	73.4	85.5	872	95.4	99.3	805	90.3	91.7
Region									
Northeast	3,657	90.8	90.2	3,858	95.3	95.2	3,626	89.2	89.5
Midwest	4,997	94.0	94.4	5,203	98.4	98.3	4,858	91.4	91.8
South	6,209	91.2	92.4	6,515	96.7	96.9	6,258	93.6	93.1
West	4,581	86.9	86.6	5,071	95.8	95.8	4,750	89.6	89.8
High grade									
Trans K, K or trans 1	527	88.1	84.9	615	99.6	99.0	561	92.6	90.3
01-12	18,464	90.9	91.3	19,536	96.6	96.6	18,479	91.4	91.4
Ungraded	43	70.1	76.8	43	70.1	76.8	38	63.3	67.9
Unknown	410	92.9	90.5	453	100.0	100.0	414	91.9	91.4

[a]Based on frame data.

[b]Reading, math, or general knowledge assessment was scorable or child was language minority (not Spanish) and scored below cut score in OLDS or disabled (IEP).

[c]Family structure portion of parent interview was completed.

[d]Questionnaire was not blank.

Table 5-6. The ECLS-K, fall-kindergarten: number of completed child-level cases and child-level completion rates, by child characteristic

Child characteristic[a]	Child assessment			Parent interview		
		Completion rates			Completion rates	
	Completes[b]	Weighted	Unweighted	Completes[c]	Weighted	Unweighted
All children	19,173	89.9	89.8	18,097	85.3	84.7
Gender						
Male	9,751	89.5	89.4	9,171	84.6	84.1
Female	9,367	90.4	90.2	8,884	86.3	85.6
Unknown gender	55	79.1	76.4	42	52.7	58.3
Race-ethnicity						
White (not Hispanic)	10,975	90.2	90.2	10,692	87.8	87.9
Black (not Hispanic)	3,021	90.8	90.6	2,784	84.3	83.5
Hispanic	2,631	89.6	89.4	2,428	82.7	82.5
Asian or Pacific Islander	1,652	88.6	88.5	1,392	74.1	74.6
American Indian or Alaskan Native	339	93.4	93.1	292	80.0	80.2
Other	221	83.3	85.3	206	79.5	79.5
Unknown race-ethnicity	334	79.5	78.0	303	68.9	70.8
Year of birth						
1992	5,466	91.3	91.1	5,206	87.2	86.7
1993	13,410	89.8	89.8	12,611	85.0	84.5
1994	53	83.5	86.9	49	77.6	80.3
Other/unknown	244	71.7	67.8	231	65.1	64.2
Kindergarten type						
Full-day	10,640	90.4	90.2	10,061	85.8	85.3
Part-day	8,533	89.5	89.5	8,036	84.9	84.3
Kindergarten type unknown	0	0.0	0.0	0	0.0	0.0

Table 5-6. The ECLS-K, fall-kindergarten: number of completed child-level cases and child-level completion rates, by child characteristic (continued)

Child characteristic[a]	Teacher questionnaire A: Completes[d]	Teacher questionnaire A: Completion rates Weighted	Teacher questionnaire A: Completion rates Unweighted	Teacher questionnaire B: Completes[d]	Teacher questionnaire B: Completion rates Weighted	Teacher questionnaire B: Completion rates Unweighted	Teacher questionnaire C: Completes[d]	Teacher questionnaire C: Completion rates Weighted	Teacher questionnaire C: Completion rates Unweighted
All children	19,444	90.8	91.0	20,647	96.6	96.7	19,492	91.4	91.3
Gender									
Male	9,871	90.3	90.5	10,536	96.6	96.6	9,942	91.2	91.2
Female	9,524	91.6	91.7	10,062	96.9	96.9	9,504	91.8	91.5
Unknown gender	49	59.6	68.1	49	59.6	68.1	46	55.7	63.9
Race-ethnicity									
White (not Hispanic)	11,114	91.0	91.4	11,861	97.5	97.5	11,104	91.4	91.3
Black (not Hispanic)	3,089	92.8	92.7	3,224	96.9	96.7	3,103	93.5	93.1
Hispanic	2,579	88.1	87.6	2,776	94.1	94.3	2,673	91.1	90.8
Asian or Pacific Islander	1,731	92.9	92.7	1,816	97.4	97.3	1,721	92.4	92.2
American Indian or Alaskan Native	359	98.6	98.6	337	94.0	92.6	344	94.3	94.5
Other	237	89.3	91.5	251	97.1	96.9	225	85.3	86.9
Unknown race-ethnicity	335	78.8	78.3	382	87.3	89.3	322	73.8	75.2
Year of birth									
1992	5,551	91.9	92.5	5,845	97.1	97.4	5,552	92.7	92.5
1993	13,581	90.8	90.9	14,449	96.8	96.8	13,635	91.3	91.3
1994	54	90.7	88.5	57	92.9	93.4	52	82.3	85.2
Other/unknown	258	71.1	71.7	296	83.0	82.2	253	71.4	70.3
Kindergarten type									
Full-day	10,930	93.0	92.6	11,370	96.4	96.4	10,866	92.2	92.1
Part-day	8,514	88.4	89.3	9,277	97.1	97.3	8,626	90.6	90.5
Kindergarten type unknown	0	0.0	0.0	0	0.0	0.0	0	0.0	0.0

[a]Demographic data were captured from school records prior to student sampling.
[b]Reading, math, or general knowledge assessment was scorable or child was language minority (not Spanish) and scored below cut score in OLDS or disabled (IEP).
[c]Family structure portion of parent interview was completed.
[d]Questionnaire was not blank.

Table 5-7 presents the results of the fall-kindergarten teacher data collection, with teachers, not children, considered the unit of analysis. As all kindergarten teachers within a school were sampled, the weight used to calculate weighted completion rates is the school weight. Teachers were slightly more willing to complete questionnaire B that asked for personal background information (94.3 percent overall) than questionnaire A that focused on classroom characteristics (90.8 percent).

The school-level response rates presented in table 5-4 can be combined with the child-level and teacher-level completion rates from tables 5-5 and 5-7 to produce overall response rates at the child and teacher levels, by school characteristic. Table 5-8 displays the results at the child level. Children enrolled in Catholic schools are noteworthy for having high response rates across all survey instruments. Also notable are the relatively high response rates for children in schools with over 90 percent minority enrollment, schools in the Midwest, and schools whose highest grade is kindergarten. Table 5-9 presents overall response rates at the teacher level, by school characteristic. The same subgroups exhibit a better than average level of response.

5.6 Spring-Kindergarten Data Collection

5.6.1 Advance Contact with Respondents

In March 1999, letters were mailed to school administrators confirming the scheduled visits for the school that were set up in the fall. Letters were mailed to the school coordinators reminding them of the upcoming visits to the school. Letters were also mailed to parents reminding them of the spring data collection activities.

5.6.2 Preassessment Contact

Field supervisors conducted most preassessment activities by telephone except in the schools that had refused to participate in the fall-kindergarten collection but had agreed to participate in the spring collection. In these converted schools, field supervisors made personal visits to select the child sample and complete the other preassessment activities as described for fall-kindergarten data collection.

Table 5-7. The ECLS-K, fall-kindergarten: number of completed teacher questionnaires and teacher-level completion rates

School characteristic[a]	Teacher questionnaire A			Teacher questionnaire B		
		Completion rates			Completion rates	
	Completes[b]	Weighted	Unweighted	Completes[b]	Weighted	Unweighted
All schools	2,931	90.8	90.4	3,047	94.3	94.0
School type						
Public	2,577	90.3	90.2	2,684	94.2	93.9
Private	354	93.3	91.9	363	94.7	94.3
Catholic	146	92.8	94.2	152	97.2	98.1
Other private	208	93.5	90.4	211	93.5	91.7
Type of locale						
Large city	668	90.5	87.2	686	92.4	89.6
Mid-size city	599	95.5	94.9	614	97.7	97.3
Urban fringe of large city	854	88.3	88.1	903	93.1	93.2
Urban fringe of mid-size city	230	95.8	95.8	237	99.0	98.8
Large town	44	95.6	93.6	45	96.7	95.7
Small town	262	95.1	94.6	273	98.2	98.6
Rural	274	84.0	87.8	289	89.7	92.6
School size (Kindergartners)						
1 to 12	60	94.2	93.8	61	95.1	95.3
13 to 24	139	84.7	90.3	140	87.3	90.9
25 to 36	101	92.9	93.5	105	96.6	97.2
37 or more	2,631	91.2	90.2	2,741	95.0	94.0
Percent minority enrolled						
0 – 10%	650	91.5	91.9	679	96.0	96.0
11 – 49%	999	93.4	92.7	1,024	95.4	95.0
50 – 89%	662	90.9	92.2	684	92.9	95.3
90 – 100%	525	86.9	84.0	548	89.9	87.7
Unknown	95	82.5	83.3	112	97.6	98.2
Region						
Northeast	437	89.6	89.2	457	92.4	93.3
Midwest	573	95.3	93.9	589	95.9	96.6
South	1,129	91.2	92.4	1,156	95.1	94.6
West	792	87.0	86.1	845	92.9	91.8
High grade						
Trans K, K or trans 1	130	90.9	86.7	145	96.4	96.7
01-12	2,743	90.6	90.5	2,842	94.1	93.8
Ungraded	11	91.7	91.7	11	91.7	91.7
Unknown	47	96.0	95.9	49	100.0	100.0

[a]Based on frame data.
[b]Nonblank questionnaire.

Table 5-8. The ECLS-K, fall-kindergarten: number of completed child-level cases and overall child-level response rates, by school characteristic

School characteristic[a]	Child assessment			Parent interview		
		Overall response rates			Overall response rates	
	Completes[b]	Weighted	Unweighted	Completes[c]	Weighted	Unweighted
All children	19,173	62.4	61.8	18,097	59.2	58.3
School type						
Public	15,229	63.0	62.9	14,283	59.6	59.0
Private	3,944	59.4	59.1	3,814	57.7	57.2
Catholic	2,007	75.8	76.1	1,947	73.5	73.8
Other private	1,937	50.1	50.5	1,867	48.6	48.7
Type of locale						
Large city	3,954	68.9	68.7	3,611	63.1	62.8
Mid-size city	4,135	67.0	66.4	3,977	64.6	63.9
Urban fringe of large city	5,452	60.7	59.5	5,092	57.2	55.5
Urban fringe of mid-size city	1,359	55.6	54.8	1,302	53.2	52.5
Large town	416	52.8	55.8	415	52.8	55.6
Small town	1,601	59.4	57.8	1,532	56.7	55.3
Rural	2,256	58.3	58.0	2,168	56.2	55.7
School size (Kindergartners)						
1 to 12	366	64.1	59.6	358	63.8	58.3
13 to 24	1,550	57.2	58.7	1,478	54.9	56.0
25 to 36	1,640	65.6	68.1	1,577	62.9	65.5
37 or more	15,617	62.6	61.8	14,684	59.2	58.1
Percent minority enrolled						
0 – 10%	5,134	59.5	58.3	5,027	58.3	57.1
11 – 49%	6,704	61.6	60.5	6,405	58.7	57.8
50 – 89%	3,637	63.8	64.7	3,381	60.0	60.1
90 – 100%	2,905	70.6	70.4	2,514	61.9	61.0
Unknown	793	55.9	56.7	770	55.3	55.1
Region						
Northeast	3,605	57.9	58.5	3,373	54.3	54.7
Midwest	4,769	69.5	69.8	4,575	66.6	67.0
South	6,126	61.3	58.4	5,853	58.8	55.8
West	4,673	60.7	61.1	4,296	56.5	56.2
High grade						
Trans K, K or trans 1	552	70.9	69.7	526	67.3	66.4
01-12	18,174	62.5	62.0	17,135	59.2	58.4
Ungraded	40	74.6	71.4	35	56.7	62.5
Unknown	407	51.8	50.7	401	51.0	50.0

[a]Based on frame data.

[b]Reading, math, or general knowledge assessment was scorable or child was language minority (not Spanish) and scored below cut score in OLDS or disabled (IEP).

[c]Family structure portion of Parent Interview was completed.

[d]Questionnaire was not blank.

Table 5-9. The ECLS-K, fall-kindergarten: number of completed child-level cases and overall child-level response rates, by school characteristic

School characteristic[a]	Teacher questionnaire A Completes[d]	Teacher questionnaire A Overall response rates Weighted	Teacher questionnaire A Overall response rates Unweighted	Teacher questionnaire B Completes[d]	Teacher questionnaire B Overall response rates Weighted	Teacher questionnaire B Overall response rates Unweighted	Teacher questionnaire C Completes[d]	Teacher questionnaire C Overall response rates Weighted	Teacher questionnaire C Overall response rates Unweighted
All children	19,444	63.0	62.6	20,647	67.0	66.5	19,492	63.4	62.8
School type									
Public	15,468	63.6	63.9	16,382	67.7	67.6	15,489	64.1	63.9
Private	3,976	60.1	59.6	4,265	64.0	64.0	4,003	60.3	60.0
Catholic	2,015	75.6	76.4	2,186	82.5	82.9	2,036	76.5	77.1
Other private	1,961	51.4	51.1	2,079	53.4	54.2	1,967	51.1	51.3
Type of locale									
Large city	4,022	69.4	70.0	4,215	72.7	73.3	3,996	69.4	69.5
Mid-size city	4,314	70.4	69.3	4,496	73.2	72.2	4,238	68.6	68.1
Urban fringe of large city	5,372	60.3	58.6	5,836	64.8	63.7	5,522	61.4	60.3
Urban fringe of mid-size city	1,452	60.2	58.6	1,497	61.7	60.4	1,388	56.8	55.9
Large town	425	55.4	57.0	452	57.5	60.6	423	53.8	56.7
Small town	1,589	59.6	57.3	1,746	65.1	63.0	1,631	60.7	58.8
Rural	2,270	54.8	58.4	2,405	61.4	61.8	2,294	59.3	58.9
School size (Kindergartners)									
1 to 12	415	72.9	67.7	420	73.3	68.5	371	65.3	60.5
13 to 24	1,561	52.0	59.1	1,604	58.1	60.7	1,564	57.5	59.2
25 to 36	1,660	66.0	68.9	1,766	70.4	73.3	1,668	66.9	69.2
37 or more	15,808	63.6	62.5	16,857	67.5	66.7	15,889	63.6	62.9
Percent minority enrolled									
0 – 10%	5,190	60.3	58.9	5,601	65.0	63.6	5,196	60.3	59.0
11 – 49%	6,808	62.8	61.4	7,254	66.4	65.5	6,804	62.4	61.4
50 – 89%	3,856	67.5	68.5	3,948	69.3	70.2	3,744	65.7	66.5
90 – 100%	2,839	68.8	68.8	2,972	71.5	72.1	2,943	71.1	71.4
Unknown	751	46.2	53.7	872	60.0	62.4	805	56.8	57.6
Region									
Northeast	3,657	59.1	59.4	3,858	62.0	62.6	3,626	58.1	58.9
Midwest	4,997	72.8	73.2	5,203	76.2	76.2	4,858	70.7	71.1
South	6,209	61.1	59.2	6,515	64.8	62.1	6,258	62.7	59.7
West	4,581	59.7	59.9	5,071	65.8	66.3	4,750	61.6	62.1
High grade									
Trans K, K or trans 1	527	69.1	66.6	615	78.1	77.6	561	72.6	70.8
01-12	18,464	63.1	63.0	19,536	67.0	66.7	18,479	63.4	63.1
Ungraded	43	70.1	76.8	43	70.1	76.8	38	63.3	67.9
Unknown	410	53.4	51.1	453	57.5	56.5	414	52.8	51.6

[a]Based on frame data.
[b]Reading, math, or general knowledge assessment was scorable or child was language minority (not Spanish) and scored below cut score in OLDS or disabled (IEP).
[c]Family structure portion of Parent Interview was completed.
[d]Questionnaire was not blank.

5.6.3 Conducting Direct Child Assessments

The spring direct child assessment was identical to the one used in the fall except that the psychomotor assessment was dropped. The direct child assessments were conducted between March and June 1999. For children with a language other than English in the home, the child's score on the fall-kindergarten OLDS determined what assessment was administered in spring. Table 5-10 summarizes the OLDS routing patterns in spring-kindergarten.

Table 5-10. OLDS routing patterns in spring-kindergarten for fall-kindergarten respondents

Home Language	Fall-Kindergarten OLDS Score	OLDS Required in Spring-Kindergarten?	Spring assessment path
English	Not applicable	No	English
Spanish	At or above cut score	No	English
	Below cut score	Yes	
		At or above cut score	English
		Below cut score	Spanish
Other language	At or above cut score	No	English
	Below cut score	Yes	
		At or above cut score	English
		Below cut score	Height/Weight

Children who scored at or above the cut point on the OLDS in fall-kindergarten were automatically routed by CAPI to take the assessment in English; the OLDS was not administered again. Children who scored below the cut point in the OLDS in fall-kindergarten were administered the OLDS again in spring-kindergarten and routed according to the new spring-kindergarten OLDS score. Children taking the direct assessment for the first time in the spring, with a language other than English in the home, were routed according to their home language, just as in fall-kindergarten (see section 5.4.3). Table 5-11 presents the percentage of children who were routed into the various assessment alternatives in spring-kindergarten. Eight percent of the sampled children were screened using the OLDS in the spring-kindergarten. Overall almost one-half (42%) of the screened children were at or above the cut score on the OLDS and were administered the English direct child assessment. Of the children whose home language was Spanish, 36 percent were at or above the cut score, and of the children whose home language was a language other than English or Spanish, 55 percent were at or above the cut score.

Table 5-11. OLDS routing results in spring-kindergarten

	Total screened (%)	At or above cut score on OLDS (%)	Below cut score on OLDS (%)
Total screened	8	42	58
Spanish language	69	36	64
Other language	31	55	45

As in the fall, children identified as requiring an accommodation for the direct child assessment were provided the accommodation during the assessment. The patterns for accommodations were the same as in the fall-kindergarten data collection. Less than one percent of children were excluded from the spring direct child assessment due to physical or mental disability. Table 5-12 presents the number of children excluded from or requiring an accommodation to the direct child assessment in the spring of kindergarten.

Table 5-12. Number of children excluded from or accommodated in the spring-kindergarten assessment

	Number of children
Excluded for disability	70
Setting	35
Scheduling/timing	70
Health care aide	8
Assistive device	5

5.6.4 Conducting the Parent Interview

Parent interview procedures mirrored those of the fall-kindergarten. The parent interview was administered, primarily by telephone interview, from March through July 1999. The parent interview averaged 65 minutes. As in the fall, the parent interview was conducted in person if the respondent did not have a telephone. Three percent of all completed parent interviews were conducted in person. As in the fall, seven percent of the completed parent interviews were conducted in a language other than English with 61 percent of completed non-English interviews conducted in Spanish. The refusal conversion workshop developed for fall-kindergarten was implemented at the end of the spring field period to improve response rates. The special effort to build parent response rates was conducted between

June 15 and July 4, 1999, and yielded an additional two percentage points. Five percent of the cases were not completed because of locating problems.

5.6.5 Teacher and School Data Collection

Data were collected from school administrators, regular classroom teachers, and special education teachers between March and June 1999. Teachers were again asked to complete individual ratings for the sampled children in their classrooms, and they were reimbursed seven dollars for each child rating (teacher questionnaire C) they completed. In addition, school staff were asked to complete a student record abstract after the school year closed. The school staff were reimbursed five dollars for every student record abstract they completed. Field supervisors also completed a facilities checklist for each sampled school.

Field supervisors distributed the school and teacher questionnaires in a variety of ways, depending on the preference of the school staff. Questionnaires were distributed during the preassessment visit (if one was held), by mail, and during the assessment visits. During the field period, field supervisors followed up with school administrators and teachers by telephone and visits to the schools to ensure completed questionnaires were mailed to Westat, the company contracted to conduct the survey. In late June, a package was mailed to all schools with outstanding school administrator questionnaires, teacher questionnaires, or student record abstracts with a request to complete and return questionnaires. Field staff were prompted by telephone and in person for the return of school administrator questionnaires and student record abstracts through October 1999.

Data about annual salary and benefits of identified teachers and principals in the sampled schools were collected between April and July 1999 from public school district administrators, headmasters of private schools, and Catholic dioceses or schools. The majority of teachers were identified in early April 1999; a smaller number were identified and fielded in mid-May 1999. The package mailed to these respondents included a list by school of the names of the principals and teachers about whom the salary and benefit data were requested. Responses were returned by facsimile or in the included return mailer. Respondents were prompted by telephone for the return of the salary and benefit information through July 1999.

5.6.6 Conducting Data Collection on Children Who Withdrew from Their Original School

During the preassessment contacts, field supervisors asked school staff to identify children who had withdrawn from the school since the fall of kindergarten. School staff were asked whether they knew the name and address of the school the child transferred into as well as any new information about the child's household address. This information was entered into the FMS and processed at Westat for refielding. A total of 1,568 children (8.2 percent) were identified as movers from their original school between fall-kindergarten and spring-kindergarten. Different data collection strategies were followed for children who moved, depending on how they were classified. Data collection was attempted for children who moved as follows:

- Parent interviews were attempted regardless of children's mover status.
- Data collected for children moving into cooperating sampled schools included the child assessment in the school, school administrator questionnaire, regular or special education teacher questionnaires, and student record abstract forms.
- Data collected for children moving into nonsampled schools in cooperating districts included the child assessment in the school, school administrator questionnaire, regular or special education teacher questionnaires, salary and benefits questionnaire, and student record abstract forms.
- Data collected for children moving into sample schools that refused, schools in sampled districts that refused, or ineligible sampled schools included only the direct child assessment conducted in the home. No school or teacher data were collected.
- Data collected for children moving into schools in nonsampled districts or dioceses depended on whether the school was within or outside of the PSU.
  - For children in schools within the PSU, data collected included only the child assessment in the home.
  - For children in schools outside the PSU, no child or teacher data were collected.
- For children who were not in school in the spring (home school or at home children), data collected included the child assessment in the home, if the children were in the sample PSU, and no data if they were outside the sample PSU.

Almost one-half (46.7 percent, 733 children) of the children who moved were identified as moving into a school in the sampled ECLS-K PSUs, 18 percent moved into a school outside the PSU, 3 percent were identified as in home school, 27 percent of the movers were unlocatable because there was not enough information from the parent locating information to search the frame, and for 5 percent the schools could not be found on the frame. Of the 1,568 movers, data collection was attempted for 52 percent (816 children) of the located movers beginning in late April 1999. Almost 70 percent of the movers were fielded for in-home assessments and the remaining 30 percent of movers were fielded for in-school assessment. Sixty-two percent (510 children) of the fielded movers had assessments completed, 31 percent were nonresponse, and 1 percent were classified as out-of-scope. Table 5-13 presents the classification of the 1,568 children who moved and the number of completed assessments.

Table 5-13. Number of children who moved by category and assessment result

Child in	Number of children	Number of complete assessments
Cooperating sampled school	103	85
Refusal school	41	24
Nonsampled district within sample PSU	589	348
Nonsampled district outside sample PSU	302	0
Home school	45	32
Nonsampled private school	38	21
Unknown location	450	0
Total	1,568	510

5.6.7 Collection of Head Start Data

As mentioned in chapter 2, children's Head Start participation data were collected in the fall parent interview and in the spring-kindergarten student record abstract. The fall-kindergarten parent interview included a directory of Head Start centers names and addresses to assist the respondent in identifying the correct Head Start center name and address. The Head Start directory was sorted by state and alphabetically within state. In the child care section of the respondent was asked if the child attended Head Start in the year before kindergarten. If the answer was "Yes," then the directory was searched for the center name. If a match was found, the center name and address was verified with the respondent, selected from the directory and the identification number stored as the response to the question. If the center could not be found on the directory, then the field staff collected the name and address of the center, the name of the director and where the Head Start center was located, for example, in its own building, or in a school or church.

Head Start Data Collection-Spring-Kindergarten

In the student record abstract, collected in spring-kindergarten, one question asked whether or not the sampled child had attended a Head Start center before entering kindergarten. If the answer was "Yes," then the school staff person was asked to record the name, address, and telephone number of the Head Start center and the name of the Head Start director.

The process of verifying the Head Start centers and child participation began in the fall of 1999 and will continue through winter 2000.

5.7 Data Collection Quality Control

Data collection quality control efforts began with the development and testing of the CATI/CAPI applications and FMS. As these applications were programmed, extensive testing of the system was conducted. This testing included review by project design staff, statistical staff, and the programmers themselves. This testing by staff members, representing different aspects of the project, was designed to ensure that the systems were working properly from all perspectives. Quality control processes continued with the development of field procedures that maximized cooperation and thereby reduced the potential for nonresponse bias.

A live pilot test of the systems and procedures, including field supervisor and assessor training, was conducted in April and May 1998 with 12 elementary schools in the Washington, DC, metropolitan area. The purpose of the pilot test was to ensure that all the systems were working properly. Modifications to the data collection procedures, training programs, and systems were made to improve efficiency and reduce respondent burden. Modifications to the parent interview to address some issues raised by pilot test respondents were also made at this time.

Quality control activities continued during training and data collection. During the assessor training, field staff practiced conducting the parent interview in pairs and practiced the direct child assessment with kindergarten children brought to the training site for this purpose. 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. A sample of the assessor-completed OLDS score sheets were rescored in the home office for quality control purposes.

5.7.1 Child Assessment Observations

Field supervisors conducted on-site observations of the child assessments. During the fall-kindergarten collection, one observation was completed for each assessor within the first two weeks of the field period. In spring-kindergarten, two observations were completed for each assessor. The first observation was within two weeks after the assessments began, and the second observation was completed within three weeks of the first observation.

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

- Rapport building and working with the child – such as use of neutral praise and the assessor's response to various child behaviors.
- Cognitive assessment activities – such as reading questions verbatim, the use of acceptable probes, the use of appropriate hand motions, and the absence of coaching.
- Specific assessment activities – such as correctly coding answers to open-ended questions in the assessment, weighing and measuring the child correctly, 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. A training video reinforcing appropriate gesturing and reading items verbatim was created and used in the spring training.

5.7.2 Parent Validations

Parent validation forms were generated for approximately ten percent of the completed parent interviews. The first parent interview completed by an assessor was always validated. 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. This ensured that ten percent of each assessor's cases were selected for validation. The parent validation was approximately five minutes long.

Field supervisors used a standardized parent validation script when calling the parents. The script covered the following topics:

- Verification of the child's name, date of birth, and gender; and
- Between eight and ten questions from the current round interview were re-asked of the parent.

During the validation process, no evidence was found of parent interviews being falsified.

5.7.3 School Validations

To ensure that assessments proceeded smoothly, a validation call was completed with the school principal in approximately ten percent of each supervisor's assigned schools in both the fall- and spring-kindergarten collections.

Field managers conducted the school validations by telephone. The first school that each team completed 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 provided, including in-person observation of the supervisor's next school if necessary.

Field managers used a standardized script when calling the school principals. The script covered the following topics:

- How well the ECLS-K supervisor organized and executed the sampling tasks;
- An overall rating of how the assessments went;
- Feedback about the study from the children and kindergarten teachers;
- Suggestions for improving procedures and making it easier for a school to participate; and
- General comments and suggestions.

5.7.4 Quality Control of the OLDS Scoring

The OLDS used to screen children for English language proficiency included the "Let's Tell Stories" subtest. This subtest involved reading the child a short story and having the child repeat it back to the assessor. The child's responses were recorded verbatim and scored by the assessor. Responses to this subtest are unique to each child, and it was important for interviewers' and coders' scoring of the child's responses to match the preLAS®2000 standards.

ECLS-K assessors were trained to conduct the OLDS using audiotapes of the stories and children's responses to the stories. Assessors listened to the audiotaped stories and to the child's responses and recorded the child's responses verbatim. Then the assessor scored the story using the preLAS®2000 rules. Reasons for scoring each story a particular way were discussed in detail. Differences between the assessor's scores and the correct scores were discussed during training, so assessors could understand the difference between the scores. Several stories in each scoring category were provided for practice to fine tune the assessor's scoring. Then the scoring ability of each assessor was tested. Only assessors who scored a 90 percent accuracy in scoring the training stories as matched against the preLAS®2000 samples were allowed to conduct the OLDS.

A ten percent sample of each assessor's OLDS stories were recoded in the home office. The coders received the same training as the assessors. Coders then scored the stories independently. If the home office coders' scores differed from the assessor, their scores were verified by their supervisor. All cases were adjudicated by lead trainers for the OLDS. Approximately 66 percent of the stories had complete score agreement between the assessor, coder, and lead trainer. The additional 33 percent of the stories had score agreement by two of the three scorers.

5.7.5 Assessor Effects

Individual Test Administrator effects and Design Effects

A multi-level analysis[3] was carried out to estimate components of variance in fall- and spring-kindergarten cognitive scores associated with the: (1) student level, (2) school level, (3) team

[3] Bryk, A. & Raudenbush, S.W. (1992). *Hierarchical Linear Models: Applications and data analysis methods.* New York, Sage Publications. Snijders, T. & Bosker, R. (1999). *Multilevel Analysis – An introduction to basic and advanced multilevel modeling.* London, Sage Publications.

leader, and (4) individual test administrator. This secondary analysis was motivated by Westat's earlier finding of larger than expected design effects. In addition, the impact on the above sources of variance of the SES indicator (parent's education) was also estimated. It was expected that much of the clustering of students within neighborhood schools (hence higher design effects) could be explained by SES.

In addition to the potential clustering effects related to shared parent SES within schools, there was a concern that the individual mode of administration might inject additional and unwanted variance to both the individual and the between school components of variance in the cognitive scores. Since it is more difficult to standardize test administrations when tests are individually administered, this source of variance could contribute to the high design effects if the individual assessors differed systematically in their modes of administration.

It was found that the component of variance associated with the individual test administration effect was negligible in all three cognitive areas and thus had little or no impact on the design effects. Much of the design effects with respect to cognitive scores could be explained by parents' SES.

5.8 Spring-Kindergarten Response Rates

Table 5-14 presents school-level response rates for spring-kindergarten data collection. Response rates increased for both public and private schools from fall to spring because of a concerted effort at refusal conversion. This was most dramatic for Catholic schools where the rate increased from 83.0 percent to 94.1 percent, but even non-Catholic private schools increased by over four percentage points. Although not directly evident from the table, all but three schools cooperating in the fall data collection continued to do so in the spring.

Table 5-15 presents child-level completion rates for the spring data collection, broken out by school characteristics. Because of the refusal converted schools, the absolute number of completed cases increased from fall- to spring-kindergarten. However, the completion rates fell by a point or two for the child assessment and parent interviews between the two rounds. Overall the child-level rates for teacher questionnaires A, B, and C dropped by four to six points between fall-kindergarten and spring, largely

Table 5-14. The ECLS-K, spring-kindergarten: number of cooperating schools and before-substitution school-level response rates

School characteristic[a]	Number cooperating[b]		Before-substitution response rates	
	Before substitution	After substitution	School weight × KG enrollment	Unweighted
All schools	940	1014	74.0	73.7
School type				
Public	678	752	74.2	74.2
Private	262	262	72.8	72.4
Catholic	113	113	94.1	94.2
Other private	149	149	60.6	61.6
Type of locale				
Large city	195	202	79.0	79.6
Mid-size city	189	206	76.8	76.5
Urban fringe of large city	271	287	71.6	70.9
Urban fringe of mid-size city	63	70	66.9	65.6
Large town	31	31	96.3	93.9
Small town	77	88	71.6	68.8
Rural	114	130	69.4	70.8
School size (Kindergartners)				
1 to 12	57	57	79.0	73.1
13 to 24	108	110	69.7	72.0
25 to 36	82	86	78.0	81.2
37 or more	693	761	73.9	73.2
All schools	940	1014	74.0	73.7
Percent minority enrolled				
0 – 10 %	249	273	73.0	71.8
11 – 49 %	302	341	71.5	71.6
50 – 89%	191	199	75.8	76.7
90 – 100%	140	143	81.3	81.4
Unknown	58	58	67.7	67.4
Region				
Northeast	168	189	68.4	69.1
Midwest	246	253	82.3	82.6
South	301	327	74.3	72.2
West	225	245	69.6	70.8
High grade				
Trans K, K or trans 1	43	43	83.8	84.3
01-12	855	929	73.9	73.8
Ungraded	4	4	100.0	100.0
Unknown	38	38	62.0	61.3

[a]Based on frame data.

[b]One or more children or parents participated, or school is one of two schools in the Archdiocese of New Orleans where student sampling was performed too late in fall-kindergarten to field child or parent data collection.

Table 5-15. The ECLS-K, spring-kindergarten: number of completed child-level cases and child-level completion rates, by school characteristic

School characteristic[a]	Child assessment			Parent interview			School administrator questionnaire		
		Completion rates			Completion rates			Completion rates	
	Completes[b]	Weighted	Unweighted	Completes[c]	Weighted	Unweighted	Completes[d]	Weighted	Unweighted
All children	19,967	88.0	88.3	18,950	83.9	83.8	19,282	85.9	85.4
School type									
Public	15,581	87.7	87.8	14,695	83.3	82.8	14,930	85.3	84.3
Private	4,386	90.1	90.0	4,255	87.9	87.3	4,352	89.9	89.5
Catholic	2,296	91.7	91.5	2,206	88.4	87.9	2,334	94.2	93.1
Other private	2,090	88.5	88.4	2,049	87.5	86.7	2,018	85.7	85.6
Type of locale									
Large city	3,933	86.6	86.9	3,599	79.4	79.5	3,289	71.4	72.9
Mid-size city	4,219	89.0	89.7	4,050	85.5	86.1	4,284	91.4	91.3
Urban fringe of large city	5,674	87.6	87.9	5,360	83.6	83.0	5,305	82.1	82.2
Urban fringe of mid-size city	1,404	86.7	87.3	1,349	83.2	83.9	1,514	95.3	94.3
Large town	636	89.2	89.1	624	87.1	87.4	666	92.6	93.3
Small town	1,696	89.5	89.4	1,641	87.2	86.5	1,683	91.7	88.9
Rural	2,405	88.6	88.4	2,327	85.8	85.6	2,541	94.5	93.7
School size (Kindergartners)									
1 to 12	383	83.9	84.2	395	87.7	86.8	405	90.9	89.4
13 to 24	1,725	89.8	91.1	1,642	85.3	86.7	1,701	93.2	90.0
25 to 36	1,780	91.0	91.1	1,712	87.6	87.6	1,852	96.6	95.0
37 or more	16,079	87.7	87.8	15,201	83.4	83.0	15,324	84.3	83.8
Percent minority enrolled									
0 – 10%	5,538	87.9	88.4	5,427	86.4	86.6	5,853	92.8	93.6
11 – 49%	6,878	87.9	87.9	6,622	84.6	84.6	7,223	93.1	92.5
50 – 89%	3,746	88.0	87.9	3,478	82.6	81.6	3,375	79.3	79.3
90 – 100%	2,952	88.7	89.0	2,592	78.6	78.1	2,054	61.9	62.1
Unknown	853	87.7	88.9	831	85.6	86.7	777	90.9	81.1
Region									
Northeast	3,746	88.2	88.5	3,513	82.5	83.0	3,417	78.4	80.9
Midwest	4,951	87.6	88.2	4,761	84.3	84.8	5,069	90.7	90.5
South	6,685	89.4	89.6	6,358	85.4	85.2	6,476	88.0	86.9
West	4,585	85.9	86.3	4,318	82.1	81.2	4,320	83.1	81.5
High grade									
Trans K, K or trans 1	597	90.3	89.4	573	87.8	85.8	582	91.2	87.3
01-12	18,880	87.9	88.3	17,895	83.8	83.7	18,227	85.7	85.4
Ungraded	43	80.5	76.8	38	62.8	67.9	56	100.0	100.0
Unknown	447	88.0	87.5	444	87.5	86.9	417	83.5	81.8
All children	19,637	87.0	87.0	20,523	90.5	90.9	19,382	86.0	85.8

Table 5-15. The ECLS-K, spring-kindergarten: number of completed child-level cases and child-level completion rates, by school characteristic (continued)

School characteristic[a]	Teacher questionnaire A			Teacher questionnaire B			Teacher questionnaire C		
		Completion rates			Completion rates			Completion rates	
	Completes[d]	Weighted	Unweighted	Completes[d]	Weighted	Unweighted	Completes[d]	Weighted	Unweighted
School type									
Public	15,389	86.9	86.9	15,880	89.7	89.6	15,233	85.9	86.0
Private	4,248	87.9	87.3	4,643	95.2	95.5	4,149	86.0	85.3
Catholic	2,296	91.7	91.5	2,206	88.4	87.9	2,334	94.2	93.1
Other private	2,069	88.3	87.7	2,197	92.7	93.2	1,955	83.7	82.9
Type of locale									
Large city	3,526	77.8	78.1	3,872	84.5	85.8	3,741	82.8	82.9
Mid-size city	4,214	89.6	89.8	4,359	92.2	92.9	4,093	87.2	87.2
Urban fringe of large city	5,718	89.2	88.6	5,855	90.7	90.8	5,462	84.8	84.7
Urban fringe of mid-size city	1,418	87.7	88.3	1,482	91.7	92.3	1,380	85.3	85.9
Large town	632	89.0	88.5	655	91.5	91.7	632	88.6	88.5
Small town	1,673	88.7	88.4	1,795	95.0	94.8	1,670	88.8	88.2
Rural	2,456	89.1	90.6	2,505	91.4	92.4	2,404	88.7	88.6
School size (Kindergartners)									
1 to 12	394	88.4	87.0	423	93.2	93.4	367	82.0	81.0
13 to 24	1,671	85.6	88.4	1,756	89.9	92.9	1,654	85.3	87.5
25 to 36	1,771	90.6	90.9	1,866	94.7	95.7	1,718	89.1	88.1
37 or more	15,801	86.9	86.4	16,478	90.1	90.1	15,643	85.8	85.5
Percent minority enrolled									
0 – 10%	5,701	90.9	91.2	5,901	94.1	94.4	5,486	87.5	87.7
11 – 49%	6,735	86.5	86.3	7,158	91.2	91.7	6,666	85.4	85.4
50 – 89%	3,682	86.7	86.5	3,810	89.5	89.5	3,648	86.8	85.7
90 – 100%	2,716	82.4	82.1	2,742	82.0	82.9	2,781	84.5	84.1
Unknown	803	83.8	83.8	912	92.6	95.2	801	82.6	83.6
Region									
Northeast	3,836	90.6	90.8	3,871	91.6	91.6	3,601	85.6	85.2
Midwest	4,773	85.3	85.2	5,176	91.9	92.4	4,870	86.8	86.9
South	6,699	89.2	89.9	6,781	90.1	91.0	6,545	87.6	87.8
West	4,329	82.3	81.6	4,695	88.6	88.6	4,366	82.5	82.3
High grade									
Trans K, K or trans 1	529	83.4	79.3	631	95.3	94.6	575	87.8	86.2
01-12	18,627	87.3	87.2	19,358	90.2	90.7	18,342	86.1	85.9
Ungraded	29	41.8	51.8	42	68.5	75.0	38	63.2	67.9
Unknown	452	87.5	88.6	492	96.6	96.5	427	82.2	83.7

[a]Based on frame data.
[b]Reading, math, or general knowledge assessment was scorable or child was language minority (not Spanish) and score below cut score in OLDS or disabled (IEP).
[c]Family structure portion of parent interview was completed.
[d]Questionnaire was not blank.

because of the data collection protocol for students who transferred to another school (for details see section 5.6.6). For teacher A and B, this disproportionately affected children in large cities, who showed a loss of 9 to 11 percentage points. Catholic schools too showed an 11-point drop in teacher B completion, though the spring-kindergarten rate was still in excess of 88 percent. The weighted completion rate for the school administrator questionnaire was 85.9 percent overall with the lowest rate observed for large cities (71.4 percent).

Table 5-16 also presents child-level response rates for the spring data collection, this time broken out by child characteristics. The same gains and losses of sample size and completion rates hold as in table 5-15 and for the same reasons. However, for the teacher A and B questionnaires, the decrease in the completion rate between fall and spring was higher for black students, i.e., in the range of 8 to 9 points. The lowest rate of completion for the school administrator questionnaire was for black and Hispanic students (76.4 and 75.1 percent, respectively), largely because of their clustering in large cities.

Table 5-17 presents teacher-level completion rates for the spring-kindergarten data collection broken out by school characteristics. Relative to the fall data collection experience, the overall completion rate for the teacher A questionnaire increased by nearly three percentage points to 93.4 percent and the teacher B rate held steady (94.1 percent). There was an 11-point gain in the teacher A rate for rural schools but a nearly four percentage point loss in the Midwest, which, however, still remained at nearly 92 percent.

Table 5-18 displays school-level completion rates for the school administrator questionnaire broken out by school characteristics. The overall weighted rate is 87.2 percent. The private school completion rate is about five points higher than that for public schools (90.7 vs. 85.5 percent). Rates tended to be lower for schools in large cities (77.6 percent complete), those having 90 percent or more minority students enrolled (65.0 percent), and those located in the West (80.1 percent).

The spring-kindergarten school-level response rates presented in table 5-14 can be combined with the child-level and teacher-level completion rates from tables 5-15 and 5-17 to produce overall response rates at the child and teacher levels, by school characteristic. Table 5-19 displays the results at the child-level. For the most part, the response patterns notable in the fall-kindergarten assessment persisted into the spring. Response rates for children enrolled in Catholic schools were higher than average across all survey instruments, as were those for children in schools in the Midwest, and schools

Table 5-16. The ECLS-K, spring-kindergarten: number of completed child-level cases and child-level completion rates, by child characteristic

Child characteristic[a]	Child assessment			Parent interview			School administrator questionnaire		
		Completion rates			Completion rates			Completion rates	
	Completes[b]	Weighted	Unweighted	Completes[c]	Weighted	Unweighted	Completes[d]	Weighted	Unweighted
All children	19,967	88.0	88.3	18,950	83.9	83.8	19,282	85.8	85.4
Gender									
Male	10,191	88.1	88.3	9,679	84.0	83.8	9,864	86.1	85.6
Female	9,720	88.2	88.5	9,231	84.3	84.1	9,370	85.7	85.5
Unknown gender	56	59.1	56.6	40	35.2	40.4	48	51.7	48.5
Race-ethnicity									
White (not Hispanic)	11,602	88.2	88.5	11,370	86.4	86.7	12,011	92.0	91.8
Black (not Hispanic)	3,130	88.7	88.6	2,847	80.8	80.6	2,601	76.4	73.8
Hispanic	2,668	88.4	88.5	2,494	83.0	82.7	2,223	75.1	73.9
Asian or Pacific Islander	1,696	89.6	89.9	1,455	77.0	77.1	1,614	86.8	85.7
American Indian or Alaskan Native	327	88.1	88.4	288	77.7	77.8	332	88.6	90.2
Other	219	83.4	85.2	201	78.0	78.2	227	90.8	88.3
Unknown race-ethnicity	325	72.8	71.4	295	64.0	64.8	274	52.8	60.2
Year of birth									
1992	5,720	88.8	89.1	5,465	85.1	85.2	5,689	88.9	88.7
1993	13,936	88.1	88.4	13,199	84.0	83.7	13,244	84.5	84.2
1994	54	85.2	85.7	52	83.9	82.5	46	82.6	75.4
Other/unknown	257	70.1	68.7	234	62.2	62.6	303	84.2	81.2
Kindergarten type									
Full day	11,036	93.0	92.9	10,246	86.4	86.3	10,149	85.4	85.5
Part day	8,543	91.1	91.1	8,016	86.0	85.5	8,017	86.8	86.1
Kindergarten type unknown	388	27.0	28.3	688	49.9	50.1	1,062	81.5	79.8

Table 5-16. The ECLS-K, spring-kindergarten: number of completed child-level cases and child-level completion rates, by child characteristic (continued)

Child characteristic[a]	Teacher questionnaire A			Teacher questionnaire B			Teacher questionnaire C		
		Completion rates			Completion rates			Completion rates	
	Completes[d]	Weighted	Unweighted	Completes[d]	Weighted	Unweighted	Completes[d]	Weighted	Unweighted
All children	19,637	87.0	87.0	20,523	90.4	90.9	19,382	85.9	85.8
Gender									
Male	10,010	87.2	86.9	10,455	90.5	90.7	9,875	85.7	85.7
Female	9,584	87.3	87.4	10,011	90.8	91.3	9,477	86.6	86.5
Unknown gender	43	46.0	43.4	57	49.5	57.6	30	35.6	30.3
Race-ethnicity									
White (not Hispanic)	11,623	89.0	88.8	12,187	92.7	93.1	11,323	86.5	86.5
Black (not Hispanic)	2,983	85.1	84.6	3,086	87.7	87.6	3,057	87.2	86.7
Hispanic	2,502	83.4	83.2	2,593	86.2	86.2	2,544	85.2	84.6
Asian or Pacific Islander	1,647	88.3	87.4	1,745	92.9	92.6	1,632	87.2	86.6
American Indian or Alaskan Native	331	89.3	89.9	313	85.8	85.1	332	90.1	90.2
Other	215	86.1	83.7	237	90.6	92.2	208	81.1	80.9
Unknown race-ethnicity	336	70.5	73.8	362	77.4	79.6	286	62.1	62.9
Year of birth									
1992	5,581	86.9	87.0	5,885	91.1	91.8	5,599	87.1	87.3
1993	13,750	87.6	87.4	14,303	90.6	90.9	13,498	86.0	85.8
1994	51	87.0	83.6	56	90.8	91.8	47	80.0	77.0
Other/unknown	255	65.7	68.4	279	71.9	74.8	238	63.0	63.8
Kindergarten type									
Full-day	11,106	94.4	93.5	11,437	96.2	96.3	10,800	91.6	91.0
Part-day	8,531	91.0	91.0	9,062	96.5	96.6	8,412	89.6	89.7
Kindergarten type unknown	0	0.0	0.0	24	1.9	1.8	170	14.2	12.8

[a]Demographic data were captured from school records prior to student sampling.

[b]Reading, math, or general knowledge assessment was scorable or child was language minority (not Spanish) and scored below cut score in OLDS or disabled (IEP).

[c]Family structure portion of parent interview was completed.

[d]Questionnaire was not blank.

Table 5-17. The ECLS-K, spring-kindergarten: number of completed teacher questionnaires and teacher-level completion rates

School characteristic[a]	Teacher questionnaire A			Teacher questionnaire B		
		Completion rates			Completion rates	
	Completes[b]	Weighted	Unweighted	Completes[c]	Weighted	Unweighted
All schools	3,215	93.4	93.0	3,243	94.1	93.8
School type						
Public	2,819	93.7	93.2	2,830	93.9	93.6
Private	396	92.4	91.2	413	95.3	95.2
Catholic	168	93.5	93.9	176	98.1	98.3
Other private	228	91.8	89.4	237	93.8	92.9
Type of locale						
Large city	688	88.8	88.1	703	92.4	90.0
Mid-size city	640	96.6	96.2	646	97.6	97.1
Urban fringe of large city	955	93.1	93.1	953	93.0	92.9
Urban fringe of mid-size city	244	95.0	95.3	253	98.7	98.8
Large town	85	89.4	89.5	86	90.6	90.5
Small town	288	94.8	95.7	291	96.4	96.7
Rural	315	94.9	94.3	311	90.8	93.1
School size (Kindergartners)						
1 to 12	64	94.2	92.8	67	96.7	97.1
13 to 24	157	88.4	90.2	158	87.7	90.8
25 to 36	114	96.7	95.8	116	96.9	97.5
37 or more	2,880	93.8	93.0	2,902	94.6	93.7
Percent minority enrolled						
0 – 10%	736	95.5	94.8	740	95.5	95.4
11 – 49%	1,070	93.3	93.6	1,083	95.0	94.8
50 – 89%	704	92.4	92.5	726	93.7	95.4
90 – 100%	596	91.9	91.0	573	89.5	87.5
Unknown	109	92.8	88.6	121	97.1	98.4
Region						
Northeast	486	93.5	92.7	490	93.1	93.5
Midwest	574	91.6	89.0	613	94.6	95.0
South	1,288	96.0	96.2	1,267	95.0	94.6
West	867	90.7	91.3	873	92.9	91.9
High grade						
Trans K, K or trans 1	134	83.4	79.8	160	94.4	95.2
01-12	3,019	94.1	93.7	3,017	94.0	93.6
Ungraded	10	77.2	76.9	12	92.4	92.3
Unknown	52	94.3	94.5	54	98.5	98.2

[a]Based on frame data.

[b]Round 2 questionnaire completed by teacher employed in a base-year probability sampled school.

[c]Fall-kindergarten or spring-kindergarten questionnaire completed by teacher employed in a base-year probability sampled school.

Table 5-18. The ECLS-K, spring-kindergarten: number of completed school administrator questionnaire and school-level completion rates

School characteristic[a]	Number cooperating[b]	Completion rates	
		Weighted[c]	Unweighted
All schools	866	87.2	85.4
School type			
Public	631	85.5	83.9
Private	235	90.7	89.7
Catholic	105	93.9	92.9
Other private	130	88.9	87.2
Type of locale			
Large city	148	77.6	73.3
Mid-size city	189	92.8	91.7
Urban fringe of large city	237	84.1	82.6
Urban fringe of mid-size city	66	95.5	94.3
Large town	29	94.7	93.5
Small town	78	88.6	88.6
Rural	119	89.5	91.5
School size (Kindergartners)			
1 to 12	50	85.0	87.7
13 to 24	98	90.2	89.1
25 to 36	82	96.0	95.3
37 or more	636	84.8	83.6
Percent minority enrolled			
0 – 10%	256	94.7	93.8
11 – 49%	315	92.2	92.4
50 – 89%	158	80.8	79.4
90 – 100%	89	65.0	62.2
Unknown	48	86.3	82.8
Region			
Northeast	154	84.7	81.5
Midwest	228	92.1	90.1
South	286	89.6	87.5
West	198	80.1	80.8
High grade			
Trans K, K or trans 1	37	87.1	86.0
01-12	793	87.2	85.4
Ungraded	4	100.0	100.0
Unknown	32	85.2	84.2

[a]Based on frame data.

[b]School is part of the original base-year probability sample and questionnaire was not blank.

[c]Weighted by school base weight.

Table 5-19. The ECLS-K, spring-kindergarten: child-level sample sizes and overall response rates, by school characteristic

School characteristics[a]	Child assessment			Parent interview			School administrator questionnaire		
		Overall response rates			Overall response rates			Overall response rates	
	Completes[b]	Weighted	Unweighted	Completes[c]	Weighted	Unweighted	Completes[d]	Weighted	Unweighted
All children	19,967	65.1	65.1	18,950	62.1	61.8	19,282	63.6	62.9
School type									
Public	15,581	65.1	65.1	14,695	61.8	61.4	14,930	63.3	62.6
Private	4,386	65.6	65.2	4,255	64.0	63.2	4,352	65.4	64.8
Catholic	2,296	86.3	86.2	2,206	83.2	82.8	2,334	88.6	87.7
Other private	2,090	53.6	54.5	2,049	53.0	53.4	2,018	51.9	52.7
Type of locale									
Large city	3,933	68.4	69.2	3,599	62.7	63.3	3,289	56.4	58.0
Mid-size city	4,219	68.4	68.6	4,050	65.7	65.9	4,284	70.2	69.8
Urban fringe of large city	5,674	62.7	62.3	5,360	59.9	58.8	5,305	58.8	58.3
Urban fringe of mid-size city	1,404	58.0	57.3	1,349	55.7	55.0	1,514	63.8	61.9
Large town	636	85.9	83.7	624	83.9	82.1	666	89.2	87.6
Small town	1,696	64.1	61.5	1,641	62.4	59.5	1,683	65.7	61.2
Rural	2,405	61.5	62.6	2,327	59.5	60.6	2,541	65.6	66.3
School size (Kindergartners)									
1 to 12	383	66.3	61.6	395	69.3	63.5	405	71.8	65.4
13 to 24	1,725	62.6	65.6	1,642	59.5	62.4	1,701	65.0	64.8
25 to 36	1,780	71.0	74.0	1,712	68.3	71.1	1,852	75.3	77.1
37 or more	16,079	64.8	64.3	15,201	61.6	60.8	15,324	62.3	61.3
Percent nonwhite enrolled									
0 – 10%	5,538	64.2	63.5	5,427	63.1	62.2	5,853	67.7	67.2
11 – 49%	6,878	62.8	62.9	6,622	60.5	60.6	7,223	66.6	66.2
50 – 89%	3,746	66.7	67.4	3,478	62.6	62.6	3,375	60.1	60.8
90 – 100%	2,952	72.1	72.4	2,592	63.9	63.6	2,054	50.3	50.5
Unknown	853	59.4	59.9	831	58.0	58.4	777	61.5	54.7
Region									
Northeast	3,746	60.3	61.2	3,513	56.4	57.4	3,417	53.6	55.9
Midwest	4,951	72.1	72.9	4,761	69.4	70.0	5,069	74.6	74.8
South	6,685	66.4	64.7	6,358	63.5	61.5	6,476	65.4	62.7
West	4,585	59.8	61.1	4,318	57.1	57.5	4,320	57.8	57.7
High grade									
Trans K, K or trans 1	597	75.7	75.4	573	73.6	72.3	582	76.4	73.6
01-12	18,880	65.0	65.2	17,895	61.9	61.8	18,227	63.3	63.0
Ungraded	43	80.5	76.8	38	62.8	67.9	56	100.0	100.0
Unknown	447	54.6	53.6	444	54.3	53.3	417	51.8	50.1

Table 5.19. The ECLS-K, spring-kindergarten: child-level sample sizes and overall response rates, by school characteristic (continued)

School characteristics[a]	Teacher questionnaire A			Teacher questionnaire B			Teacher questionnaire C		
	Completes[d]	Overall response rates		Completes[d]	Overall response rates		Completes[d]	Overall response rates	
		Weighted	Unweighted		Weighted	Unweighted		Weighted	Unweighted
All children	19,637	64.4	64.1	20,523	67.0	67.0	19,382	63.6	63.2
School type									
Public	15,389	64.5	64.5	15,880	66.6	66.5	15,233	63.7	63.8
Private	4,248	64.0	63.2	4,643	69.3	69.1	4,149	62.6	61.8
Catholic	2,179	82.3	82.0	2,446	92.1	91.9	2,194	83.3	82.4
Other private	2,069	53.5	54.0	2,197	56.2	57.4	1,955	50.7	51.1
Type of locale									
Large city	3,526	61.5	62.2	3,872	66.8	68.3	3,741	65.4	66.0
Mid-size city	4,214	68.8	68.7	4,359	70.8	71.1	4,093	67.0	66.7
Urban fringe of large city	5,718	63.9	62.8	5,855	64.9	64.4	5,462	60.7	60.1
Urban fringe of mid-size city	1,418	58.7	57.9	1,482	61.3	60.5	1,380	57.1	56.4
Large town	632	85.7	83.1	655	88.1	86.1	632	85.3	83.1
Small town	1,673	63.5	60.8	1,795	68.0	65.2	1,670	63.6	60.7
Rural	2,456	61.8	64.1	2,505	63.4	65.4	2,404	61.6	62.7
School size (Kindergartners)									
1 to 12	394	69.8	63.6	423	73.6	68.3	367	64.8	59.2
13 to 24	1,671	59.7	63.6	1,756	62.7	66.9	1,654	59.5	63.0
25 to 36	1,771	70.7	73.8	1,866	73.9	77.7	1,718	69.5	71.5
37 or more	15,801	64.2	63.2	16,478	66.6	66.0	15,643	63.4	62.6
Percent nonwhite enrolled									
0 – 10%	5,701	66.4	65.5	5,901	68.7	67.8	5,486	63.9	63.0
11 – 49%	6,735	61.8	61.8	7,158	65.2	65.7	6,666	61.1	61.1
50 – 89%	3,682	65.7	66.3	3,810	67.8	68.6	3,648	65.8	65.7
90 – 100%	2,716	67.0	66.8	2,742	66.7	67.5	2,781	68.7	68.5
Unknown	803	56.7	56.5	912	62.7	64.2	801	55.9	56.3
Region									
Northeast	3,836	62.0	62.7	3,871	62.7	63.3	3,601	58.6	58.9
Midwest	4,773	70.2	70.4	5,176	75.6	76.3	4,870	71.4	71.8
South	6,699	66.3	64.9	6,781	66.9	65.7	6,545	65.1	63.4
West	4,329	57.3	57.8	4,695	61.7	62.7	4,366	57.4	58.3
High grade									
Trans K, K or trans 1	529	69.9	66.8	631	79.9	79.7	575	73.6	72.7
01-12	18,627	64.5	64.4	19,358	66.7	66.9	18,342	63.6	63.4
Ungraded	29	41.8	51.8	42	68.5	75.0	38	63.2	67.9
Unknown	452	54.3	54.3	492	59.9	59.2	427	51.0	51.3

[a]Based on frame data.
[b]Reading, math, or general knowledge assessment was scorable or child was language minority (not Spanish) and scored below the cut score in the OLDS or disabled (IEP).
[c]Family structure portion of Parent Interview was completed.
[d]Questionnaire was not blank.

whose highest grade is kindergarten. For schools with over 90 percent minority enrollment, response rates were higher than average for all but the school administrator questionnaire, which was a full 13 points below the average rate. Table 5-20 presents overall response rates at the teacher-level, by school characteristic. The same subgroups exhibit a better than average level of response.

In addition to the child assessment, parent interview, teacher questionnaires, and school administrator questionnaires whose completion rates have been summarized in the preceding tables, various other types of data were collected during spring of kindergarten as well. Table 5-21 presents counts of completes and weighted and unweighted completion rates at the overall student level for these other data collection efforts. The facilities checklist has a 96.6 percent weighted completion rate at the student level. (At the school level, it is 95.7 percent.) Also very high are the special education teacher questionnaire A and the school principal's salary and benefits information, at 94.1 and 91.0 percent completion respectively. The salary and benefits data for special education teachers is largely unavailable (31.4 percent completion) as is the Adaptive Behavior Scale (45.6 percent), which was collected only for students who were not assessed in spring of kindergarten because of physical or mental disability. The Adaptive Behavior Scale was collected for a very small number of children, less than 100, and seemed to be burdensome for the special education teachers to complete. The majority of teachers were identified for teacher and salary benefits data collection in early April 1999; a smaller number were identified and fielded in mid-May 1999. It appears that the timing of this data collection effort may not have included all of the special education teachers. In addition, these teachers often had other sources of payment about which the school district could not report; for example, the county had a special education cooperative that paid the teacher salary or the salary was paid by other special services contracts.

5.9 Base Year Response Rates

Thus far cooperation rates have been presented for the two rounds of data collection separately. It is also reasonable to consider such rates from the perspective of the base year as a whole. Table 5-22 presents base year rates for schools, children, parents, and teachers. Within this table a child, parent, or teacher is considered a respondent if he or she cooperated in either the fall- or spring-kindergarten data collection. A school is counted as responding if one or more parents or children from that school cooperated in either round. Rates are given both weighted and unweighted. Weighted rates were calculated using the school base weight for schools and teachers and the child base weight for

Table 5-20. The ECLS-K, spring-kindergarten: number of completed teacher questionnaires and teacher-level response rates

School characteristic[a]	Teacher questionnaire A			Teacher questionnaire B		
		Response rates			Response rates	
	Completes[b]	Weighted	Unweighted	Completes[b]	Weighted	Unweighted
All schools	3,215	69.1	68.5	3,243	69.6	69.1
School type						
Public	2,819	69.5	69.2	2,830	69.7	69.5
Private	396	67.3	66.0	413	69.4	68.9
Catholic	168	88.0	88.5	176	92.3	92.6
Other private	228	55.6	55.1	237	56.8	57.2
Type of locale						
Large city	688	70.2	70.1	703	73.0	71.6
Mid-size city	640	74.2	73.6	646	75.0	74.3
Urban fringe of large city	955	66.7	66.0	953	66.6	65.9
Urban fringe of mid-size city	244	63.6	62.5	253	66.0	64.8
Large town	85	86.1	84.0	86	87.2	85.0
Small town	288	67.9	65.8	291	69.0	66.5
Rural	315	65.9	66.8	311	63.0	65.9
School size (Kindergartners)						
1 to 12	64	74.4	67.8	67	76.4	71.0
13 to 24	157	61.6	64.9	158	61.1	65.4
25 to 36	114	75.4	77.8	116	75.6	79.2
37 or more	2,880	69.3	68.1	2,902	69.9	68.6
Percent minority enrolled						
0 – 10%	736	69.7	68.1	740	69.7	68.5
11 – 49%	1,070	66.7	67.0	1,083	67.9	67.9
50 – 89%	704	70.0	70.9	726	71.0	73.2
90 – 100%	596	74.7	74.1	573	72.8	71.2
Unknown	109	62.8	59.7	121	65.7	66.3
Region						
Northeast	486	64.0	64.1	490	63.7	64.6
Midwest	574	75.4	73.5	613	77.9	78.5
South	1,288	71.3	69.5	1,267	70.6	68.3
West	867	63.1	64.6	873	64.7	65.1
High grade						
Trans K, K or trans 1	134	69.9	67.3	160	79.1	80.3
01-12	3,019	69.5	69.2	3,017	69.5	69.1
Ungraded	10	77.2	76.9	12	92.4	92.3
Unknown	52	58.5	57.9	54	61.1	60.2

[a]Based on frame data.
[b]Nonblank questionnaire.

Table 5-21. The ECLS-K, spring-kindergarten: number of completed instruments and child-level completion rates for additional data collected

Instrument	Completes	Completion rates	
		Weighted	Unweighted
School Facilities Checklist	21,140	96.6	93.6
Student Record Abstract	16,902	75.9	74.8
Salary and Benefits (principal)	19,628	91.0	86.9
Salary and Benefits (regular teacher)	18,700	88.1	87.3
Salary and Benefits (special education teacher)	240	31.4	30.0
Special Education Teacher Questionnaire A	737	94.1	92.2
Special Education Teacher Questionnaire B	698	87.2	87.4
Adaptive Behavior Scale	32	45.6	45.7

Table 5-22. The ECLS-K: number sampled and number and percent cooperating during fall- or spring-kindergarten, with percent of fall respondents who also responded in the spring

	Sampled	Responded fall or spring kindergarten	Base year response rate		Percent of fall respondents who also responded in spring	
			Weighted	Unweighted	Weighted	Unweighted
Schools						
Before substitution	1,277	944	74.2	73.9	99.3	99.5
After substitution	1,277	1,018	80.7	79.7	99.4	99.6
Children	22,782	20,929	92.0	91.9	95.1	95.3
Parents	22,782	20,141	88.8	88.4	93.7	93.7
Teachers	3,551	3,305	93.5	93.1	-[a]	-[a]

[a] The conditional response rate is not presented for teachers because in the spring a teacher was considered to be a respondent if the teacher questionnaire B had been completed in either the spring or the fall. Therefore, the nominal fall to spring retention rate would be meaningless, i.e., 100 percent.

children and parents. In a longitudinal study such as the ECLS-K, another informative measure is the rate at which respondents in one round of data collection continue to cooperate in the next round. Typically the retention rates from round to round are higher than the initial cooperation rate. This is definitely true of the ECLS-K where the fall-kindergarten child assessment completion rate was 89.9 percent (see table 5-5), but 95.1 percent of fall-assessed children were also assessed in spring-kindergarten. The difference for parents was even more dramatic: 85.3 percent cooperated in the fall, of whom 95.1 percent also participated in the spring. Of schools that cooperated in the fall, over 99 percent continued in the study during the spring.

6. DATA PREPARATION

As described in chapter 5, two types of data collection instruments were used for the Early Childhood Longitudinal Study-Kindergarten Class of 1998-99 (ECLS-K)—computer-assisted and self-administered paper forms (hard copy). The data preparation approach differs with the type of data collection. The direct child assessment and parent interview were conducted using computer-assisted personal interviewing (CAPI) and computer-assisted telephone interviewing (CATI). Editing specifications were built into the computer program used by field staff to collect the data. The teacher and school administrator forms were self-administered. When the field supervisors returned these forms, coders recorded the receipt of these forms into a project-specific forms tracking system. Coders reviewed the questionnaires to ensure data readability for transfer into an electronic format. The visual review included changing any “other” responses that actually fit within the numerical answer structure of the question. (This is called upcoding.) There were some items where upcoding was conducted after the data were keyed due to the large volume of other responses. Once they finished this review, the coders sent the instruments to data entry to be manually transferred to an electronic format and reviewed for range and logic consistency. This section presents the data preparation activities for both types of data collection in more detail.

6.1 Coding and Editing Specifications for CATI/CAPI

The very nature of designing a computer-assisted interview forces decisions about edit specifications to be made up front. Both acceptable ranges and logic consistency checks were pre-programmed into the electronic questionnaire. The next few sections describe the coding and editing of the data collected using CATI/CAPI.

6.1.1 Range Specifications

Within the CATI/CAPI instruments, respondent answers were subjected to both “hard” and “soft” range edits during the interviewing process. A “soft range” is one that represents the reasonable expected range of values but does not include all possible values. Reponses outside the soft range were confirmed with the respondent and entered a second time. For example, the number of hours each week a child attends a day care center on a regular basis had a soft range of 1 to 50. A value outside this range could be entered and confirmed as correct by the assessor as long as it was within the hard range of values (1 to 70).

"Hard ranges" are those that have a finite set of parameters for the values that can be entered into the computer, for example, "1-13 pounds" for birth weight. Out-of-range values for closed-ended questions were not accepted. If the respondent insisted that a response outside the hard range was correct, the assessor could enter the information in a comments data file. Data preparation and project staff reviewed these comments. Out-of-range values were accepted if the comments supported the response.

6.1.2 Consistency Checks (Logical Edits)

Consistency checks, or logical edits, examine the relationship between responses to ensure that they do not conflict with one another or that the response to one item does not make the response to another item unlikely. For example, in the household roster, one could not be recorded as a mother and male. When a logical error such as this occurred during a session, the assessor saw a message requesting verification of the last response and a resolution of the discrepancy. In some instances, if the verified response still resulted in a logical error, the assessor recorded the problem either in a comment or on a problem report.

6.1.3 Coding

Additional coding was required for some items of data collected in the CAPI/CATI instruments. These items included "Other, specify" text responses, occupation, race-ethnicity, and language. Staff were recruited and trained to code these data using coding manuals designed by Westat and NCES to support the coding process. In this section, we describe the coding activities for the CAPI/CATI instruments.

Review of "Other, specify" Items

The "Other, specify" open-ended parent interview responses were reviewed to determine if they should be coded into one of the existing code categories. During the data collection, when a respondent selected an "other" response in the parent interview, the assessor entered the text into a "specify" overlay that appeared on the screen. These text "specify" responses were reviewed by the data preparation staff and, where appropriate, coded into one of the existing response categories. If a text "specify" response for which there was no code occurred frequently enough, new codes were added.

Fall-Kindergarten Parent Interview

Review of the "specify" text responses revealed over 1,000 cases where respondents spoke languages other than the ones specified in the questionnaire as spoken in the household. Over 50 languages beyond the options provided were recorded, and there were over 100 cases with other languages (892 for other languages, 306 for primary language). Given the frequency with which some of the languages appeared, groups of languages were created based on geographic boundaries. These additions were: African language; Eastern European language; Native American language; Sign language; Middle Eastern language; Western European language; Indian subcontinent language; Southeast Asian language; Pacific Islander language; and other language.

Spring-Kindergarten Parent Interview

No additional codes were added from the review of the spring-kindergarten parent interview "Other, specify" items.

Parent Occupation Coding

Occupations were coded using the "Manual for Coding Industries and Occupations," March 1999 (National Household Education Survey, NHES: 99). This coding manual was created for NHES and used an aggregated version of industry and occupation codes. The industry and occupation codes used by NHES were originally developed for the National Postsecondary Student Aid Study (NPSAS, 1990) and contained one to four digits. Analysis of the NPSAS categories revealed that some categories had very small numbers of cases and some categories that are similar in industry or occupation had similar participation rates, suggesting that the separate codes could be collapsed without significant loss of information. The NHES industry and occupation code categories use a one-digit code, the highest level of aggregation, to have sufficient numbers of cases to support analysis without collapsing categories. There are 13 industry codes and 22 occupation codes in the NHES coding scheme. If an industry or occupation could not be coded using this manual, the "Index of Industries and Occupations - 1980" and "Standard Occupational Classification Manual—1980" were used. Both of these manuals use an expanded coding system and at the same time are directly related to the much more condensed NHES coding scheme. The 1980 manuals were used for reference in cases where the NHES coding scheme did not adequately cover a particular situation (see chapter 7, section 7.4 for an expanded description of the industry and occupation codes).[1]

Occupation coding began with an autocoding procedure using a computer string match program developed for the NHES. The program searched the responses for strings of text for each record/case and assigned an appropriate code. About 25 percent of the cases were autocoded and were not verified because there was an exact match between the respondent's answer and the occupation code.

Cases that could not be coded using the autocoding system were coded manually by coders using a customized coding utility program designed for coding occupations. The customized coding utility program brought up each case for coders to assign the most appropriate codes. In addition to the text strings, other information such as main duties, highest level of education, and name of the employer was available for the coders. The coders used this information to ensure that the occupation code assigned to each case was appropriate.

Verification of coding is an important tool to assure quality control and as an extension of coder training. As a verification step, two coders independently assigned codes, i.e., double-blind coding, to industry and occupation cases. A coding supervisor arbitrated disagreements between the two codes, the initial code and the verification code. In the early stages of each coder's work, 100 percent of each coder's work was reviewed. Once the coder's error rate had dropped to one percent or less, ten percent of the coder's work was reviewed.

---

[1] Office of Budget and Management, Executive Office of the President (1980). *Standard Industrial Classification Manual.* Springfield, VA, and Office of Federal Statistical Policy and Standards, U.S. Department of Commerce (1980). *Standard Occupational Classification Manual* (2nd ed.). Washington, DC: Superintendent of Documents, U.S. Government Printing Office.

Partially Complete Parent Interviews

A "completed" parent instrument was defined by whether the section on family structure (FSQ) was completed by the interviewer. Only completed interviews were retained in the final data file. A small number of interviews in each round, approximately 50 (<1 percent) in fall-kindergarten and 75 (<1 percent) in spring-kindergarten, terminated the parent interview after the FSQ section but before the end of the instrument. These interviews were defined as "partially complete" cases and were included in the data file. All instrument items after the interview termination point were set to -9 for "not ascertained."

Household Roster in the Parent Interview

Several tests were run on the household roster to look for missing or inaccurate information. First, the relationship of an individual to the focal child was compared to the individual's listed age and gender. There were 235 cases with inconsistencies (such as a male mother or a biological mother over age 80) that were examined more closely, and any problems found were corrected wherever possible; 231 cases were corrected. Second, households with more than one mother or more than one father were scrutinized for errors. While it is possible to have more than one mother in a household—for example, a household could contain one biological and one foster mother of the focal child—such cases warranted closer inspection. There were 37 cases with more than one mother or father in the household. Corrections were made wherever clear errors existed. Twenty-one cases were corrected, and the other 16 appeared to be correct. Lastly, the relationship of an individual to both the focal child and the reference person was examined, as there are cases in which the relationship of an individual to the focal child conflicts with his status as the spouse/partner of the reference person. For example, in a household containing a child's grandparents, but not his or her parents, we may designate the grandmother as the "mother" figure, and the grandfather thus becomes the "father" by virtue of his marriage to the grandmother. These cases were examined but left unchanged. Both the original—and correct (grandfather)—relationship data and the new "parent-figure" designation (father) that had been constructed were kept.

Race-Ethnicity Coding

Just under 5,000 "Other, Specify" responses were received on the race-ethnicity questions because respondents were allowed to indicate that they belonged to more than one race. Many of these “others” included more than one response (e.g., African American/Asian or American Indian/white). The open responses were coded into one or more of the following seven categories: one Hispanic category; White, non-Hispanic; Black or African American, non-Hispanic; American Indian or Alaskan Native; Asian; Native Hawaiian, or other Pacific Islander; and one unspecified multirace-ethnicity category.

The same coding rules were used to code all race-ethnicity variables for children, resident parents, and nonresident parents. See chapter 7, section 7.4.1 for details on how the race variables were coded and how the race-ethnicity composite was created.

6.2 Coding and Editing Specifications for Hard Copy Questionnaires

6.2.1 Receipt Control

In order to monitor the more than 125,000 documents that were to be received in the base year, a project-specific receipt and document control system was established. The receipt and document control system was initially loaded with the identifying information such as identification numbers for schools, teachers, and children; the links between teachers and children; and the questionnaires that were expected from each school and teacher, for each cooperating school in the sample. As data were collected in the field, field staff completed transmittal forms for each school to indicate which questionnaires were being mailed to the home office. Once data collection started, receipt control clerks compared the transmittal forms to the questionnaires sent in from the field for accuracy and completeness. The identification number on each form was matched against the identification numbers in the tracking system to verify that the appropriate number of forms for each school was returned. The forms were then logged into the receipt and document control system. Once forms were logged in, if they had any data (some forms had no data due to refusal by the respondent to complete them), they were then coded; the data were entered into electronic format, after which the data were edited. The following sections describe the coding, data entry, and editing processes for hard copy questionnaires.

6.2.2 Coding

The hard copy questionnaires required coding of race-ethnicity for teachers, review of “Other, specify” text responses, and a quick visual review of particular questions in each questionnaire.

The quick visual review was to assure that the questionnaire values were accurate and complete and were consistent across variables and that the numbers were converted to the appropriate unit of measurement prior to converting data to an electronic format. The coding staff were trained on the coding procedures and had coding manuals to support the coding process. This staff also did the data editing after data entry was complete. Senior coders verified coding. The verification rate was set at 100 percent for each coder until accuracy of less than one percent error rate was established. After that point, work was reviewed at a rate of ten percent.

Review of “Other, specify” Items

The “Other, specify” text responses were reviewed by the data editing staff and, where appropriate, upcoded into one of the existing response categories. Reexamination of uncoded specify text responses in the fall-kindergarten teacher questionnaires, the spring-kindergarten teacher questionnaires, and the school administrator questionnaire revealed that all responses could be coded into preexisting categories. The small number of specify responses that remained after upcoding those did not fit into any preexisting category and were of insufficient numbers to warrant an additional category. No new codes were added.

Coding Teacher Race-Ethnicity

“Other, specify” text responses for race-ethnicity in the teacher questionnaire B were coded using the procedures described above in section 6.1.3.

Coding Teacher Language

“Other, specify” text responses for language in the teacher questionnaire A were coded using the procedures described above in section 6.1.3.

Coding the Psychomotor Assessment

In fall-kindergarten, the copy form items—copy a plus sign, a square, a triangle, an open-square with a circle, and an asterisk—and the draw-a-person item from the fine motor section of the psychomotor assessment required special coding. The coding protocol for the measure from which the fine motor assessment was adapted was used to code all drawings. A coding sheet was developed that contained both the scoring decisions for the copy forms and the draw-a-person. The coding sheet and coding protocol were reviewed prior to coder training by the authors of the measure from which the fine motor assessment was adapted.

Each copy form was scored separately based on the coding protocol. The first 1,100 forms were sent to one of the authors of the measure from which the fine motor assessment was adapted for additional review as a check on the quality of coding. These forms had ambiguous responses that required more expert adjudication and establishment of additional coding rules. The author made decisions about the drawings, and she coded the forms and established rules that were followed for the remainder of the forms. The coding was verified by the coding supervisor at 100 percent until the coder error rate fell below one percent. Coders were retrained as necessary to follow the coding protocol.

Coding the Facilities Checklist

On question 3a, parts a, b, c, and d, the valid options were “a little,” “some” and “a lot.” There was no response option for the value “none.” However, in approximately 35 to 45 percent of the cases the field staff wrote in a category of “none.” Thus a new category of “none” was included as a response option, even though it was not directly provided in the questionnaire.

6.2.3 Data Entry

Westat data entry staff keyed the forms in each batch. The data were rekeyed by more senior data entry operators at a rate of 100 percent to verify the data entry. The results of the two data entry passes were compared and differences identified. The hard copy form was pulled and examined to determine what corrections had to be made to the keyed data. These corrections were rekeyed. An accuracy rate exceeding 99 percent was the result of this process. The verified batches were then transmitted electronically to Westat's computer system for data editing.

6.2.4 Data Editing

The data editing process consisted of running range edits for soft and hard ranges, running consistency edits, and reviewing frequencies of the results.

Range Specifications

Hard copy range specifications are the set parameters for high and low acceptable values for a question. Where values were preprinted on the forms, these were used as the range parameters. For open-ended questions, high and low ranges were established as acceptable values. Data frequencies were run on the range of values to identify any errors. Values outside the range fell out as an error and were printed on hard copy for a data editor to review. Cases identified with range errors were extracted from the data set, and the original response was updated. Data frequencies were then rerun and reviewed. This iterative process was repeated until no further range errors were found.

Consistency Checks (Logical Edits)

Consistency between variables not involved in a skip pattern was accomplished by programming logical edits between variables (for example, in teacher questionnaire A, the sum of the number of boys and girls in the class could not exceed the total number of children enrolled in the classroom). These logical edits were run on the whole database after all data entry and range edits were complete. The logical edits were run separately for each form. All batches of data were combined into one large data file, and data frequencies were produced. The frequencies were reviewed to ensure the data remained logically consistent within the form. When an inconsistency was found, the case was identified and the inconsistency was printed on paper for an editor to review. The original value was replaced, and the case was then rerun through the consistency edits. Once the case passed the consistency edits, it was appended back into the main data set. The frequencies were then rerun and reviewed. This was an iterative process; it was repeated until no further inconsistencies were found.

Frequency and Cross-Tabulation Review

Frequencies and cross-tabulations were run to determine consistency and accuracy across the various forms and matched against the data in the field management system. If discrepancies could not be explained, no changes were made to the data. For example, in teacher questionnaire A, an item asking about languages other than English spoken in the classroom includes a response option of "No language other than English." If a respondent circled that response but also answered that other languages besides English were spoken in the classroom, then the response was left as recorded by the respondent.

7. DATA FILE CONTENT AND COMPOSITE VARIABLES

This chapter describes the content of the Early Childhood Longitudinal Study-Kindergarten Class of 1998-99 (ECLS-K) Base Year Public-Use Data Files. There are three data files: one each for school, teacher, and child. The school file contains one record for each of the 866 schools providing a school administrator questionnaire; the teacher file contains one record for each of the 3,305 responding teachers; and the child file contains one record for each of the 21,260 responding students. School- and teacher-level data, including composites, are also stored on the child catalog for the convenience of users performing child-level analyses that require school or teacher data. Each of these data files is stored in the root directory of the CD-ROM as an ASCII file; the file names are school.dat, teacher.dat, and child.dat. However, it is strongly recommended that users access the data using the Electronic Code Book (ECB) software available on the CD-ROM rather than access the ASCII files directly. Appendix B contains the record layout for the school, teacher, and child files.

7.1 Identification Variables

Each data file contains an identification (ID) variable that uniquely identifies each record. For the school file, the ID variable is S_ID; for the teacher file, the ID variable is T_ID; and for the child file, the ID variable is CHILD_ID.

Each type of respondent has a unique ID number. The school ID number is the base for all the subsequent ID numbers as children, parents, and teachers were sampled from schools. The school ID number is a four-digit number assigned sequentially to sampled schools. The school ID has a series of ranges: 0001-1299 for originally sampled schools; 2000 series for new schools added to the sample during the sample freshening process; 3000 series for substitute schools that replaced nonresponding original sample schools; and 4000 series for transfer schools, which were assigned during processing at the home office. (See chapter 4 for a complete description of the ECLS-K sample.)

The child ID number is a concatenation of the school ID where the child was sampled, a three-digit student number and the letter “C.” For example, 0001001C is the ID number of the first child sampled in school 0001. The teacher ID number is a concatenation of the school ID where the teacher was sampled, the letter “T” and a two-digit teacher number. For example, 0001T01 is the ID number for the first teacher sampled in school 0001. The parent ID number is linked to the child ID number and is a concatenation of the school ID, the child ID and the letter "P." For example, 0001001P is the ID number of the parent of the first child sampled in school 0001. If twins are sampled in a particular household, the ID of the first child sampled is used to generate the parent ID. For twins, there will be two child-level records with the same parent ID. Children with the same teacher can be identified by finding all children on the child file with the same teacher ID.

7.2 Missing Values

All variables on these files use a standard scheme for missing values. Codes are used to indicate item nonresponse, legitimate skips, unit nonresponse, and suppressed data.

-1	Not Applicable (legitimate skip)
-2	Data Suppressed
-7	Refused
-8	Don't Know
-9	Not Ascertained
(blank)	System Missing

The "not applicable" code (-1) indicates that the respondent did not answer the question due to skip instructions within the instrument or external reasons that lead a respondent to not participate. For the child file where the respondents were directly interviewed, a "not applicable" is coded for items that were not asked of the respondent because of a previous answer given. For example, an item about a sibling's age is not asked when the respondent has indicated that the child has no siblings. A "not applicable" code is also used in the direct child assessment if a child did not participate in any section due to language or a disability. For the teacher and school files where the instruments are self-administered, a "not applicable" is coded for items that the respondent left blank because the written directions instructed them to skip the item due to a certain response on a previous item.

Depending on the research question being addressed, cases with a value of -1 may need to be recoded. For example, a researcher interested in the average number of hours kindergarten children spend in relative care, would include both children who are in such care and those who do not receive relative care, i.e. children whose overall exposure is zero hours. These cases would have a -1 for the question on number of hours in relative care, since they would have been skipped out of the series once the respondent had answered that the child had no relative care. Thus, the -1s would first have to be converted to zero before being included in the computations. If the same researcher was interested in the average time children in relative care spend in such care, he would simply average the cases with valid data values, i.e. those with no missing values or -1s.

The “data suppressed” code (-2) indicates that the data for that variable are suppressed in order to protect the identity of the respondent or child. When the data for a variable are suppressed, all the cases have a value of -2 for that variable. A comment, “This data is suppressed for respondent confidentiality,” is displayed in the comment field in the electronic code book.

The “refused” code (-7) indicates that the respondent specifically told the interviewer that he or she would not answer the question. This, along with the “don’t know” code and the “not ascertained” code, indicates item nonresponse. This code rarely appears in the school and teacher files because it indicates the respondent specifically wrote something on the questionnaire indicating an unwillingness to answer the question.

The “don’t know” code (-8) indicates that the respondent specifically told the interviewer that he or she does not know the answer to the question (or in rare cases on the self-administered questionnaires, “I don’t know” was written in for the item). The “don’t know” code was also used in the direct child assessment when children did not answer a particular question after procedures had been followed to repeat the question and try it again. For items where “don’t know” is one of the options explicitly provided, a “-8” will not be coded for those that choose this option; instead the “don’t know” response will be coded as indicated in the value label information for that item.

The “not ascertained” code (-9) indicates that the respondent left the item blank that he or she should have answered. For the school and teacher self-administered questionnaires, this is the primary code for item nonresponse. For data outside the instruments, e.g., direct assessment scores, a –9 means that a value was not ascertained or could not be calculated due to nonresponse.

System missing appears as a blank when viewing code book frequencies and in the ASCII data files. System missing codes (blanks) indicate that an entire instrument or assessment is missing due to unit nonresponse. For example, if a child’s parent did not participate in the parent interview, then all items from the parent interview will have a system missing. These may be translated to another value when the data are extracted into specific processing packages. For instance, SAS will translate these blanks into periods (".") for numeric variables.

Missing values for composite variables were coded in a similar fashion. If a particular composite was inappropriate for a given household—as the variable P1MOMID was for a household with no resident mother—that variable was given a value of -1. In instances where a variable was appropriate, but complete information to construct the composite was not available, the composite was given a value of -9. The "Refused" and "Don't Know" codes were not used for the composites, except in the calculations of the height, weight, and body mass index (BMI) composites for fall-kindergarten and spring-kindergarten.[1]

The ECLS-K Base Year Public-Use Data Files are provided on a CD-ROM and are accessible through an ECB that allows data users to view variable frequencies, tag variables for extraction, and create the SAS, SPSS for Windows, or STATA code needed to create an extract each file for analysis. The three data files on the ECB—school, teacher, and child—are each referred to as a "catalog." Instructions for using the CD-ROM and ECB are provided in chapter 8.

7.3 Variable Naming Conventions

Variables were named according to the data source, (e.g., parent interview, teacher questionnaire) and the data collection point (e.g., fall-kindergarten, spring-kindergarten). These variable names are used consistently throughout all three catalogs. In general, variable names start with the following prefixes:

A1	Data collected/derived from fall-kindergarten teacher questionnaire A.
A2	Data collected/derived from spring-kindergarten teacher questionnaire A.
B1	Data collected/derived from fall-kindergarten teacher questionnaire B.

[1] Children's height and weight measurements were each taken twice to prevent error and provide an accurate reading. Children's BMI was calculated based on height and weight. The rules for using "don't know" and "not ascertained" codes for these values was as follows. If both the first and second measurement of height in the child assessment were coded as -8 (don't know), then the height composite was coded as -8 (don't know). If both the first and second measurements of weight were coded as -8 (don't know), the weight composite was coded as -8 (don't know). If either the height or weight composites were coded as not ascertained (-9), the BMI composite was coded as not ascertained (-9). If neither the height nor weight composites were coded as not ascertained, and either the height or weight composite was coded as -8 (don't know), then the BMI composite was coded as -8 (don't know).

B2	Data collected/derived from spring-kindergarten teacher questionnaire B.
BY	Base year panel weight variables.
C1	Data/scores collected/derived from fall-kindergarten direct child assessment and fall-kindergarten weight variables.
C2	Data/scores collected/derived from spring-kindergarten direct child assessment and spring-kindergarten weight variables.
F1	Data from fall-kindergarten field management system (FMS).
F2	Data from spring-kindergarten FMS.
FK	Data from base year (cross round) FMS.
IF	Imputation flags.
K2	Data collected/derived from spring-kindergarten school facility checklist.
P1	Data/scores collected/derived from fall-kindergarten parent interview.
P2	Data/scores collected/derived from spring-kindergarten parent interview.
R1	Derived child demographic or child status variables for fall-kindergarten.
R2	Derived child demographic or child status variables for spring-kindergarten.
S2	Data collected/derived from spring-kindergarten school administrator questionnaire.
T1	Data/scores collected/derived from fall-kindergarten teacher questionnaire C.
T2	Data/scores collected/derived from spring-kindergarten teacher questionnaire C.
WK	Base year (cross round) parent composite variables.

A few exceptions that do not follow the above-mentioned prefix convention are:

The identifiers CHILDID, PARENTID, T1_ID, T2_ID, S1_ID, and S2_ID.

School demographic variables from the sample frame that are named CREGION (Census region), and CS_TYPE2 (type of school), KURBAN (location type, e.g., large city, small town).

In general, all composites derived from a given source maintain the same prefix as the source. Some composite variables, however, combined information across data collection points and/or several sources and are consequently not associated with any prefixes. Derived child demographic variables, gender, race-ethnicity, and date of birth, were created from the best source of data and are named GENDER, RACE, DOBMM, DOBDD, and DOBYY. Other such derived variables include CREGION, FKCHGSCH, FKCHGTCH, KURBUN, R1_KAGE and R2_KAGE. Sources and other details for these and all other composite variables can be found in table 7-6 .

7.4 Composite Variables

To facilitate analysis of the survey data, a group of composite variables were created and added to the child, teacher and school data files. Variables based on the child assessment include height, weight, and BMI. Variables based on the teacher data include class type (e.g., AM, PM, or all-day kindergarten class), student-teacher instructional ratio, and teacher's age. Variables constructed from the school data include the percentage of minority students, school type, and school instructional level. Variables constructed from the parent interview data include parent identifiers, parent demographics, household composition, household income and poverty, childcare, and child demographics.

Table 7-6 lists all the composite variables. All basic child demographic items (gender, age, race-ethnicity, and date of birth) are listed first. Child care variables follow the demographics, and then household composition and imputed variables are listed. Demographics for parents are next; resident father and mother characteristics are first, followed by characteristics of nonresident biological parents and nonresident adoptive parents. Teacher, classroom, and school variables are listed last. Once the user identifies the composites of interest, he or she can refer to table 8-8 for instructions on accessing the variables from the ECB.

7.4.1 Child Variables

Child Height Composite (C1HEIGHT and C2HEIGHT)

Children's height was measured twice. For the height composite (C1HEIGHT and C2HEIGHT), if the two height values from the instrument were less than two inches apart, the average of the two height values was computed and used as the composite value. Otherwise, the value that was closest to 43 inches (the average height for a five year old child) was used as the composite.

Child Weight Composite (C1WEIGHT and C2WEIGHT)

Children's weight was measured twice. For the weight composite (C1WEIGHT and C2WEIGHT), if the two weight values from the instrument were less than five pounds apart, the average of the two values was computed and used as the composite value. Otherwise, the value that was closest to 40 pounds (the average weight for a five year old child) was used as the composite value.

Child BMI Composite (C1BMI and C2BMI)

Composite BMI (C1BMI and C2BMI) was calculated by multiplying the composite weight in pounds by 703.0696261393 and dividing by the square of the child's composite height in inches.

Child Date of Birth Composite (DOBYY, DOBMM, and DOBDD)

The child date of birth composite was created using parent interview data and, in cases in which the parent interview data did not exist or were outside of the criteria for inclusion, using the FMS data. If the date of birth given was before June 1, 1990, or after March 31, 1995, the data were excluded.

Child Age at Assessment Composite (R1_KAGE, R2_KAGE)

The child's age was calculated by determining the number of days between the child assessment date and the child's date of birth. The value was then divided by 30 to calculate the age in months.

Gender Composite (GENDER)

The gender composite was derived using the gender indicated in the parent interview and, if it was missing, the FMS. Also, if the gender was different in the fall-kindergarten parent interview and spring-kindergarten parent interview, the FMS gender was used.

Race-Ethnicity Composites

The data on race-ethnicity is presented in the ECLS-K files in two ways. Since a respondent was allowed to indicate that they belonged to more than one of the five race categories (White, Black or African American, American Indian or Alaskan Native, Asian, Native Hawaiian or other Pacific Islander), a series of five dichotomous race variables were created that indicated separately whether the respondent belonged to any of the five specified race groups. In addition one more dichotomous variable was created for those who had simply indicated that they were multiracial without specifying the race (e.g., biracial).

Data was collected on ethnicity as well. Respondents were asked if they were Hispanic or not. Using the six race dichotomous variables and the Hispanic ethnicity variable, a race-ethnicity composite variable was created. The categories were: White, non Hispanic; black or African American, non-Hispanic; Hispanic, race specified; Hispanic, no race specified; Asian; Native Hawaiian or other Pacific Islander; American Indian or Alaskan Native, and more than one race specified, non-Hispanic.

The retention of the dichotomous on the file allows users to create different composites as needed.

7.4.2 Family and Household Variables

The list of all composite variables, including variables used to derive the composites, is given in table 7-6 at the end of this chapter. Several household and family composite variables were created. The creation of one of these, socioeconomic status (SES), is described below.

The socioeconomic scale (SES) variable was computed at the household level for the set of parents who completed the parent interview in fall-kindergarten or spring-kindergarten. The SES variable reflects the socioeconomic status of the household at the time of data collection for spring-kindergarten (Spring 1999). The components used for the creation the SES were:

Father/male guardian’s education;

Mother/female guardian’s education;

Father/male guardian's occupation;

Mother/female guardian's occupation; and

Household income.

The parent's occupation was recoded to reflect the average of the 1989 General Social Survey (GSS) prestige score[2] of the occupation. It was computed by averaging the corresponding prestige score of the 1980 Census occupational category codes covered by the ECLS-K occupation. Table 7-6 provides details on the prestige score values.

The variables were collected as follows:

1. **Income.** The information about income was collected in spring-kindergarten. As a result, income is missing for all households with parents who did not participate in the survey in spring-kindergarten.

2. **Parent's education.** The information about parent's education was collected in round 1. For households not interviewed in fall-kindergarten (e.g., parents of children in refusal-converted schools), this information was collected in spring-kindergarten.

3. **Parent's occupation.** The information about parent's education was collected in fall-kindergarten only.

Because not all the parents responded to all the questions or were respondents in both rounds, there were missing values for some of the components of the SES indicator. The amounts of missing data for these variables were small percentages, with income having the largest percentage missing (see table 7-1 below).

Table 7-1. Missing data for SES variables

Variable	Number Missing	Percent
Mother's Education	414	2.1%
Father's Education	756	3.8%
Mother's Occupation	2256	11.3%
Father's Occupation	2252	11.3%
Household Income	5630	28.2%

[2] Nakao, K., and Treas, J. (1992). *The 1989 Socioeconomic Index of Occupations: Construction from the 1989 Occupational Prestige Scores:* General Social Survey Methodological Report No. 74. Chicago: NORC.

A hot deck imputation methodology was used to impute for missing values of all components of the SES. In hot deck imputation, the value reported by a respondent for a particular item is given or “donated” to a “similar” person who failed to respond to that question. Groups or cells use auxiliary information known for both donors and nonrespondents. Ideally, donors and nonrespondents have similar characteristics in the cell. The value to impute is from the randomly selected donor among the respondents within the cell.

The SES component variables were highly correlated so a multivariate analysis was more appropriate for examining the relationship of the characteristics of donors and nonrespondents. A categorical search algorithm called CHAID (Chi-squared Automatic Interaction Detector) was used to divide the data into cells based on the distribution of the variable to be imputed. The analysis used the records with no missing values for the variable being imputed. CHAID not only analyzed and determined the best predictors but also created the cells that were used for hot deck imputation.

The variables were imputed in a sequential order and separately by type of household (female single parent, male single parent, and both parents present). For households with both parents present, the mother’s and father’s variables were imputed separately. The new imputed values were used in the creation of the imputation cells if these values had been already imputed. If this was not the case, an “unknown” or missing category was created as an additional level for the CHAID analysis. As a rule, no imputed value was used as a donor. In addition, the same donor was not used more than two times. The order of the imputation for all the variables was from the lowest percent missing to the highest. Occupation imputation involved two steps. First, the labor force status of the parent was imputed, i.e., whether the parent was employed or not. Then the parent’s occupation was imputed only for those parents whose status was identified as employed either through the parent interview or the first imputation step. The variable for income was imputed last using a three-stage procedure, where if a respondent provided some partial information about income, that was used in the imputation process as well. For example, some respondents did respond in the parent interview to the item on whether they earned more or less than $25,000 but did not respond to the items asking for further details.

For example, for households where both parents are present, variables were imputed in the following order:

Mother’s education;

Father’s education;

Mother's labor force status;

Mother's occupation if the mother was employed or imputed as employed;

Father's labor force status;

Father's occupation if the father was employed or imputed as employed;

Income figure for households where the detailed income ranges are known;

Income figure for households where the broad income ranges are known; and

Income figure for households with no information about income.

Once the components of the SES variable were imputed, their corresponding *z*-score or normalized value was computed. The expression of *z*-score z_{hi} for the *h*-th component in the *i*-th household is

$$z_{hi} = \frac{x_{hi} - \bar{x}_w}{se(\bar{x}_w)}$$

where x_{hi} is the value of the *h*-th SES component for the *i*-th household;
w_i is the base weight for the *i*-th record;
$\bar{x}_w$ is the weighted mean of x_{hi}; and
$se(\bar{x}_w)$ is the standard error of $\bar{x}_w$.

That is, each component converted to a *z*-score with mean of 0 and a standard deviation of one. For income, the component x_i is defined as the logarithm of the income for *i*-th household. The logarithm of income was used as its distribution is less skewed than the direct income values. The SES value for the *i*-th household was then computed as

$$SES_i = \frac{\sum_{h=1}^{m_i} z_{hi}}{m_i}$$

where m_i is number of non-missing SES components for the *i*-th household. WKSESL is a continuous variable that ranges from –4.75 to 2.75. As described, the SES composite is the average of up to five measures, each of which was standardized to have a mean of 0 and a standard deviation of 1, hence the negative valves. For analyses that require a continuous SES measure, such as multivariate regressions, WKSESL is the variable to use. For categorical analyses, use WKSESQ5.

Note that for households with only one parent present, not all the components were defined. In these cases, the SES was computed averaging the available components.

A categorical SES variable was created to indicate the quintile for the value of the SES. The quintiles were computed at the child level using the spring-kindergarten parent weights.

The imputed income variable was also used to create a household-level poverty variable. Income was compared to preliminary Census poverty thresholds for 1998, which vary by household size. Households whose income fell below the appropriate threshold were classified as poor (see table 7-2). For example, if a household contained two members, and the household income was lower than $10,973, then the household was considered to be in poverty.

Table 7-2. Preliminary Census poverty thresholds for 1998[1]

Household Size	Weighted Average Thresholds
2	$10,973
3	$13,001
4	$16,655
5	$19,682
6	$22,227
7	$25,188
8	$28,023
9+	$33,073

[1] U.S. Census Bureau, Current Population Survey. http://www.census.gov/hhes/poverty/threshld/thresh98.html.

7.4.3 Teacher Variables

The list of all composite variables, including the variables used to derive the composites, is given in table 7-6 at the end of this chapter. Details about how one of the teacher composites, class type was created are provided below.

Base Year Class Type Composite (KGCLASS)

This composite indicates the type(s) of kindergarten class(es) teachers taught, (i.e., AM only, PM only, AM and PM, all day). If the fall-kindergarten class type and spring-kindergarten class type were the same, then that class type was used as KGCLASS. If either data collection point was missing, the remaining class type was used as KGCLASS. If one of the data collection points showed an answer that indicated overlapping times of day for instruction (the teacher indicated that he or she taught AM and all day classes, PM and all day classes, or AM, PM, and all day classes) and the other data collection point had a more typical answer (AM only, PM only, AM and PM, or all day), then the typical answer was used as KGCLASS. If the two data collection points both had consistent answers but they were different from each other, then the FMS data were used if it was in agreement with one of the data collection points. If there were two different overlapping answers in the two data collection points or the FMS was not in agreement with either of the two data collection points, then the fall-kindergarten answer was used. Overlapping answers, e.g., AM and all day, were used only when they were provided in both fall- and spring-kindergarten. If both data collection points had missing data, then the FMS data were used. Since KGCLASS is a teacher level variable it is only present in the teacher level dataset.

There were 1,063 respondents who chose the “other specify” category for question 3e in the teacher questionnaire C, student information section. A new variable, TZTT1OTG, was created that classified the “other specifies” from question 3e into the following five categories: ESL/Bilingual, Handicapped/Special Education, School-wide Title I, Computer, and other. It is important to remember that these responses were not offered to all respondents, and thus, should not be used to make estimates such as percentage of children participating in Title I computer programs. The goal of this variable was to provide some information on the large number of “other specifies” in this question.

7.4.4 School and Class Characteristics

The list of all composite variables, including the variables used to derive the composites, is given in table 7-6 at the end of this chapter. Variables on school and class characteristics were constructed from the teacher and school data. Details on how some of the variables were created follow.

Percent Minority Students (S2KMINOR)

The percent of minority students was derived from answers to the spring-kindergarten school administrator questionnaire by determining the percentage of children who were either of Hispanic origins (question 14) or who were American Indian or Alaskan Native, Asian, Black or African-American, or Native Hawaiian or Other Pacific Islander (question 15) to create the percent minority composite. In the questionnaire design, it was assumed that the school administrator would allow for overlap between the ethnicity and race. For example, 20 percent of the children could be listed as Hispanic, and these same children's races could be indicated in the next question, such that the percentages for different races in the school would add to 100 percent. However, this is not how all the questionnaires were answered; therefore, rules were established for the following five types of schools to accommodate different answer patterns. It is important to note that the term "race variables" refers to the variables obtained from question 15 that pertain solely to race, whereas the terms "ethnicity" or "Hispanic" refer to the ethnicity variable obtained from question 14. The term "race-ethnicity" variables refer to the sum of the variables from questions 14 and 15, i.e., both the race and ethnicity variables.

- **Type 1: Schools with missing race or Hispanic data.** For the schools that had all race variables missing, the Hispanic variables missing, or both race and the Hispanic variables missing, data from the school sample frame was used to create the percent minority composite. If there was no school sample frame data for certain records, then the composite was not created for those records.

- **Type 2: Schools with race and Hispanic variables totaling less than 97 percent.** For the schools that had race and Hispanic variables with percentages totaling less than 97 percent, school sample frame data were used to create composite.

- **Type 3: Schools with race-ethnicity variables (questions 14 and 15) greater than 103 percent but race variables sum to less than 97 percent.** For the schools that had a total of race and Hispanic variable percentages greater than 103 percent but with the sum of the race variables less than 97 percent, school sample frame data were used to create the composite.

- **Type 4: Schools with total race-ethnicity percentages (sum of questions 14 and 15) between 97 percent and 103 percent.** Several schools had the total percentages of Hispanic and race variables (question 14 and question 15) total to between 97 percent and 103 percent. Overall, these cases included two types of cases: a) where items in question 14 (Hispanicity) and question 15 (race) cumulatively added up to 100 percent, it was assumed that this indicated that the school administrator did not allow an overlap between ethnicity and race and therefore none of the percentages for whites were meant to indicate Hispanic whites; and b) cases where the percentages given for Hispanic was so small that it was less relevant how the principal interpreted these items and the total of question 14 and question 15 was not equal to 100 percent, but lay between 97 and 103 percent. Of these cases, some add up to exactly 100 percent, and so no adjustment was needed prior to calculating percentage minority. For the latter cases, the percent white was adjusted so that the total equaled 100. In order to do so, the percent white was divided by reported total (sum of question 14 and question 15) then multiplied by 100 to get adjustment. Then, 100 minus adjusted percent white was used for the percent minority.

- **Type 5: Schools with race-ethnicity percentages greater than 103 percent and race variables sum to between 97 percent and 103 percent.** Many schools that had Hispanic and race variables that added to more than 103 percent and totals for race variables that were between 97 percent and 103 percent. This indicated that the school administrator interpreted these items to allow for overlap (e.g., there could be Hispanic whites, Hispanic African-Americans, etc.)—some of the whites may be Hispanic whites. It could not be known with certainty how many of the whites were Hispanic and how many were non-Hispanic. For these schools, the first step was to determine the range of possible values for percent minority (100 minus non-Hispanic whites) given the data. The following example illustrates this process:

  Example: Hispanic = 60 percent; and Asian = 1 percent, Native Indian = 1 percent, African-American = 12 percent, Pacific Islander = 0 percent, white = 88 percent, and Total Races = 100 percent. Therefore, total of percent given for all races and Hispanic = 160 percent.

The next step was to determine the highest possible percent minority. This value was calculated assuming that all Hispanics are white-Hispanics. The Hispanic percent was subtracted from the percent white. (If percent Hispanic is greater than percent white, then percent white becomes zero.) Then remaining white percentage was subtracted from 100 to equal highest possible percent minority.

Example: Non-Hispanic whites = 88 percent (whites) – 60 percent (Hispanic) = 28 percent. Therefore, highest percent minority = 100 percent - 28 percent (non-Hispanic whites) = 72 percent.

The next step was to determine the lowest possible percent minority. This value was calculated assuming that all of the African-Americans were African-American-Hispanics and that any remaining Hispanics were white-Hispanics. Total African-Americans were subtracted from Hispanics, then only those remaining Hispanics were subtracted from whites to estimate the lowest possible percentage of non-Hispanic whites. Next, non-Hispanic whites were subtracted from 100 to get the lowest possible percent minority.

Example: Non-African-American Hispanics = 60 percent (Hispanics) - 12 percent (African-Americans) = 48 percent; Non-Hispanic whites = 88 percent (whites) - 48 percent (non-African-American Hispanics) = 40. Therefore, lowest percent minority = 100 percent - 40 percent (non-Hispanic whites) = 60 percent.

The last step was to compare this range of possible percent minority values to the school sample frame value. If the school sample frame value was within range of the possible values, then the school sample frame value was used as the percent minority composite. If the school sample frame value was outside of the range of possible values, then school sample frame values were not used because the range of values in the school administrator questionnaire offered at least some improvement for an estimate over the school sample frame alone. If the school sample frame value was lower than the range of values, the lowest possible percent minority was used as the composite estimate. If the school sample frame value was higher than the range of values, the highest possible percent minority was used as the composite estimate. For example, if the range of possible percent minority was 60 to 72 and the school sample frame value was 73, then the percent minority was estimated as 72.

School Type Composite (S2KSCTYP)

Questions 5 (whether school is public) and 7 (type of private school) from the school administrator questionnaire, along with school sample frame data, were used to create this variable. If the response to question 5 (Is this a public school?) was “Yes,” then S2KSCTYP was coded “public.” If the response to question 5 was “No,” and the response to question 7.a. (Is your school a Catholic school) was

"Yes," then the school was coded as "Catholic." If the responses to both questions 5 and 7.a. were No," and the response to question 7.b. (Is your school private with another religious affiliation?) was "Yes," then S2KSCTYP was coded as "private, other religious." Otherwise, a school administrator questionnaire with valid answers to questions 5 and 7, then S2KSCTYP was coded as "private, non-religious." If there was no school administrator questionnaire, then school sample frame data were used to create the composite.

School Instructional Level Composite (S2KSCLVL)

The purpose of this composite is to classify schools based on the highest grade taught in the school. Question 4 (grade levels included in the school) of the school administrator questionnaire was used to create this composite. The highest grade level circled on the form was determined, and the grade level was classified accordingly. If question 4 was left blank, question 10 from the school administrator questionnaire (grade levels that participate in special programs, if applicable) was used as a proxy for question 4. If the respondent did not answer questions 4 and 10, or there was no school administrator questionnaire, then school sample frame data were used to determine the value for the composite.

7.4.5 Parent Identifiers and Household Composition

The construction of parent identifiers and the household composition variables from the parent interview data was a two-step process. First, individuals identifying themselves as the child's mother/father were located within the household roster, and their type (biological, adoptive, foster, step, partner, or unknown) was established. For households containing more than one father or mother, a hierarchy was used to designate the "current," or residential, parent of each sex. The biological parent, if present, was always the current mother or father. In the absence of a biological parent, the current mother/father designation was given to the adoptive, step, foster/guardian, partner, or "unknown-type" parent. This information, along with household size and presence or absence of grandparents, siblings, and other relatives was used to construct the household composition variables P1HPARNT, P2HPARNT, P1HDAD, P2HDAD, P1HMOM, P2HMOM, P1HFAMIL, and P2HFAMIL and parent type variables P1MOMTYP, P1DADTYP, P2MOMTYP, and P2DADTYP.

After the actual residential parents were identified and the composite variables were constructed, in any household without a parent the household respondent (and his or her spouse, if applicable) was assigned as a "parent figure." Parent demographic variables (including age, race-ethnicity, and education) were then constructed for all parents/parent "figures." It should be noted, however, that these parent "figures" were not incorporated into the construction of the household composition variables described above.

It should be noted that because the composite construction identifies only one resident mother and/or one resident father, same sex parents are not readily identified in the composites themselves. Two approaches can be used to identify these couples. First, the user should search the relationship variables (P1REL_1, P2REL_1, etc.) to identify households in which more than one person identifies himself or herself as a father/mother to the focal child. Secondly, since not all same-sex partners identify themselves as "mother" or "father" to the focal child, the user should also search for households in which the respondent (identified by P1PER_1, P2PER_1, etc.) is the child's parent and the respondent's spouse (identified from P1SPOUSE, P2SPOUSE) is the same sex as the respondent.

There are four sections in the parent interview that asked about parent-figure-specific questions:

ALQ	Years child lived with person
PLQ	Primary language
PEQ	Parent education
EMQ	Employment

Each of these sections may be completed during the parent interview about a different household member but about no more than four household members. Rather than reserve space (mostly unused) for these sections with each of the 17 household members (the FSQ section), these sections are repeated four times, with "pointer" or "foreign key" variables that hold the original number of the household member. For instance, if household member #3 is the first person to receive the ALQ section, then the pointer variable P1ALQHH1 will hold the value "3," and the actual ALQ variable P1YRS_1 will hold the number of years the child spent with the household member specified by the value of P1ALQHH1. Table 7-3 identifies the pointer variables.

Table 7-3. Pointers to parent figure questions

Person Pointer		Interview Item	
P1ALQHH1	P1 ALQ010 HH PERSON POINTER 1	P1YRS_1	P1 ALQ010 YEARS CHD LIVED WITH PERSON 1
		P1MON_1	P1 ALQ010 MONTHS CHD LIVED WITH PERSON 1
P1ALQHH2	P1 ALQ010 HH PERSON POINTER 2	P1YRS_2	P1 ALQ010 YEARS CHD LIVED WITH PERSON 2
		P1MON_2	P1 ALQ010 MONTHS CHD LIVED WITH PERSON 2
P1ALQHH3	P1 ALQ010 HH PERSON POINTER 3	P1YRS_3	P1 ALQ010 YEARS CHD LIVED WITH PERSON 3
		P1MON_3	P1 ALQ010 MONTHS CHD LIVED WITH PERSON 3
P1ALQHH4	P1 ALQ010 HH PERSON POINTER 4	P1YRS_4	P1 ALQ010 YEARS CHD LIVED WITH PERSON 4
		P1MON_4	P1 ALQ010 MONTHS CHD LIVED WITH PERSON 4
P1PLQHH1	P1 PLQ083-090 HH PERSON POINTER 1	P1RES_1	P1 PLQ083 PERSON 1 LANGUAGE TO CHILD
		P1CHL_1	P1 PLQ090 CHILD'S LANGUAGE TO PERSON 1
P1PLQHH2	P1 PLQ083-090 HH PERSON POINTER 2	P1RES_2	P1 PLQ083 PERSON 2 LANGUAGE TO CHILD
		P1CHL_2	P1 PLQ090 CHILD'S LANGUAGE TO PERSON 2
P1PLQHH3	P1 PLQ083-090 HH PERSON POINTER 3	P1RES_3	P1 PLQ083 PERSON 3 LANGUAGE TO CHILD
		P1CHL_3	P1 PLQ090 CHILD'S LANGUAGE TO PERSON 3
P1PLQHH4	P1 PLQ083-090 HH PERSON POINTER 4	P1RES_4	P1 PLQ083 PERSON 4 LANGUAGE TO CHILD
		P1CHL_4	P1 PLQ090 CHILD'S LANGUAGE TO PERSON 4
P1PEQHH1	P1 PEQ020-080 HH PERSON POINTER 1	P1HIG_1	P1 PEQ020 PERS 1 HIGHEST EDUCATION LEVEL
		P1HIS_1	P1 PEQ030 IF PERS 1 HIGH SCHOOL DIPLOMA
		P1ENR_1	P1 PEQ050 IF PERS 1 ENROLLED IN COURSES
		P1FPT_1	P1 PEQ060 PERS 1 COURSE FULL/PART TIME
		P1TRN_1	P1 PEQ070 IF PERSON 1 GETS JOB TRAINING
		P1WKL_1	P1 PEQ080 PERS 1 WEEK HOURS IN TRAINING
P1PEQHH2	P1 PEQ020-080 HH PERSON POINTER 2	P1HIG_2	P1 PEQ020 PERS 2 HIGHEST EDUCATION LEVEL
		P1HIS_2	P1 PEQ030 IF PERS 2 HIGH SCHOOL DIPLOMA
		P1ENR_2	P1 PEQ050 IF PERS 2 ENROLLED IN COURSES
		P1FPT_2	P1 PEQ060 PERS 2 COURSE FULL/PART TIME
		P1TRN_2	P1 PEQ070 IF PERSON 2 GETS JOB TRAINING
		P1WKL_2	P1 PEQ080 PERS 2 WEEK HOURS IN TRAINING
P1PEQHH3	P1 PEQ020-080 HH PERSON POINTER 3	P1HIG_3	P1 PEQ020 PERS 3 HIGHEST EDUCATION LEVEL
		P1HIS_3	P1 PEQ030 IF PERS 3 HIGH SCHOOL DIPLOMA
		P1ENR_3	P1 PEQ050 IF PERS 3 ENROLLED IN COURSES
		P1FPT_3	P1 PEQ060 PERS 3 COURSE FULL/PART TIME
		P1TRN_3	P1 PEQ070 IF PERSON 3 GETS JOB TRAINING
		P1WKL_3	P1 PEQ080 PERS 3 WEEK HOURS IN TRAINING

Table 7-3. Pointers to parent figure questions (continued)

Person Pointer		Interview Item	
P1PEQHH4	P1 PEQ020-080 HH PERSON POINTER 4	P1HIG_4	P1 PEQ020 PERS 4 HIGHEST EDUCATION LEVEL
		P1HIS_4	P1 PEQ030 IF PERS 4 HIGH SCHOOL DIPLOMA
		P1ENR_4	P1 PEQ050 IF PERS 4 ENROLLED IN COURSES
		P1FPT_4	P1 PEQ060 PERS 4 COURSE FULL/PART TIME
		P1TRN_4	P1 PEQ070 IF PERSON 4 GETS JOB TRAINING
		P1WKL_4	P1 PEQ080 PERS 4 WEEK HOURS IN TRAINING
P1EMQHH1	P1 EMQ020-150 HH PERSON POINTER 1	P1PAY_1	P1 EMQ020 PERS 1 HAD PAID JOB LAST WEEK
		P1VAC_1	P1 EMQ030 IF PERS 1 ON LEAVE PAST WEEK
		P1JOB_1	P1 EMQ040 PERSON 1 NUMBER OF ALL JOBS
		P1HRS_1	P1 EMQ050 PERSON 1 HOURS/WK AT ALL JOBS
		P1LOK_1	P1 EMQ060 PERS 1 SOUGHT JOB LAST 4 WEEKS
		P1DO1_1	P1 EMQ070 PERS 1 CHKD W/PUB EMPL AGNCY
		P1DO2_1	P1 EMQ070 PERS 1 CHKD W/PRIV EMP AGNCY
		P1DO3_1	P1 EMQ070 PERS 1 CHKD W/EMPLOYR DIRECTLY
		P1DO4_1	P1 EMQ070 PERS 1 CHKD W/FRIENDS & REL
		P1DO5_1	P1 EMQ070 PERS 1 PLACED OR ANSWERED ADS
		P1DO6_1	P1 EMQ070 PERS 1 READ WANT ADS
		P1DO7_1	P1 EMQ070 PERS 1 DID SOMETHING ELSE
		P1DOW_1	P1 EMQ080 WHAT PERSON 1 DOING LAST WEEK
		P1TAK_1	P1 EMQ100 PERS 1 JOB AVAILABLE LAST WEEK
		P1OCC_1	P1 EMQ130-50 1ST PERSON OCCUPATION CODE
P1EMQHH2	P1 EMQ020-150 HH PERSON POINTER 2	P1PAY_2	P1 EMQ020 PERS 2 HAD PAID JOB LAST WEEK
		P1VAC_2	P1 EMQ030 IF PERS 2 ON LEAVE PAST WEEK
		P1JOB_2	P1 EMQ040 PERSON 2 NUMBER OF ALL JOBS
		P1HRS_2	P1 EMQ050 PERSON 2 HOURS/WK AT ALL JOBS
		P1LOK_2	P1 EMQ060 PERS 2 SOUGHT JOB LAST 4 WEEKS
		P1DO1_2	P1 EMQ070 PERS 2 CHKD W/PUB EMPL AGNCY
		P1DO2_2	P1 EMQ070 PERS 2 CHKD W/PRIV EMP AGNCY
		P1DO3_2	P1 EMQ070 PERS 2 CHKD W/EMPLOYR DIRECTLY
		P1DO4_2	P1 EMQ070 PERS 2 CHKD W/FRIENDS & REL
		P1DO5_2	P1 EMQ070 PERS 2 PLACED OR ANSWERED ADS
		P1DO6_2	P1 EMQ070 PERS 2 READ WANT ADS
		P1DO7_2	P1 EMQ070 PERS 2 DID SOMETHING ELSE
		P1DOW_2	P1 EMQ080 WHAT PERSON 2 DOING LAST WEEK
		P1TAK_2	P1 EMQ100 PERS 2 JOB AVAILABLE LAST WEEK
		P1OCC_2	P1 EMQ130-50 2ND PERSON OCCUPATION CODE

7.4.6 Changes in Household Composition between Fall-Kindergarten and Spring-Kindergarten

Changes in household composition between fall-kindergarten and spring-kindergarten may complicate analyses. A household may have two different mothers in the base year file, one for each data collection point. For most composite variables, a change in parent across rounds will result in missing values in spring-kindergarten, since much of the information collected in fall-kindergarten is either not collected in spring-kindergarten or is only asked of new households in spring-kindergarten. Consequently, the household with a change in mother between rounds will have a nonmissing value for mother's employment in fall-kindergarten, but a missing value for mother's employment in spring-kindergarten.

For several composites—education, mother's marital status, and mother's work history—a cross-round, "base year" composite was constructed to compensate for missing values resulting from household changes. These composites have a prefix of "w." In general, if a nonmissing value existed for spring-kindergarten, the composite was assigned that value; otherwise, if a nonmissing value existed for fall-kindergarten, that value was used; otherwise the variable was set to missing. The result of this scheme maximizes available information across rounds, on the theory that "replacement" parents in spring-kindergarten do not differ much demographically from their fall-kindergarten counterparts.

Error in the Child Data File

The following error was identified in the ECLS-K Base Year Public-Use *Child* data file (child.dat). Five children in the spring-kindergarten had their parent interviews conducted under the wrong parent identification number. As a result the child records for these children have incorrect parent data, although they are correct for the other components (child data, non-parent weights, etc.). None of the five children have fall-kindergarten parent data. Two of these children should not have had any parent data since they were round 2 nonrespondents, while three of the cases had their parent data stored under other children's identification numbers. The IDs are:

	Case (Child ID)	Has parent data belonging too …	Should have parent data from…
1.	0162001C	3056007C	No parent data (R2 nonrespondent)
2.	0192001C	0192003C	No parent data (R2 nonrespondent)
3.	0192003C	0192015C	0192001C
4.	0192015C	None	0192003C
5.	3056007C	None	0162001C

There are two recommended strategies for correcting this error. Users may go ahead and use the cases since the effect on any analysis will be trivial. The other option would be to discard these five cases during analysis.

7.4.7 The Supplemental Section (SPQ) in the Spring-Kindergarten Parent Interview

The SPQ section was administered in spring-kindergarten to households that were nonrespondents in fall-kindergarten. The SPQ was constructed to allow respondents participating for the first time in spring-kindergarten to provide data for some of the more vital items from fall-kindergarten, rather than respond to the entire fall-kindergarten instrument. For example, an entire section of fall-kindergarten is devoted to education of parents and other adults in the household, while the spring-kindergarten SPQ contains only four questions on the education of the respondent and respondent's spouse (if applicable).

The topics covered in SPQ section include child care, Head Start participation, child health and health services received, marital status of the biological parents at the time of the child's birth, the primary language spoken in the home, respondent (and spouse) education, and receipt of WIC benefits.

Relevant items from the SPQ section were used to construct composites for families new to the survey in spring-kindergarten. Composites were calculated for child care (WKCAREPK), parental education (WKMOMED, WKDADED, WKPARED), primary language (WKLANGST), mother's work status (WKHEARLY), and parent's marital status at the time of the focal child's birth (WKHMOMAR, WKBMOMAR). These "cross-round" composites facilitate analyses across fall-kindergarten and spring-kindergarten.

When using SPQ data for cross-round analyses, users should combine the data items from SPQ with the appropriate items from fall-kindergarten. Table 7-4 presents instrument item matches across the two data collection points.

Care should be exercised when doing cross-round analysis of parental education, since the spring-kindergarten SPQ education items are asked only of the respondent and the respondent's spouse, and it is conceivable that the respondent is not the parent. The cross-round parental education composites (WKPARED, WKMOMED, and WKDADED) are more appropriate for such analyses.

The only section that is asked in its entirety to fall nonrespondents is the FSQ section, where information on the household roster is collected for the first time.

7.4.8 Industry and Occupation Codes Used in ECLS-K

This section describes the aggregated categories that were used for coding occupation in the ECLS-K.[3]

- **Executive, Administrative, and Managerial Occupations**

This category includes top and middle management occupations and occupations that directly support management. Top level managers are persons concerned with policy making, planning, staffing, directing, and/or controlling activities. Middle managers include persons who plan, or organize, direct and/or control activities at the operational level. Legislators are also included in this category. Workers in this category are not directly concerned with the fabrication of products or with the provision of services. Other officials and administrators include consultants, library directors, customer-house brokers, and location managers.

[3] Office of Budget and Management, Executive Office of the President (1980). *Standard Industrial Classification Manual*. Springfield, VA, and Office of Federal Statistical Policy and Standards, U.S. Department of Commerce (1980). *Standard Occupational Classification Manual* (2nd ed.). Washington, DC: Superintendent of Documents, U.S. Government Printing Office.

Table 7-4. SPQ item matches to fall-kindergarten items

Fall-kindergarten Variable	Fall-kindergarten Item	Spring-kindergarten Variable	Spring-kindergarten Item
P1SAMETW	CCQ003	P2SAMETW	SPQ005
P1RPREK	CCQ025	P2RPREK	SPQ010
P1NPREK	CCQ130	P2NPREK	SPQ020
P1HSPREK	CCQ215	P2HSPREK	SPQ030
P1CPREK	CCQ280	P2CPREK	SPQ040
P1WEIGHP	CHQ005	P2WEIGHP	SPQ060
P1WEIGHO	CHQ005	P2WEIGHO	SPQ065
P1WEIGH5	CHQ010	P2WEIGH5	SPQ070
P1WEIGH6	CHQ015	P2WEIGH6	SPQ080
P1PREMAT	CHQ025	P2PREMAT	SPQ090
P1EARLY	CHQ030	P2EARLY	SPQ100
P1EARDAY	CHQ030	P2EARDAY	SPQ105
P1SIGHT	CHQ285	P2SIGHT	SPQ107
P1CORREC	CHQ315	P2CORREC	SPQ108
P1DIFFHR	CHQ230	P2DIFFHR	SPQ109
P1THERAP	CHQ340	P2THERAP	SPQ110
P1THERA2	CHQ345A	P2THERA2	SPQ120A
P1THERA3	CHQ345B	P2THERA3	SPQ120B
P1THERA5	CHQ345D	P2THERA5	SPQ120D
P1THERA6	CHQ345E	P2THERA6	SPQ120E
P1THERA7	CHQ345F	P2THERA7	SPQ120F
P1THERA8	CHQ345G	P2THERA8	SPQ120G
P1THERA9	CHQ345H	P2THERA9	SPQ120H
P1THER10	CHQ345I	P2THER10	SPQ120I
P1THER11	CHQ345J	P2THER11	SPQ120J
P1THER12	CHQ345K	P2THER12	SPQ120K
P1THER13	CHQ345L	P2THER13	SPQ120L
P1THER14	CHQ375	P2THER14	SPQ130A
P1THER15	CHQ375U	P2THER15	SPQ130B
P1BIOLOG	MHQ060	P2BIOLOG	SPQ150
P1ANYLNG	PLQ020	P2ANYLNG	SPQ155
P1PRMLNG	PLQ060	P2PRMLNG	SPQ157
P1MOMWRK	EMQ170	P2MOMWRK	SPQ200
P1WICMOM	WPQ030	P2WICMOM	SPQ210
P1WICCHD	WPQ040	P2WICCHD	SPQ220

- **Natural Scientists and Mathematicians**

This category includes those engaged primarily in the application of scientific principles to research and development. Natural scientists are those in the physical sciences (e.g., chemistry, physics) and the life sciences (e.g., biology, agriculture, medicine). In addition, this category includes those in computer science, mathematics (including statistics), and operations research.

- **Social Scientists, Social Workers, Religious Workers, and Lawyers**

This division includes occupations concerned with the social needs of people and in basic and applied research in the social sciences.

- **Teachers: College, University, and other Postsecondary Institution; Counselors, Librarians, Archivists**

This NHES category includes those who teach at higher education institutions and at other postsecondary (after high school) institutions, such as vocational institutes. In addition, vocational and educational counselors, librarians, and archivists are included here.

- **Teachers, except Postsecondary Institution**

This category includes prekindergarten and kindergarten teachers, elementary and secondary teachers, special education teachers, instructional coordinators, and adult education teachers (outside postsecondary).

- **Health Diagnosing and Treating Practitioners**

This category includes health care professionals who diagnose and treat patients. In addition to physicians, dentists, and veterinarians, this category includes optometrists, podiatrists, and other diagnosing and treating professionals, such as optometrists, podiatrists, chiropractors, hypnotherapists, and acupuncturists.

- **Registered Nurses, Pharmacists, Dieticians, Therapists, and Physician's Assistants**

This category includes occupations concerned with the maintenance of health, the prevention of illness, and the care of the ill through the provision and supervision of nursing care; compounding drugs, planning food service or nutritional programs; providing assistance to physicians; and the provision of therapy and treatment as directed by physicians.

- **Writers, Artists, Entertainers, and Athletes**

This occupational category includes occupations concerned with creating and executing artistic works in a personally interpreted manner, by painting, sculpturing, drawing, engraving, etching, and other methods; creating designs for products and interior decorations; designing and illustrating books, magazines, and other publications; writing; still, motion picture and television photography/filming; producing, directing, staging, acting, dancing, singing in entertainment; and participating in sports and athletics as competitor or player and administering and directing athletic programs.

- **Health Technologists and Technicians**

This category includes occupations concerned with providing technical assistance in the provision of health care. For example, clinical laboratory technologists and technicians, dental hygienists, radiologic technicians, licensed practical nurses (LPN's), and other health technologists are included here.

- **Technologists and Technicians, except Health**

This group includes those providing technical assistance in engineering and scientific research, development, testing, and related activities, as well as operating and programming technical equipment and systems.

- **Marketing and Sales Occupations**

This category includes occupations involving selling goods or services, purchasing commodities and property for resale, and conducting wholesale or retail business.

- **Administrative Support Occupations, including Clerical**

Occupations involving preparing, transcribing, transferring, systematizing, and preserving written communications and records; collecting accounts; gathering and distributing information; operating office machines and data processing equipment; operating switchboards; distributing mail and messages; and other support and clerical duties such as bank teller, data entry keyer, etc.

- **Service Occupations**

The category includes occupations providing personal and protective services to individuals, and current maintenance and cleaning for building and residences. Some examples include food service, health service (e.g., aides or assistants), cleaning services other than household, and personal services.

- **Agricultural, Forestry, and Fishing Occupations**

This category is concerned with the production, propagation (breeding/growing), gathering, and catching of animals, animal products, and plant products (timber, crop, and ornamental); the provision of services associated with agricultural production; and game farms, fisheries, and wildlife conservation. "Other agricultural and related occupations" include occupations concerned with the production and propagation of animals, animals products, plants and products (crops and ornamental).

- **Mechanics and Repairers**

Mechanics and repairers are persons who do adjustment, maintenance, part replacement, and repair of tools, equipment, and machines. Installation may be included if installation is usually done in conjunction with other duties of the repairers.

- **Construction and Extractive Occupations**

This category includes occupations that normally are performed at a specific site, which will change over time, in contrast to production workers, where the work is usually at a fixed location. Construction workers include those in overall construction, brickmasons, stonemasons, carpenters, electricians, drywall installers, paperhangers and painters, etc. Extractive occupations include oil well drillers, mining machine operators, and so on.

- **Precision Production Occupations**

Precision production includes occupations concerned with performing production tasks that require a high degree of precision or attainment of rigid specification and operating plants or large systems. Examples are tool and die makers, pattern and model makers, machinists, jewelers, engravers, and so on. Also included are some food-related occupations including butchers and bakers. Plant and system operators include water and sewage, gas, power, chemical, petroleum, and other plant or system operators.

- **Production Working Occupations**

This category includes occupations concerned with setting up, operating, and tending of machines and hand production work usually in a factory or other fixed place of business.

- **Transportation and Material Moving Occupations**

This category includes occupations concerned with operating and controlling equipment used to facilitate the movement of people or materials and the supervising of those workers.

- **Handlers, Equipment Cleaners, Helpers, and Laborers**

This category includes occupations that involve helping other workers and performing routine nonmachine tasks. A wide variety of helpers, handlers, etc., are included in this category. Examples include construction laborers, freight, stock, and material movers, garage and service station related occupations, parking lot attendants, and vehicles washers and equipment cleaners.

7.5 Using Teacher Data Collected Either in Fall-kindergarten or Spring-kindergarten

Teachers were asked about classroom characteristics in teacher questionnaire part B (TQB). If the data were collected in fall-kindergarten, they were not collected again in spring-kindergarten. If they were not collected in fall-kindergarten, i.e., a teacher was added to the sample in spring-kindergarten, then they were collected in spring-kindergarten. There are flags on the file to indicate in which round the data were collected. The TQB flags and data are set on the teacher and child catalogs as follows:

- **Teacher catalog:** There are two flags: B1TQUEX and B2TQUEX. For each, a "1" means that the TQB for that teacher was collected at that data collection point. All teachers will have one but not both of these variables set. Each teacher record has only one set of TQB data, regardless of data collection point.

- **Child catalog:** There are two flags and corresponding sets of data. The first flag, B1TQUEX, and variables (approximately 80 variables all starting with the B1 prefix) are set if the teacher with whom the child was linked in fall-kindergarten has a completed TQB. This is regardless of which round the TQB on the teacher record was collected. The second flag, B2TQUEX, and variables (also approx. 80 variables, all starting with B2) are set if the teacher with whom the child was linked in spring-kindergarten has a TQB. That is, the B1TQUEX and B2TQUEX on the child catalog refer to the data collection point in which the child was linked to the teacher, not the data collection point in which that teacher's TQB data was collected.

7.6 The Relationship Among Composite Variables on the Child, Teacher, and School Catalogs

The school catalog contains ONLY school composites.

The teacher catalog contains all teacher composites and five of the school composites (S2KSCTYP, S2KENRLS, S2KMINOR, S2KSCLVL, and S2KPUPR).

The child catalog contains all child, teacher (except KGCLASS which is in the teacher catalog only), and school composites, but it should be noted that some of the teacher variables on the child catalog are slightly altered when they are brought to the child level. While some teacher variables are class specific, i.e., there will be three versions of each of these class-specific variables (AM, PM, and all day) on the teacher file, the child file will only contain data from only one of each of the three variables—depending on whether the child attends an AM, PM, or all day class. For example, there are three variables (one for AM, one for PM, and one for all day classes) on the teacher catalog that represent percent of African-Americans in the teacher's class. Since the child can only be in one class, the child-file version of percent African-American in class is only one variable. The same is true for percent Hispanic and percent minority variables. The opposite is true for the teacher age variable—since a child could have two teachers, one in each data collection point, there are two teacher ages (B1AGE, B2AGE) on the child file, but only one teacher age (B1AGE) on the teacher file. Table 7-5 presents the relationship of the composite variables on the three catalogs.

Table 7-5. Relationship among composite variables on the three ECB catalogs

School Catalog	Teacher Catalog	Child Catalog	Variable Name	Variable Label
		X	DOBMM	CHILD COMPOSITE DOB MONTH
		X	DOBDD	CHILD COMPOSITE DOB DAY
		X	DOBYY	CHILD COMPOSITE DOB YEAR
		X	GENDER	CHILD COMPOSITE GENDER
		X	RACE	CHILD COMPOSITE RACE
		X	R1_KAGE	R1 COMPOSITE CHILD ASSESSMENT AGE(MNTHS)
		X	R2_KAGE	R2 COMPOSITE CHILD ASSESSMENT AGE(MNTHS)
		X	R2ELIG	R2 CHILD ELIGIBILITY IN ROUND 2
		X	FKCHGTCH	FK CHD CHANGED TCHRS BETWEEN ROUND 1 & 2
		X	FKCHGSCH	FK CHD CHANGED SCHLS BETWEEN ROUND 1 & 2
		X	C1HEIGHT	C1 ROUND 1 CHILD COMPOSITE HGT (INCHES)
		X	C1WEIGHT	C1 ROUND 1 CHILD COMPOSITE WGT (POUNDS)
		X	C1BMI	C1 ROUND 1 CHILD COMPOSITE BMI
		X	C2HEIGHT	C2 ROUND 2 CHILD COMPOSITE HGT (INCHES)
		X	C2WEIGHT	C2 ROUND 2 CHILD COMPOSITE WGT (POUNDS)
		X	C2BMI	C2 ROUND 2 CHILD COMPOSITE BMI
		X	P1HMEMP	P1 CURRENT MOTHER EMPLOYMENT STATUS
		X	P1HDEMP	P1 CURRENT FATHER EMPLOYMENT STATUS
		X	P1HMAGE	P1 AGE - CURRENT MOTHER (YRS)
		X	P1HDAGE	P1 AGE - CURRENT FATHER (YRS)
		X	P1DADID	P1 HOUSEHOLD ROSTER NUMBER OF FATHER
		X	P1MOMID	P1 HOUSEHOLD ROSTER NUMBER OF MOTHER
		X	P1HMRACE	P1 RACE OF CURRENT MOTHER
		X	P1HMHISP	P1 MOTHER HISPANIC
		X	P1HMIND	P1 MOTHER AMERICAN INDIAN
		X	P1HMASN	P1 MOTHER ASIAN
		X	P1HMBLCK	P1 MOTHER BLACK
		X	P1HMPACI	P1 MOTHER PACIFIC ISLANDER
		X	P1HMWHT	P1 MOTHER WHITE
		X	P1HMMT1R	P1 MOTHER MORE THAN ONE RACE
		X	P1HDRACE	P1 RACE OF CURRENT FATHER
		X	P1HDHISP	P1 FATHER HISPANIC
		X	P1HDIND	P1 FATHER AMERICAN INDIAN
		X	P1HDASN	P1 FATHER ASIAN
		X	P1HDBLCK	P1 FATHER BLACK
		X	P1HDPACI	P1 FATHER PACIFIC ISLANDER
		X	P1HDWHT	P1 FATHER WHITE
		X	P1HDMT1R	P1 FATHER MORE THAN ONE RACE
		X	P1HMAFB	P1 AGE AT 1ST BIRTH - CURRENT MOM (YRS)
		X	P1BMAGE	P1 AGE OF NONRES BIO MOTHER (YRS)
		X	P1BDAGE	P1 AGE OF NONRES BIO FATHER (YRS)
		X	P1BMRACE	P1 RACE OF NONRES BIOLOGICAL MOTHER
		X	P1BMHISP	P1 NONRES BIOLOGICAL MOTHER HISPANIC

Table 7-5. Relationship among composite variables on the three ECB catalogs (continued)

School Catalog	Teacher Catalog	Child Catalog	Variable Name	Variable Label
		X	P1BMIND	P1 NONRES BIO MOTHER AMERICAN INDIAN
		X	P1BMASN	P1 NONRES BIOLOGICAL MOTHER ASIAN
		X	P1BMBLCK	P1 NONRES BIOLOGICAL MOTHER BLACK
		X	P1BMPACI	P1 NONRES BIO MOTHER PACIFIC ISLANDER
		X	P1BMWHT	P1 NONRES BIOLOGICAL MOTHER WHITE
		X	P1BDMT1R	P1 NONRES BIO FATHER MORE THAN 1 RACE
		X	P1BDRACE	P1 RACE OF NONRES BIOLOGICAL FATHER
		X	P1BDHISP	P1 NONRES BIOLOGICAL FATHER HISPANIC
		X	P1BDIND	P1 NONRES BIO FATHER AMERICAN INDIAN
		X	P1BDASN	P1 NONRES BIOLOGICAL FATHER ASIAN
		X	P1BDBLCK	P1 NONRES BIOLOGICAL FATHER BLACK
		X	P1BDPACI	P1 NONRES BIO FATHER PACIFIC ISLANDER
		X	P1BDWHT	P1 NONRES BIOLOGICAL FATHER WHITE
		X	P1BMMT1R	P1 NONRES BIO MOTHER MORE THAN 1 RACE
		X	P1BMAFB	P1 AGE AT 1ST BIRTH NONRES BIO MOM (YRS)
		X	P1AMOMED	P1 EDUCATION LVL NONRES ADOPTIVE MOTHER
		X	P1ADADED	P1 EDUCATION LVL NONRES ADOPTIVE FATHER
		X	P1BMOMED	P1 EDUCATION LVL NONRES BIOLOGICAL MOM
		X	P1BDADED	P1 EDUCATION LVL NONRES BIOLOGICAL DAD
		X	P1MOMOCC	P1 RESIDENT MOTHER’S OCCUPATION
		X	P1DADOCC	P1 RESIDENT FATHER’S OCCUPATION
		X	P1AGEENT	P1 AGE (MONTHS) AT KINDERGARTEN ENTRY
		X	P1AGEFRS	P1 AGE (MNTHS) AT FIRST NONPARENTAL CARE
		X	P1CARNOW	P1 CURRENT NONPARENTAL CARE ARRANGEMENTS
		X	P1NUMNOW	P1 # NONPARENTAL CARE ARRANGEMENTS NOW
		X	P1HRSNOW	P1 # HOURS SPENT IN NONPARENTAL CARE NOW
		X	P1PRIMNW	P1 PRIMARY TYPE OF NONPARENTAL CARE
		X	P1CENTER	P1 CHILD EVER IN CENTER-BASED CARE
		X	P1DISABL	P1 CHILD W/ DISABILITY
		X	P1HMLANG	P1 MOTHER’S LANGUAGE TO CHILD
		X	P1HMLTOM	P1 CHILD’S LANGUAGE TO MOTHER
		X	P1HDLANG	P1 FATHER’S LANGUAGE TO CHILD
		X	P1HDLTOD	P1 CHILD’S LANGUAGE TO FATHER
		X	P1LANGUG	P1 BOTH PARENT LANGUAGE TO CHILD
		X	P1HMOM	P1 RESIDENT MOTHER TYPE
		X	P1HDAD	P1 RESIDENT FATHER TYPE
		X	P1NUMSIB	P1 NUMBER OF SIBLINGS IN HOUSEHOLD
		X	P1LESS18	P1 NUMBER IN HOUSEHOLD AGED <18
		X	P1OVER18	P1 NUMBER IN HOUSEHOLD AGED 18+
		X	P1HTOTAL	P1 TOTAL NUMBER IN HOUSEHOLD
		X	P1HFAMIL	P1 FAMILY TYPE
		X	P1HPARNT	P1 TYPES OF PARENTS IN HOUSEHOLD
		X	P1MOMTYP	P1 TYPE OF RESIDENT MOTHER
		X	P1BMVITL	P1 BIOLOGICAL MOTHER IS LIVING/DEAD
		X	P1ABSMOM	P1 TYPE OF NONRESIDENT MOTHER(S)

Table 7-5. Relationship among composite variables on the three ECB catalogs (continued)

School Catalog	Teacher Catalog	Child Catalog	Variable Name	Variable Label
		X	P1DADTYP	P1 TYPE OF RESIDENT FATHER
		X	P1BDVITL	P1 BIOLOGICAL FATHER IS LIVING/DEAD
		X	P1ABSDAD	P1 TYPE OF NONRESIDENT FATHER(S)
		X	P1TWINST	P1 CHILD BIRTH STATUS
		X	P1FIRKDG	P1 FIRST-TIME KINDERGARTNER
		X	P2HMOM	P2 RESIDENT MOTHER TYPE
		X	P2HDAD	P2 RESIDENT FATHER TYPE
		X	P2NUMSIB	P2 NUMBER OF SIBLINGS IN HOUSEHOLD
		X	P2LESS18	P2 NUMBER IN HOUSEHOLD AGED <18
		X	P2OVER18	P2 NUMBER IN HOUSEHOLD AGED 18+
		X	P2HTOTAL	P2 TOTAL NUMBER IN HOUSEHOLD
		X	P2HFAMIL	P2 FAMILY TYPE
		X	P2HPARNT	P2 TYPES OF PARENTS IN HOUSEHOLD
		X	P2MARSTA	P2 MARITAL STATUS
		X	P2MOMID	P2 HOUSEHOLD ROSTER NUMBER OF MOTHER
		X	P2DADID	P2 HOUSEHOLD ROSTER NUMBER OF FATHER
		X	P2HMAGE	P2 AGE - CURRENT MOTHER (YRS)
		X	P2HDAGE	P2 AGE - CURRENT FATHER (YRS)
		X	P2HMRACE	P2 RACE OF CURRENT MOTHER
		X	P2HMHISP	P2 MOTHER HISPANIC
		X	P2HMIND	P2 MOTHER AMERICAN INDIAN
		X	P2HMASN	P2 MOTHER ASIAN
		X	P2HMBLCK	P2 MOTHER BLACK
		X	P2HMPACI	P2 MOTHER PACIFIC ISLANDER
		X	P2HMWHT	P2 MOTHER WHITE
		X	P2HMMT1R	P2 MOTHER MORE THAN ONE RACE
		X	P2HDRACE	P2 RACE OF CURRENT FATHER
		X	P2HDHISP	P2 FATHER HISPANIC
		X	P2HDIND	P2 FATHER AMERICAN INDIAN
		X	P2HDASN	P2 FATHER ASIAN
		X	P2HDBLCK	P2 FATHER BLACK
		X	P2HDPACI	P2 FATHER PACIFIC ISLANDER
		X	P2HDWHT	P2 FATHER WHITE
		X	P2HDMT1R	P2 FATHER MORE THAN ONE RACE
		X	P2DADCHK	P2 FATHER IN HOUSEHOLD
		X	P2DADTYP	P2 TYPE OF RESIDENT FATHER
		X	P2MOMCHK	P2 MOTHER IN HOUSEHOLD
		X	P2MOMTYP	P2 TYPE OF RESIDENT MOTHER
		X	WKMOMED	WK MOTHER'S EDUCATION LEVEL
		X	IF_M_ED	IF MOTHER'S EDUCATION IMPUTED
		X	WKDADED	WK FATHER'S EDUCATION LEVEL
		X	IF_D_ED	IF FATHER'S EDUCATION IMPUTED
		X	WKHEARLY	WK MOM WORKED BTWN CHILD'S BIRTH & KINDG
		X	WKPARED	WK PARENT HIGHEST EDUCATION LEVEL
		X	WKHMOMAR	WK MOTHER MARRIED AT TIME OF BIRTH
		X	WKBMOMAR	WK NONRES MOM MARRIED AT TIME OF BIRTH
		X	WKRACETH	WK CHILD RACE- COMPOSITE

Table 7-5. Relationship among composite variables on the three ECB catalogs (continued)

School Catalog	Teacher Catalog	Child Catalog	Variable Name	Variable Label
		X	WKHISP	WK CHILD HISPANIC
		X	WKAMERIN	WK CHILD AMERICAN INDIAN
		X	WKASIAN	WK CHILD ASIAN
		X	WKBLACK	WK CHILD BLACK
		X	WKPACISL	WK CHILD PACIFIC ISLANDER
		X	WKWHITE	WK CHILD WHITE
		X	WKMT1RAC	WK CHILD MORE THAN ONE RACE
		X	WKCAREPK	WK NONPARENTAL CARE ARRANGEMENTS PRE-K
		X	WKNUMPRK	WK # NONPARENTAL CARE ARRANGEMENTS PRE-K
		X	P1HRSPRK	P1 # HRS SPENT IN NONPARENTAL CARE PRE-K
		X	P1PRIMPK	P1 PRIMARY TYPE NONPARENTAL CARE PRE-K
		X	P1COSTPK	P1 COST NONPARENTAL CHILD CARE PRE-K ($)
		X	WKLANGST	WK HOME LANGUAGE OF CHILD
		X	WKMOMSCR	WK AVG PRESTIGE SCORE FOR MOM OCCUPATION
		X	IF_M_OCS	IF MOM OCCUPATION PRESTIGE SCORE IMPUTED
		X	WKDADSCR	WK AVG PRESTIGE SCORE FOR DAD OCCUPATION
		X	IF_D_OCS	IF DAD OCCUPATION PRESTIGE SCORE IMPUTED
		X	WKSESL	WK CONTINUOUS SES MEASURE
		X	WKSESQ5	WK CATEGORICAL SES MEASURE
		X	WKPOVRTY	WK POVERTY LEVEL
		X	WKINCOME	WK INCOME (IMPUTED)
		X	IF_INC	IF INCOME VALUE IMPUTED
		X	A1PBLK	A1 PERCENT OF BLACKS IN CLASS
	X		A1APBLK	A1 PERCENT OF BLACKS IN CLASS-AM
	X		A1PPBLK	A1 PERCENT OF BLACKS IN CLASS-PM
	X		A1DPBLK	A1 PERCENT OF BLACKS IN CLASS-AD
		X	A1PHIS	A1 PERCENT OF HISPANICS IN CLASS
	X		A1APHIS	A1 PERCENT OF HISPANICS IN CLASS-AM
	X		A1PPHIS	A1 PERCENT OF HISPANICS IN CLASS-PM
	X		A1DPHIS	A1 PERCENT OF HISPANICS IN CLASS-AD
		X	A1PMIN	A1 PERCENT OF MINORITIES IN CLASS
	X		A1APMIN	A1 PERCENT OF MINORITIES IN CLASS-AM
	X		A1PPMIN	A1 PERCENT OF MINORITIES IN CLASS-PM
	X		A1DPMIN	A1 PERCENT OF MINORITIES IN CLASS-AD
	X	X	B1AGE	B1 TEACHER’S AGE
		X	B2AGE	B2 TEACHER’S AGE
	X		KGCLASS	KINDERGARTEN TEACHER CLASS COMPOSITE
X	X	X	S2KSCTYP	S2 SCHOOL TYPE FROM THE SCH ADMIN QUEST
X		X	S2KENRLK	S2 TOTAL SCHOOL K ENROLLMENT
X	X	X	S2KENRLS	S2 TOTAL SCHOOL ENROLLMENT
X	X	X	S2KMINOR	S2 PERCENT MINORITY STUDENTS
X		X	S2KFLNCH	S2 PCT FREE LUNCH ELIGIBLE STUDENTS
X		X	S2KRLNCH	S2 PCT REDUCED LUNCH ELIGIBLE STUDENTS
X	X	X	S2KSCLVL	S2 SCH INSTRUCTNL LEVEL FROM ADMIN QUEST
X		X	S2KLNGTH	S2 SCHOOL YEAR LENGTH
X		X	S2KGFTED	S2 PERCENT GIFTED/TALENTED STUDENTS
X	X	X	S2KPUPRI	S2 PUBLIC OR PRIVATE SCHOOL

7.7 Creating a Child File for Classroom Level Analysis

Using T1_ID or T2_ID in the child catalog a user can identify all students belonging to a given teacher. And given a T1_ID and A1CLASS or T2_ID and A2CLASS combination, a classroom level file can also be created in the child catalog. For example, in the child catalog, records with T1_ID = 0001T01 and A1CLASS = AM are all children from teacher 0001T01's morning class.

7.8 Creating a Classroom Level File Using Teacher Data

For researchers interested in doing classroom level analysis on the characteristics of kindergarten classrooms, a classroom level file can be constructed. Each record in the teacher file represents one teacher regardless of whether that teacher teaches one or two classes. (Approximately 16 percent of the kindergarten teachers in the ECLS-K sample teach both a morning and an afternoon class.) For situations when the *class* rather than the *teacher* is the unit of analysis of interest to the researcher, a new class-level file may be constructed so that each record represents a kindergarten class. Teachers provided separate information about each of their kindergarten classes – morning, afternoon, or all-day – for a set of items on the fall part A and spring part A teacher questionnaires. These class-specific variables contain information about class characteristics such as class size and race-ethnicity composition. In order to create a classroom level file a user must perform the following steps:

1. Identify the variables that are to be included in the classroom level file including the AM, PM, and all day versions of identical variables. These are the variables that will comprise the classroom file, so non-class specific data, e.g., teacher's educational background, should be included as well. A1AASIAN, A1PASIAN, and A1DASIAN are examples of class-specific variables.

2. Create three files, an AM, PM, and all day file, each with the appropriate class specific variables (e.g., keep only A1AASIAN for the AM file, drop A1PASIAN and A1DASIAN) and the remaining non-class specific variables. The aim of this exercise is to isolate into three groups those records with data belonging to each of the three groups. Those teachers with more than one class type will appear in more than one file. Thus on the AM file, there will be data for all the non-class specific variables, and the AM variables. The PM and all day variables will be dropped.

3. Rename the class-specific variables in each of the three files to be the same across each of the files. For example, on the AM file, one might rename the variable A1AASIAN (the number of Asian children in the AM class) to A1ASIAN. Similarly, A1PASIAN (the number of Asian children in the PM class) gets renamed to A1ASIAN, and A1DASIAN gets renamed to A1ASIAN in the all day file.

4. On each of these three files, create a class type variable with the same name across the three files, that indicates the class type. The class type variable can use values of 1=AM, 2=PM, and 3=all day. In each of the files the class type variable value will be the same for all cases (e.g., all records in the AM file will have a value of ‘AM’ for class type), but when the records from the three files (AM, PM, and all day) are merged, the class type variable will be used to identify whether the classroom is an AM, PM, or all day class.

5. Merge the three files. Now each case represents one classroom. This file will have no classroom ID. An ID can be created if needed but is not necessary for analysis.

With this file users can do analysis such as examining characteristics of classrooms with high percent minority, etc.

KGCLASS, F1CLASS, F2CLASS, A1CLASS, and A2CLASS are all class type variables—AM/PM/All-day. However, they carry different information and serve different purposes. KGCLASS was created to reflect the teacher’s base year class type and is not related to the child data. (See section 7.4 for details.) F1CLASS and F2CLASS are the child’s fall- and spring-kindergarten class types respectively in the FMS. A1CLASS and A2CLASS were created for linking teacher and child data in fall- and spring-kindergarten, respectively.

The teacher report of class type in the teacher questionnaire A defines the teacher’s reported class type. This report does not always match the child’s class type. A1CLASS and A2CLASS were created to report a child class type. In fall-kindergarten, if the teacher’s reported class type in teacher questionnaire A matches F1CLASS, then A1CLASS is defined as the matching class type. If the teacher’s reported class type does not match F1CLASS and the teacher reported only one class in teacher questionnaire A, then A1CLASS is defined as the teacher’s reported class type and the child is linked to the only class that the teacher reported. If the teacher’s reported class type does not match F1CLASS and the teacher reported more than one class in questionnaire A, then A1CLASS is missing and the teacher and child are not linked. The same rules were followed with F2CLASS and A2CLASS with spring-kindergarten data to link the teacher and child data.

Table 7-6 lays out the composite and derived variables that are on the ECLS-K child, teacher, and school datasets. Please note that some of the variables specified in the ‘derived from’ column are intermediary variables that were not included in the final datasets. This column also contains the item numbers from the questionnaire, which help in identifying the items that were used in the creation of these composites and derived variables.

Table 7-6. Composite variables

Variable Name	Category	Description	Derived from	Values
CHILDID	Child	Child Identification Number	OCHILDID	
R1_KAGE R2_KAGE	Child	Child's age at the time the direct child assessment occurred. This age will be in month and years, no days.	DOBMO, DOBDA, DOBYR, C1ASMTMM, C1ASMTDD, C1ASMTYY, C2ASMTMM, C2ASMTDD, C2ASMTYY	Continuous
GENDER	Child	Child's gender	FMS and parent interview fall (INQ050) spring (INQ 160)	1 = Male 2 = Female
DOBMM DOBDD DOBYY	Child Child Child	Child's date of birth month Child's date of birth day Child's date of birth year	FMS, parent interview fall (INQ060) spring (INQ 170)	1-12 1-31
WKRACETH	Child	Race and ethnicity of the child	P1RC_1 (FSQ190), P1RC_2 (FSQ190), P1RC_3 (FSQ190), P1RC_4 (FSQ190), P1RC_5 (FSQ190), P1RC_6 (FSQ190), RACEOS (FSQ190), P1HSP (FSQ190), P2RC_1 (FSQ195), P2RC_2 (FSQ195), P2RC_3 (FSQ195), P2RC_4 (FSQ195), P2RC_5 (FSQ195), RACEOS (FSQ198), P2HSP (FSQ190)	1 = White, non-Hispanic 2 = Black or African American, non-Hispanic 3 = Hispanic, race specified 4 = Hispanic, no race specified 5 = Asian 6 = Native Hawaiian or other Pacific Islander 7 = American Indian or Alaskan Native 8 = More than 1 race, non-Hispanic

Table 7-6. Composite variables (continued)

Variable Name	Category	Description	Derived from	Values
WKAMERIN	Child	Child is American Indian or Native Alaskan	P1RC_1 (FSQ190), P1RC_2 (FSQ190), P1RC_3 (FSQ190), P1RC_4 (FSQ190), P1RC_5 (FSQ190), P1RC_6 (FSQ190), RACEOS (FSQ190), P1HSP (FSQ190), P2RC_1 (FSQ195), P2RC_2 (FSQ195), P2RC_3 (FSQ195), P2RC_4 (FSQ195), P2RC_5 (FSQ195), RACEOS (FSQ198), P2HSP (FSQ190)	1 = Yes 2 = No
WKASIAN	Child	Child is Asian	P1RC_1 (FSQ190), P1RC_2 (FSQ190), P1RC_3 (FSQ190), P1RC_4 (FSQ190), P1RC_5 (FSQ190), P1RC_6 (FSQ190), RACEOS (FSQ190), P1HSP (FSQ190), P2RC_1 (FSQ195), P2RC_2 (FSQ195), P2RC_3 (FSQ195), P2RC_4 (FSQ195), P2RC_5 (FSQ195), RACEOS (FSQ198), P2HSP (FSQ190)	1 = Yes 2 = No

Table 7-6. Composite variables (continued)

Variable Name	Category	Description	Derived from	Values
WKPACISL	Child	Child is Native Hawaiian or other Pacific Islander	P1RC_1 (FSQ190), P1RC_2 (FSQ190), P1RC_3 (FSQ190), P1RC_4 (FSQ190), P1RC_5 (FSQ190), P1RC_6 (FSQ190), RACEOS (FSQ190), P1HSP (FSQ190), P2RC_1 (FSQ195), P2RC_2 (FSQ195), P2RC_3 (FSQ195), P2RC_4 (FSQ195), P2RC_5 (FSQ195), RACEOS (FSQ198), P2HSP (FSQ190)	1 = Yes 2 = No
WKBLACK	Child	Child is African American	P1RC_1 (FSQ190), P1RC_2 (FSQ190), P1RC_3 (FSQ190), P1RC_4 (FSQ190), P1RC_5 (FSQ190), P1RC_6 (FSQ190), RACEOS (FSQ190), P1HSP (FSQ190), P2RC_1 (FSQ195), P2RC_2 (FSQ195), P2RC_3 (FSQ195), P2RC_4 (FSQ195), P2RC_5 (FSQ195), RACEOS (FSQ198), P2HSP (FSQ190)	1 = Yes 2 = No

Table 7-6. Composite variables (continued)

Variable Name	Category	Description	Derived from	Values
WKWHITE	Child	Child is white	P1RC_1 (FSQ190), P1RC_2 (FSQ190), P1RC_3 (FSQ190), P1RC_4 (FSQ190), P1RC_5 (FSQ190), P1RC_6 (FSQ190), RACEOS (FSQ190), P1HSP (FSQ190), P2RC_1 (FSQ195), P2RC_2 (FSQ195), P2RC_3 (FSQ195), P2RC_4 (FSQ195), P2RC_5 (FSQ195), RACEOS (FSQ198), P2HSP (FSQ190)	1 = Yes 2 = No
WKHISP	Child	Child is Hispanic	P1RC_1 (FSQ190), P1RC_2 (FSQ190), P1RC_3 (FSQ190), P1RC_4 (FSQ190), P1RC_5 (FSQ190), P1RC_6 (FSQ190), RACEOS (FSQ190), P1HSP (FSQ190), P2RC_1 (FSQ195), P2RC_2 (FSQ195), P2RC_3 (FSQ195), P2RC_4 (FSQ195), P2RC_5 (FSQ195), RACEOS (FSQ198), P2HSP (FSQ190)	1 = Yes 2 = No

Table 7-6. Composite variables (continued)

Variable Name	Category	Description	Derived from	Values
WKMT1RAC	Child	Child is more than one race	P1RC_1 (FSQ190), P1RC_2 (FSQ190), P1RC_3 (FSQ190), P1RC_4 (FSQ190), P1RC_5 (FSQ190), P1RC_6 (FSQ190), RACEOS (FSQ190), P1HSP (FSQ190), P2RC_1 (FSQ195), P2RC_2 (FSQ195), P2RC_3 (FSQ195), P2RC_4 (FSQ195), P2RC_5 (FSQ195), RACEOS (FSQ198), P2HSP (FSQ190)	1 = Yes 2 = No
C1BMI C2BMI	Child Child	Round 1 body mass index (BMI) Round 2 body mass index (BMI)	C1HEIGHT, C1WEIGHT C2HEIGHT, C2WEIGHT	Continuous
C1HEIGHT	Child	Child's round 1 composite height	HEIGHTIN, HEIGHTI2	Continuous
C2HEIGHT		Child's round 2 composite height	HEIGHTI3, HEIGHTI4	Continuous
C1WEIGHT	Child	Child's round 1 composite weight	WEIGHTIN, WEIGHTI2	Continuous
C2WEIGHT		Child's round 2 composite weight	WEIGHTI3, WEIGHTI4	Continuous
P1DISABL	Child	Child currently has a disability	P1DIAGNO (CHQ120), P1PROFFD (CHQ155), P1CLIMB (CHQ185), P1COMMU2 (CHQ215), P1DIFFH3 (CHQ245), P1VISIO2 (CHQ300), P1THERAP (CHQ340)	1 = Yes 2 = No
P1TWINST	Child	Multiple birth status	P1MULTIP (CHQ035), P1IDENTI (CHQ070)	1 = Fraternal twin, same sex 2 = Fraternal twin, different sex 3 = Identical twin 4 = Higher order multiple birth (triplet, quadruplet, etc.) 9 = Single birth

Table 7-6. Composite variables (continued)

Variable Name	Category	Description	Derived from	Values
P1AGEFRS	Child	The age in months when the focal child first began any type of nonparental care on a regular basis	P1AGEMRL (CCQ020), P1YRSFIR (CCQ125), P1AGEMHS (CCQ259), P1CAGEMO (CCQ275)	Continuous
P1CARNOW	Child	Focal child is currently receiving any nonparental care	P1RELAT (CCQ010), P1NONREL (CCQ115), P1CTRNOW (CCQ260)	1 = Yes 2 = No
WKCAREPK	Child	Focal child received any nonparental care the year before entering kindergarten	P1RLCAR2 (CCQ070), P1NONRL2 (CCQ.120), P1HEADST (CCQ215), P1CPREK (CCQ.280), P2RPREK (SPQ010), P2NPREK (SPQ020), P2HSPREK (SPQ030), P2CPREK (SP040)	1 = Yes 2 = No
P1CENTER	Child	Child has ever been in center-based care	P1EVERHD (CCQ210), P1HEADST (CCQ215), P1CTRNOW (CCQ260), P1ATTEND (CCQ265)	1 = Yes 2 = No
P1COSTPK	Child	The total amount paid for all primary care arrangements for the child the year before kindergarten	P1AMTREL (CCQ053), P1PAYNON (CCQ158), P1AMTHS (CCQ256), P1AMNTF4 (CCQ318)	Continuous
P1HRSNOW	Child	Total number of hours per week the focal child currently spends in all primary, nonparental/non-kindergarten care	P1NUMHR2 (CCQ090), P1HRSNR (CCQ190), P1NUMHRS (CCQ335)	Continuous
P1HRSPRK	Child	Total number of hours per week the focal child spent in all types of primary, nonparental care the year before kindergarten	P1NUMHRS (CCQ335), P1HRSNON (CCQ150), P1HSHRS (CCQ251), P1NUMHR4 (CCQ310)	Continuous
P1NUMNOW	Child	Total number of all types of care arrangements the focal child currently has on a regular basis	P1NUMREG (CCQ060), P1NUMCAR (CCQ165), P1NUMNO2 (CCQ325)	Continuous

Table 7-6. Composite variables (continued)

Variable Name	Category	Description	Derived from	Values
WKNUMPRK	Child Care	Total number of all types of care arrangements the focal child had on a regular basis the year before starting kindergarten (Note: Information for this variable was only collected in round 1. Data for round 2 participants are coded as “not ascertained” or -9.)	P1NUMARR (CCQ.030), P1NUMNON (CCQ.135), P1HEADST (CCQ.215), P1NUMCNT (CCQ.285),	Continuous
P1PRIMPK	Child	Primary, nonparental arrangement in which the child spent the most hours per week during the year before kindergarten	P1NHRSPK (CCQ150), P1NUMHR4 (CCQ310), P1HSHRS (CCQ251), P1HOMREL (CCQ035)	0 = No non-parental care 1 = Relative care in child’s home 2 = Relative care in another home 3 = Non-relative care in child’s home 4 = Non-relative care in another home 5 = Head Start program 6 = Center-based program 7 = 2 or more programs 8 = Location of care varies
P1PRIMNW	Child	Primary, nonparental/non-kindergarten arrangement in which the child currently spends the most hours per week	P1NUMHR2 (CCQ090), P1RLCAR2 (CCQ070), P1HOMNON (CCQ140)	0 = No non-parental care 1 = Relative care in child’s home 2 = Relative care in another home 3 = Non-relative care in child’s home 4 = Non-relative care in another home 5 = Center-based program 6 = 2 or more programs 7 = Location of care varies
P1AGEENT	Child	The age in months at which the focal child first entered kindergarten. For children in their first year of kindergarten, this will be their age as of September 1, 1998; for those in their second year of kindergarten, it will be their age as of September 1, 1997.	CHILDDOB (INQ.060), P1WHICHY (PIQ080)	Continuous
P1FIRKDG	Child	First time kindergartner	P1WHICHY (PIQ080)	1 = Yes 2 = No

Table 7-6. Composite variables (continued)

Variable Name	Category	Description	Derived from	Values
A1CLASS	Child	Student attend a morning, afternoon or full-day kindergarten class in the fall	FMS data, A1ACLASS, A1DCLASS, A1PCLASS	1 = Morning class 2 = Afternoon class 3 = Full-day class
A2CLASS	Child	Student attends a morning, afternoon, or full-day kindergarten class in the spring	FMS data, A2ACLASS, A2DCLASS, A2PCLASS	1 = Morning class 2 = Afternoon class 3 = Full-day class
FKCHGTCH	Child	Child changed teachers between round 1 and round 2	R1T_ID, R2T_ID	0 = No change 1 = Changed teachers
FKCHGSCH	Child	Child changed schools between round 1 and round 2	R1S_ID, R2S_ID	0 = No change 1 = Changed schools
P1MOMID P2MOMID	Family/HH	Household roster number of resident mother	P1REL (FSQ130), P1UNR (FSQ180), P2REL (FSQ130), P2UNR (FSQ180)	1-17
P1DADID P2DADID	Family/HH	Household roster number of resident father	P1REL (FSQ130), P1UNR (FSQ180), P2REL (FSQ130), P2UNR (FSQ180)	1-17
P1HPARNT P2HPARNT	Family/HH	Classification of the focal child's parents who reside in the household	P1REL (FSQ130), P1UNR (FSQ180) P1HMOM, P1HDAD (composites), P2REL (FSQ130), P2UNR (FSQ180), P2HMOM, P2HDAD (composites)	1 = Biological mother and biological father 2 = Biological mother and other father (step-, adoptive, foster) 3 = Biological father and other mother (step-, adoptive, foster) 4 = Biological mother only 5 = Biological father only 6 = Two adoptive parents 7 = Single adoptive parent or adoptive parent and stepparent 8 = Related guardian(s) 9 = Unrelated guardian(s)
P1HFAMIL P2HFAMIL	Family/HH	Family type categories using both parent and sibling information	P1REL (FSQ130), P1UNR (FSQ180) P1HMOM, P1HDAD, P1NUMSIB (composites) P2REL (FSQ130), P2UNR (FSQ180), P2HMOM, P2HDAD, P2NUMSIB (composites)	1 = Two parents and sibling(s) 2 = Two parents, no siblings 3 = One parent and sibling(s) 4 = One parent, no siblings 5 = Other

Table 7-6. Composite variables (continued)

Variable Name	Category	Description	Derived from	Values
P1NUMSIB P2NUMSIB	Family/HH	Total number of siblings with whom the focal child lives, including anyone reporting him/herself as the child of the focal child's foster parent/guardian	P1REL (FSQ130), P1UNR (FSQ180), P1AGE, P2REL (FSQ130), P2UNR (FSQ180), P2AGE	Continuous
P1LESS18 P2LESS18	Family/HH	Total number of household members younger than 18 years old	P1REL (FSQ130), P1UNR (FSQ180), P1AGE, P2REL (FSQ130), P2UNR (FSQ180), P2AGE	Continuous
P1OVER18 P2OVER18	Family/HH	Total number of household members older than 18 years old	P1REL (FSQ130), P1UNR (FSQ180), P1AGE, P2REL (FSQ130), P2UNR (FSQ180), P2AGE	Continuous
P1HTOTAL P2HTOTAL	Family/HH	Total number of household members	P1REL (FSQ130), P1UNR (FSQ180), P2REL (FSQ130), P2UNR (FSQ180)	Continuous
P1TWIN P2TWIN	Family/HH	Household has sampled twins	P1PER, P2PER	0 = No twin in household 1 = Twin in household
WKPOVRTY	Family/HH	Poverty indicator	WKINCOME, P2HTOTAL (composites), and Census-defined thresholds	1 = Below poverty threshold 2 = At or above poverty threshold
WKINCOME	Family/HH	Household income	INCOME (PAQ100). Imputed using hotdeck imputation if missing. Imputation used partial income information collected in the parent interview if available.	Continuous

Table 7-6. Composite variables (continued)

Variable Name	Category	Description	Derived from	Values
WKSESL	Family/HH	Socioeconomic status (SES) scale	Derived from the logarithm of WKINCOME, WKMOMED, WKDADED, WKMOMSCR, (mother's occupation GSS prestige score), WKDADSCR (father's occupation GSS prestige score) (composites)	Continuous (Higher values indicate higher SES)
WKSESQ5	Family/HH	Quintile indicator for WKSESL (SES scale)	Derived from WKSESL at the child level using round 2 P2 weights	1 = First quintile (lowest SES) 2 = Second quintile 3 = Third quintile 4 = Fourth quintile 5 = Fifth quintile (highest SES)
WKPARED	Family/HH	Highest level of education for the child's parents or nonparent guardians who reside in the household. If only one parent or guardian resides in the household, WKPARED reflects that parent's education level.	WKMOMED, WKDADED (composites)	1 = 8th grade or below 2 = 9th to 12th grade 3 = High school diploma/equivalent 4 = Voc/Tech program 5 = Some college 6 = Bachelor's degree 7 = Graduate/professional school/no degree 8 = Master's degree 9 = Doctorate or professional degree

Table 7-6. Composite variables (continued)

Variable Name	Category	Description	Derived from	Values
WKMOMSCR	Family/HH	Mother's occupation GSS prestige score	Average of the 1998 GSS prestige scores of the 1998 census occupational category codes previously matched to the ECLS-K occupation codes. The ECLS-K occupation codes were collected in EMQ120, EMQ130, and EMQ140	29.6 Handler, Equip, Cleaner, Helpers, Labor 33.42 Production Working Occupation 34.95 Service Occupations 35.63 Agriculture, Forestry, Fishing Occupations 35.78 Marketing & Sales Occupation 35.92 Transportation, Material Moving 37.67 Precision Production Occupation 38.18 Administrative Support, Including Clerk 39.18 Mechanics & Repairs 39.2 Construction & Extractive Occupations 48.69 Technologists, Except Health 52.54 Writers, Artists, Entertainers, Athletes 53.5 Executive, Admin, Managerial Occupation 57.83 Health Technologists & Technicians 59 Social Scientist/Workers, Lawyers 61.56 Registered Nurses, Pharmacists 62.87 Natural Scientists & Mathematicians 63.43 Teacher, Except Postsecondary 64.89 Engineers, Surveyors, & Architects 72.1 Teachers; College, Postsecondary Counselors, Librarians 77.5 Physicians, Dentists, Veterinarians

Table 7-6. Composite variables (continued)

Variable Name	Category	Description	Derived from	Values
WKDADSCR	Family/HH	Father's occupation GSS prestige score	Average of the 1998 GSS prestige scores of the 1998 census occupational category codes previously matched to the ECLS-K occupation codes. The ECLS-K occupation codes were collected in EMQ120, EMQ130, and EMQ140	29.6 Handler, Equip, Cleaner, Helpers, Labor 33.42 Production Working Occupation 34.95 Service Occupations 35.63 Agriculture, Forestry, Fishing Occupations 35.78 Marketing & Sales Occupation 35.92 Transportation, Material Moving 37.67 Precision Production Occupation 38.18 Administrative Support, Including Clerk 39.18 Mechanics & Repairs 39.2 Construction & Extractive Occupations 48.69 Technologists, Except Health 52.54 Writers, Artists, Entertainers, Athletes 53.5 Executive, Admin, Managerial Occupation 57.83 Health Technologists & Technicians 59 Social Scientist/Workers, Lawyers 61.56 Registered Nurses, Pharmacists 62.87 Natural Scientists & Mathematicians 63.43 Teacher, Except Postsecondary 64.89 Engineers, Surveyors, & Architects 72.1 Teachers; College, Postsecondary Counselors, Librarians 77.5 Physicians, Dentists, Veterinarians

Table 7-6. Composite variables (continued)

Variable Name	Category	Description	Derived from	Values
WKLANGST	Family/HH	The primary language other than English spoken in the child's home	P1ANYLNG (PLQ.020), P1PRMLNG (PLQ060), PRIMARYLA (PLQ060), OTHERLAN (PLQ040), P2ANYLNG (SPQ155) P2PRMLNG (SPQ157)	1 = Non-English language 2 = English language
P1LANGUG	Family/HH	Language(s) spoken most often at home by the parent(s)/guardian(s) in the household	P1ANYLNG (PLQ020), OTHERLAN (PLQ040)	1 = Both only speak English language 2 = 1 (of 2) parents mainly speaks a non-English language 3 = Both only speak a non-English language
P2MARSTA	Family/HH	Marital status of parent(s) in household	P2CURMAR (FSQ200)	1 = Married 2 = Separated 3 = Divorced 4 = Widowed 5 = Never married 7 = No biological or adoptive parents in household
P2ERRFLG	Family/HH	A dichotomous variable indicating clear errors in the household roster	P1REL (FSQ130), P1UNR (FSQ180), P2JOI, P2RDP, P2REAS1	0 = False 1 = True
P2KSHCHG	Family/HH	Household had roster change. A dichotomous variable set to 1 if a household experienced a true change between rounds. In this context, true change means that at least one individual entered or left the household between rounds, and this change did not involve interviewer corrections for round 1 roster errors.	P1REL (FSQ130), P1UNR (FSQ180), P2JOI, P2REAS1	0 = False 1 = True
P2SPQDAT	Family/HH	Record in section SPQ (fall nonrespondent)	PARENTID	0 = No record in SPQ 1 = Record in SPQ
P1DADTYP P2DADTYP	Family/HH	Type of resident father	P1REL (FSQ130), P1UNR (FSQ180), P2REL (PSQ130) P2UNR (FSQ180)	1 = Biological 2 = Other 3 = None

Table 7-6. Composite variables (continued)

Variable Name	Category	Description	Derived from	Values
P1HDAD P2HDAD	Family/HH	Indicates whether the birth, adoptive, step, or foster father of the focal child resides in the household with the focal child	P1REL (FSQ130), P1UNR (FSQ180), P2REL (FSQ130), P2UNR (FSQ180)	1 = Birth father 2 = Adoptive father 3 = Step father 4 = Foster father 5 = Partner father 6 = Father, unknown type 7 = No resident father
P1HDAGE P2HDAGE	Family/HH	Age of resident father	P1AGE P2AGE	Continuous
P1HDRACE P2HDRACE	Family/HH	Race and ethnicity of the father or male guardian in the household	P1RC_1 (FSQ190), P1RC_2 (FSQ190), P1RC_3 (FSQ190), P1RC_4 (FSQ190), P1RC_5 (FSQ190), P1RC_6 (FSQ190), RACEOS (FSQ190), P1HSP (FSQ190), P2RC_1 (FSQ195), P2RC_2 (FSQ195), P2RC_3 (FSQ195), P2RC_4 (FSQ195), P2RC_5 (FSQ195), RACEOS (FSQ198), P2HSP (FSQ190)	1 = White, non-Hispanic 2 = Black or African American, non-Hispanic 3 = Hispanic, race specified 4 = Hispanic, no race specified 5 = Asian 6 = Native Hawaiian or other Pacific Islander 7 = American Indian or Alaskan Native 8 = More than one race, non-Hispanic

Table 7-6. Composite variables (continued)

Variable Name	Category	Description	Derived from	Values
P1HDASN P2HDASN	Family/HH	Father or male guardian in the household is Asian	P1RC_1 (FSQ190), P1RC_2 (FSQ190), P1RC_3 (FSQ190), P1RC_4 (FSQ190), P1RC_5 (FSQ190), P1RC_6 (FSQ190), RACEOS (FSQ190), P1HSP (FSQ190), P2RC_1 (FSQ195), P2RC_2 (FSQ195), P2RC_3 (FSQ195), P2RC_4 (FSQ195), P2RC_5 (FSQ195), RACEOS (FSQ198), P2HSP (FSQ190)	1 = Yes 2 = No
P1HDBLCK P2HDBLCK	Family/HH	Father or male guardian in the household is African American	P1RC_1 (FSQ190), P1RC_2 (FSQ190), P1RC_3 (FSQ190), P1RC_4 (FSQ190), P1RC_5 (FSQ190), P1RC_6 (FSQ190), RACEOS (FSQ190), P1HSP (FSQ190), P2RC_1 (FSQ195), P2RC_2 (FSQ195), P2RC_3 (FSQ195), P2RC_4 (FSQ195), P2RC_5 (FSQ195), RACEOS (FSQ198), P2HSP (FSQ190)	1 = Yes 2 = No

Table 7-6. Composite variables (continued)

Variable Name	Category	Description	Derived from	Values
P1HDIND P2HDIND	Family/HH	Father or male guardian in the household is American Indian or Native Alaskan	P1RC_1 (FSQ190), P1RC_2 (FSQ190), P1RC_3 (FSQ190), P1RC_4 (FSQ190), P1RC_5 (FSQ190), P1RC_6 (FSQ190), RACEOS (FSQ190), P1HSP (FSQ190), P2RC_1 (FSQ195), P2RC_2 (FSQ195), P2RC_3 (FSQ195), P2RC_4 (FSQ195), P2RC_5 (FSQ195), RACEOS (FSQ198), P2HSP (FSQ190)	1 = Yes 2 = No
P1HDPACI P2HDPACI	Family/HH	Father or male guardian in the household is Native Hawaiian or other Pacific Islander	P1RC_1 (FSQ190), P1RC_2 (FSQ190), P1RC_3 (FSQ190), P1RC_4 (FSQ190), P1RC_5 (FSQ190), P1RC_6 (FSQ190), RACEOS (FSQ190), P1HSP (FSQ190), P2RC_1 (FSQ195), P2RC_2 (FSQ195), P2RC_3 (FSQ195), P2RC_4 (FSQ195), P2RC_5 (FSQ195), RACEOS (FSQ198), P2HSP (FSQ190)	1 = Yes 2 = No

Table 7-6. Composite variables (continued)

Variable Name	Category	Description	Derived from	Values
P1HDWHT P2HDWHT	Family/HH	Father or male guardian in the household is white	P1RC_1 (FSQ190), P1RC_2 (FSQ190), P1RC_3 (FSQ190), P1RC_4 (FSQ190), P1RC_5 (FSQ190), P1RC_6 (FSQ190), RACEOS (FSQ190), P1HSP (FSQ190), P2RC_1 (FSQ195), P2RC_2 (FSQ195), P2RC_3 (FSQ195), P2RC_4 (FSQ195), P2RC_5 (FSQ195), RACEOS (FSQ198), P2HSP (FSQ190)	1 = Yes 2 = No
P1HDHISP P2HDHISP	Family/HH	Father or male guardian in the household is Hispanic	P1RC_1 (FSQ190), P1RC_2 (FSQ190), P1RC_3 (FSQ190), P1RC_4 (FSQ190), P1RC_5 (FSQ190), P1RC_6 (FSQ190), RACEOS (FSQ190), P1HSP (FSQ190), P2RC_1 (FSQ195), P2RC_2 (FSQ195), P2RC_3 (FSQ195), P2RC_4 (FSQ195), P2RC_5 (FSQ195), RACEOS (FSQ198), P2HSP (FSQ190)	1 = Yes 2 = No

Table 7-6. Composite variables (continued)

Variable Name	Category	Description	Derived from	Values
P1HDMT1R P2HDMT1R	Family/HH	Father or male guardian in the household is more than one race	P1RC_1 (FSQ190), P1RC_2 (FSQ190), P1RC_3 (FSQ190), P1RC_4 (FSQ190), P1RC_5 (FSQ190), P1RC_6 (FSQ190), RACEOS (FSQ190), P1HSP (FSQ190), P2RC_1 (FSQ195), P2RC_2 (FSQ195), P2RC_3 (FSQ195), P2RC_4 (FSQ195), P2RC_5 (FSQ195), RACEOS (FSQ198), P2HSP (FSQ190)	1 = Yes 2 = No
WKDADED	Family/HH	Father's highest level of education	P1HIG (PEQ020), P1HIS (PEQ030) P2HIGHGR (SPQ160), P2HSDIPL (SPQ170), P2HIGHSP (SPQ180), P2PRTDIP (SPQ190), Imputed using hotdeck imputation if missing.	1 = 8th grade or below 2 = 9th to 12th grade 3 = High school diploma/equivalent 4 = Voc/Tech program 5 = Some college 6 = Bachelor's degree 7 = Graduate/professional school/no degree 8 = Master's degree 9 = Doctorate or professional degree

Table 7-6. Composite variables (continued)

Variable Name	Category	Description	Derived from	Values
P1HDEMP	Family/HH	The work status of the father/male guardian in the household.	P1HRS (EMQ050), P1PAY_1 (EMQ020), P1VAC (EMQ030), P1LOK (EMQ060), P1DO1 (EMQ070), P1DO2 (EMQ070), P1DO3 (EMQ070), P1DO4 (EMQ070), P1DO5 (EMQ070), P1DO6 (EMQ070), P1DO7 (EMQ070)	1 = 35 hours or more per week (EMQ.050) 2 = Less than 35 hours per week (EMQ.050) 3 = Looking for work (EMQ.060 & EMQ.070 does not =6) 4 = Not in the labor force
P1DADOCC	Family/HH	Father's occupation	FORWHOM (EMQ120), KINDOFIN (EMQ130), and KINDOFWO (EMP140)	1 Handler, Equip, Cleaner, Helpers, Labor 2 Production Working Occupation 3 Service Occupations 4 Agriculture, Forestry, Fishing Occupations 5 Marketing & Sales Occupation 6 Transportation, Material Moving 7 Precision Production Occupation 8 Administrative Support, Including Clerk 9 Mechanics & Repairs 10 Construction & Extractive Occupations 11 Technologists, Except Health 12 Writers, Artists, Entertainers, Athletes 13 Executive, Admin, Managerial Occupation

Table 7-6. Composite variables (continued)

Variable Name	Category	Description	Derived from	Values
P1DADOCC (continued)	Family/HH	Father's occupation	FORWHOM (EMQ120), KINDOFIN (EMQ130), and KINDOFWO (EMP140)	14 Health Technologists & Technicians 15 Social Scientist/Workers, Lawyers 16 Registered Nurses, Pharmacists 17 Natural Scientists & Mathematicians 18 Teacher, Except Postsecondary 19 Engineers, Surveyors, & Architects 20 Teachers; College, Postsecondary Counselors, Librarians 21 Physicians, Dentists, Veterinarians 22 Unemployed, Retired
P1HDLANG	Family/HH	Language used by the father when speaking to the child	P1ANYLNG (PLQ020), RESLA (PLQ.080)	1 = Never speaks non-English language 2 = Sometimes speaks non-English language 3 = Often speaks non-English language 4 = Very often speaks non-English language

Table 7-6. Composite variables (continued)

Variable Name	Category	Description	Derived from	Values
P1HDLTOD	Family/HH	Language used by the child when speaking to the father	P1ANYLNG (PLQ020), RESLA (PLQ.080)	1 = Never speaks non-English language 2 = Sometimes speaks non-English language 3 = Often speaks non-English language 4 = Very often speaks non-English language
P1MOMTYP P2MOMTYP	Family/HH	Type of resident mother	P1REL (FSQ130), P1UNR (FSQ180), P2REL (FSQ130), P2UNR (FSQ180)	1 = Biological 2 = Other 3 = None
P1HMOM P2HMOM	Family/HH	Indicates whether the birth, adoptive, step, or foster mother of the focal child resides in the household with the focal child	P1REL (FSQ130), P1UNR (FSQ180), P2REL (FSQ130), P2UNR (FSQ180)	1 = Birth mother 2 = Adoptive mother 3 = Step mother 4 = Foster mother 5 = Partner mother 6 = Mother, unknown type 7 = No resident mother
P1HMAGE P2HMAGE	Family/HH	Age of resident mother	P1AGE P2AGE	Continuous
P1HMRACE P2HMRACE	Family/HH	Race and ethnicity of the mother or female guardian in the household	P1RC_1 (FSQ190), P1RC_2 (FSQ190), P1RC_3 (FSQ190), P1RC_4 (FSQ190), P1RC_5 (FSQ190), P1RC_6 (FSQ190), RACEOS (FSQ190), P1HSP (FSQ190), P2RC_1 (FSQ195), P2RC_2 (FSQ195), P2RC_3 (FSQ195), P2RC_4 (FSQ195), P2RC_5 (FSQ195), RACEOS (FSQ198), P2HSP (FSQ190)	1 = White, non-Hispanic 2 = Black or African American, non-Hispanic 3 = Hispanic, race specified 4 = Hispanic, no race specified 5 = Asian 6 = Native Hawaiian or other Pacific Islander 7 = American Indian or Alaskan Native 8 = More than 1 race, non-Hispanic 9 = Other

Table 7-6. Composite variables (continued)

Variable Name	Category	Description	Derived from	Values
P1HMASN P2HMASN	Family/HH	Mother or female guardian in the household is Asian	P1RC_1 (FSQ190), P1RC_2 (FSQ190), P1RC_3 (FSQ190), P1RC_4 (FSQ190), P1RC_5 (FSQ190), P1RC_6 (FSQ190), RACEOS (FSQ190), P1HSP (FSQ190), P2RC_1 (FSQ195), P2RC_2 (FSQ195), P2RC_3 (FSQ195), P2RC_4 (FSQ195), P2RC_5 (FSQ195), RACEOS (FSQ198), P2HSP (FSQ190)	1 = Yes 2 = No
P1HMBLCK P2HMBLCK	Family/HH	Mother or female guardian in the household is African American	P1RC_1 (FSQ190), P1RC_2 (FSQ190), P1RC_3 (FSQ190), P1RC_4 (FSQ190), P1RC_5 (FSQ190), P1RC_6 (FSQ190), RACEOS (FSQ190), P1HSP (FSQ190), P2RC_1 (FSQ195), P2RC_2 (FSQ195), P2RC_3 (FSQ195), P2RC_4 (FSQ195), P2RC_5 (FSQ195), RACEOS (FSQ198), P2HSP (FSQ190)	1 = Yes 2 = No

Table 7-6. Composite variables (continued)

Variable Name	Category	Description	Derived from	Values
P1HMIND P2HMIND	Family/HH	Mother or female guardian in the household is American Indian or Native Alaskan	P1RC_1 (FSQ190), P1RC_2 (FSQ190), P1RC_3 (FSQ190), P1RC_4 (FSQ190), P1RC_5 (FSQ190), P1RC_6 (FSQ190), RACEOS (FSQ190), P1HSP (FSQ190), P2RC_1 (FSQ195), P2RC_2 (FSQ195), P2RC_3 (FSQ195), P2RC_4 (FSQ195), P2RC_5 (FSQ195), RACEOS (FSQ198), P2HSP (FSQ190)	1 = Yes 2 = No
P1HMPACI P2HMPACI	Family/HH	Mother or female guardian in the household is Native Hawaiian or other Pacific Islander	P1RC_1 (FSQ190), P1RC_2 (FSQ190), P1RC_3 (FSQ190), P1RC_4 (FSQ190), P1RC_5 (FSQ190), P1RC_6 (FSQ190), RACEOS (FSQ190), P1HSP (FSQ190), P2RC_1 (FSQ195), P2RC_2 (FSQ195), P2RC_3 (FSQ195), P2RC_4 (FSQ195), P2RC_5 (FSQ195), RACEOS (FSQ198), P2HSP (FSQ190)	1 = Yes 2 = No

Table 7-6. Composite variables (continued)

Variable Name	Category	Description	Derived from	Values
P1HMWHT P2HMWHT	Family/HH	Mother or female guardian in the household is white	P1RC_1 (FSQ190), P1RC_2 (FSQ190), P1RC_3 (FSQ190), P1RC_4 (FSQ190), P1RC_5 (FSQ190), P1RC_6 (FSQ190), RACEOS (FSQ190), P1HSP (FSQ190), P2RC_1 (FSQ195), P2RC_2 (FSQ195), P2RC_3 (FSQ195), P2RC_4 (FSQ195), P2RC_5 (FSQ195), RACEOS (FSQ198), P2HSP (FSQ190)	1 = Yes 2 = No
P1HMHISP P2HMHISP	Family/HH	Mother or female guardian in the household is Hispanic	P1RC_1 (FSQ190), P1RC_2 (FSQ190), P1RC_3 (FSQ190), P1RC_4 (FSQ190), P1RC_5 (FSQ190), P1RC_6 (FSQ190), RACEOS (FSQ190), P1HSP (FSQ190), P2RC_1 (FSQ195), P2RC_2 (FSQ195), P2RC_3 (FSQ195), P2RC_4 (FSQ195), P2RC_5 (FSQ195), RACEOS (FSQ198), P2HSP (FSQ190)	1 = Yes 2 = No

Table 7-6. Composite variables (continued)

Variable Name	Category	Description	Derived from	Values
P1HMMT1R P2HMMT1R	Family/HH	Mother or female guardian in the household is more than one race	P1RC_1 (FSQ190), P1RC_2 (FSQ190), P1RC_3 (FSQ190), P1RC_4 (FSQ190), P1RC_5 (FSQ190), P1RC_6 (FSQ190), RACEOS (FSQ190), P1HSP (FSQ190), P2RC_1 (FSQ195), P2RC_2 (FSQ195), P2RC_3 (FSQ195), P2RC_4 (FSQ195), P2RC_5 (FSQ195), RACEOS (FSQ198), P2HSP (FSQ190)	1 = Yes 2 = No
WKMOMED	Family/HH	Mother's highest level of education	P1HIG (PEQ020), P1HIS (PEQ030), P2HIGHGR (SPQ160), P2HSDIPL (SPQ170), P2HIGHSP (SPQ180), P2PRTDIP (SPQ190), Imputed using hotdeck imputation if missing.	1 = 8th grade or below 2 = 9th to 12th grade 3 = High school diploma/equivalent 4 = Voc/Tech program 5 = Some college 6 = Bachelor's degree 7 = Graduate/professional school/no degree 8 = Master's degree 9 = Doctorate or professional degree

Table 7-6. Composite variables (continued)

Variable Name	Category	Description	Derived from	Values
P1HMEMP	Family/HH	The work status of the mother or female guardian in the household	P1PAY_1 (EMQ020), P1HRS (EMQ050), P1VAC (EMQ030), P1LOK (EMQ060), P1DO1 (EMQ070), P1DO2 (EMQ070), P1DO3 (EMQ070), P1DO4 (EMQ070), P1DO5 (EMQ070), P1DO6 (EMQ070), P1DO7 (EMQ070)	1 = 35 hours or more per week (EMQ.050) 2 = Less than 35 hours per week (EMQ.050) 3 = Looking for work (EMQ.060 & EMQ.070 does not =6) 4 = Not in the labor force
P1MOMOCC	Family/HH	Mother's occupation	Combination of FORWHOM (EMQ120), KINDOFIN (EMQ130), and KINDOFWO (EMP140)	1 Handler, Equip, Cleaner, Helpers, Labor 2 Production Working Occupation 3 Service Occupations 4 Agriculture, Forestry, Fishing Occupations 5 Marketing & Sales Occupation 6 Transportation, Material Moving 7 Precision Production Occupation 8 Administrative Support, Including Clerk 9 Mechanics & Repairs 10 Construction & Extractive Occupations 11 Technologists, Except Health 12 Writers, Artists, Entertainers, Athletes 13 Executive, Admin, Managerial Occupation

Table 7-6. Composite variables (continued)

Variable Name	Category	Description	Derived from	Values
P1MOMOCC (continued)	Family/HH	Mother's occupation	Combination of FORWHOM (EMQ120), KINDOFIN (EMQ130), and KINDOFWO (EMP140)	14 Health Technologists & Technicians 15 Social Scientist/Workers, Lawyers 16 Registered Nurses, Pharmacists 17 Natural Scientists & Mathematicians 18 Teacher, Except Postsecondary 19 Engineers, Surveyors, & Architects 20 Teachers; College, Postsecondary Counselors, Librarians 21 Physicians, Dentists, Veterinarians 22 Unemployed, Retired
P1HMLANG	Family/HH	Language used by the mother when speaking to the child	P1ANYLNG (PLQ020), RESLA (PLQ.080)	1 = Never speaks non-English language 2 = Sometimes speaks non-English language 3 = Often speaks non-English language 4 = Very often speaks non-English language

Table 7-6. Composite variables (continued)

Variable Name	Category	Description	Derived from	Values
P1HMLTOM	Family/HH	Language used by the child when speaking to the mother	P1ANYLNG (PLQ020), RESLA (PLQ.080)	1 = Never speaks non-English language 2 = Sometimes speaks non-English language 3 = Often speaks non-English language 4 = Very often speaks non-English language
P1HMAFB	Family/HH	Resident biological mother's age when she first gave birth	P1AGE, P1OLDMOM (MHQ165)	Continuous
WKHEARLY	Family/HH	Mother worked between the child's birth and the start of kindergarten	P1MOMWRK (EMQ170), P2MOMWRK (SPQ200)	1 = Yes 2 = No
WKHMOMAR	Family/HH	Resident biological mother was married to the biological father at the time of the child's birth	P1LEGMAR (MHQ020), P1BIOLOG (MHQ.060), P1BIOPAR (MHQ125), P1WHNMAR (MHQ025), P1BIOLO2 (MHQ065), P1BIOLO3 (MHQ065), P1MDWHN (MHQ130), P1MDWH2 (MHQ130), P2BIOLOG (SPQ150)	1 = Yes 2 = No
P1ABSDAD	Family/HH	Type of nonresident father	P1REL (FSQ130), P1BIO (HRQ030)	1 = Biological only 2 = Both biological and adoptive
P1BDVITL	Family/HH	The vital status of the nonresident biological father	P1BIO (HRQ030)	1 = Living 2 = Dead
P1BDAGE	Family/HH	Age of biological nonresident father	P1DB1 (HRQ040), P1BIOLOG (HRQ030)	Continuous

Table 7-6. Composite variables (continued)

Variable Name	Category	Description	Derived from	Values
P1BDRACE	Family/HH	Race and ethnicity of nonresident biological father	P1BM1 (HRQ100), P1BM2 (HRQ100), P1BM3 (HRQ100), P1BM4 (HRQ100), P1BM5 (HRQ100), P1BM6 (HRQ100), P1BMH (HRQ090)	1 = White, non-Hispanic 2 = Black or African American, non-Hispanic 3 = Hispanic, race specified 4 = Hispanic, no race specified 5 = Asian 6 = Native Hawaiian or other Pacific Islander 7 = American Indian or Alaskan Native 8 = More than one race, non-Hispanic 9 = Other
P1BDASN	Family/HH	Nonresident biological father is Asian	P1BM1 (HRQ100), P1BM2 (HRQ100), P1BM3 (HRQ100), P1BM4 (HRQ100), P1BM5 (HRQ100), P1BM6 (HRQ100), P1BMH (HRQ090)	1 = Yes 2 = No
P1BDBLCK	Family/HH	Nonresident biological father is African American	P1BM1 (HRQ100), P1BM2 (HRQ100), P1BM3 (HRQ100), P1BM4 (HRQ100), P1BM5 (HRQ100), P1BM6 (HRQ100), P1BMH (HRQ090)	1 = Yes 2 = No
P1BDIND	Family/HH	Nonresident biological father is American Indian or Alaskan Native	P1BM1 (HRQ100), P1BM2 (HRQ100), P1BM3 (HRQ100), P1BM4 (HRQ100), P1BM5 (HRQ100), P1BM6 (HRQ100), P1BMH (HRQ090)	1 = Yes 2 = No

Table 7-6. Composite variables (continued)

Variable Name	Category	Description	Derived from	Values
P1BDPACI	Family/HH	Nonresident biological father is Native Hawaiian or other Pacific Islander	P1BM1 (HRQ100), P1BM2 (HRQ100), P1BM3 (HRQ100), P1BM4 (HRQ100), P1BM5 (HRQ100), P1BM6 (HRQ100), P1BMH (HRQ090)	1 = Yes 2 = No
P1BDWHT	Family/HH	Nonresident biological father is white	P1BM1 (HRQ100), P1BM2 (HRQ100), P1BM3 (HRQ100), P1BM4 (HRQ100), P1BM5 (HRQ100), P1BM6 (HRQ100), P1BMH (HRQ090)	1 = Yes 2 = No
P1BDHISP	Family/HH	Nonresident biological father is Hispanic	P1BM1 (HRQ100), P1BM2 (HRQ100), P1BM3 (HRQ100), P1BM4 (HRQ100), P1BM5 (HRQ100), P1BM6 (HRQ100), P1BMH (HRQ090)	1 = Yes 2 = No
P1BDMT1R	Family/HH	Nonresident biological father is more than one race	P1BM1 (HRQ100), P1BM2 (HRQ100), P1BM3 (HRQ100), P1BM4 (HRQ100), P1BM5 (HRQ100), P1BM6 (HRQ100), P1BMH (HRQ090)	1 = Yes 2 = No

Table 7-6. Composite variables (continued)

Variable Name	Category	Description	Derived from	Values
P1BDADED	Family/HH	Education of nonresident biological father	P1HIG (PEQ020), P1HIS (PEQ030)	1 = 8th grade or below 2 = 9th to 12th grade 3 = High school diploma/equivalent 4 = Voc/Tech program 5 = Some college 6 = Bachelor's degree 7 = Graduate/professional school/no degree 8 = Master's degree 9 = Doctorate or professional degree
P1ABSMOM	Family/HH	Type of nonresident mother	P1REL (FSQ130), P1BIO (HRQ030)	1 = Biological only 2 = Both biological and adoptive
P1BMVITL	Family/HH	The vital status of the nonresident biological mother	P1BIO (HRQ030)	1 = Living 2 = Dead
P1BMAGE	Family/HH	Age of biological nonresident mother	P1DB1 (HRQ040), P1BIO (HRQ030), MONTHINT	Continuous
P1BMRACE	Family/HH	Race and ethnicity of nonresident biological mother	P1BM1 (HRQ100), P1BM2 (HRQ100), P1BM3 (HRQ100), P1BM4 (HRQ100), P1BM5 (HRQ100), P1BM6 (HRQ100), P1BMH (HRQ090)	1 = White, non-Hispanic 2 = Black or African American, non-Hispanic 3 = Hispanic, race specified 4 = Hispanic, no race specified 5 = Asian 6 = Native Hawaiian or other Pacific Islander 7 = American Indian or Alaskan Native 8 = More than 1 race, non-Hispanic 9 = Other
P1BMASN	Family/HH	Nonresident biological mother is Asian	P1BM1 (HRQ100), P1BM2 (HRQ100), P1BM3 (HRQ100), P1BM4 (HRQ100), P1BM5 (HRQ100), P1BM6 (HRQ100), P1BMH (HRQ090)	1 = Yes 2 = No

Table 7-6. Composite variables (continued)

Variable Name	Category	Description	Derived from	Values
P1BMBLCK	Family/HH	Nonresident biological mother is African American	P1BM1 (HRQ100), P1BM2 (HRQ100), P1BM3 (HRQ100), P1BM4 (HRQ100), P1BM5 (HRQ100), P1BM6 (HRQ100), P1BMH (HRQ090)	1 = Yes 2 = No
P1BMIND	Family/HH	Nonresident biological mother is American Indian or Alaskan Native	P1BM1 (HRQ100), P1BM2 (HRQ100), P1BM3 (HRQ100), P1BM4 (HRQ100), P1BM5 (HRQ100), P1BM6 (HRQ100), P1BMH (HRQ090)	1 = Yes 2 = No
P1BMPACI	Family/HH	Nonresident biological mother is Native Hawaiian or other Pacific Islander	P1BM1 (HRQ100), P1BM2 (HRQ100), P1BM3 (HRQ100), P1BM4 (HRQ100), P1BM5 (HRQ100), P1BM6 (HRQ100), P1BMH (HRQ090)	1 = Yes 2 = No
P1BMWHT	Family/HH	Nonresident biological mother is white	P1BM1 (HRQ100), P1BM2 (HRQ100), P1BM3 (HRQ100), P1BM4 (HRQ100), P1BM5 (HRQ100), P1BM6 (HRQ100), P1BMH (HRQ090)	1 = Yes 2 = No
P1BMHISP	Family/HH	Nonresident biological mother is Hispanic	P1BM1 (HRQ100), P1BM2 (HRQ100), P1BM3 (HRQ100), P1BM4 (HRQ100), P1BM5 (HRQ100), P1BM6 (HRQ100), P1BMH (HRQ090)	1 = Yes 2 = No

Table 7-6. Composite variables (continued)

Variable Name	Category	Description	Derived from	Values
P1BMMT1R	Family/HH	Nonresident biological mother is more than one race	P1BM1 (HRQ100), P1BM2 (HRQ100), P1BM3 (HRQ100), P1BM4 (HRQ100), P1BM5 (HRQ100), P1BM6 (HRQ100), P1BMH (HRQ090)	1 = Yes 2 = No
P1BMOMED	Family/HH	Education of nonresident biological mother	P1HIG (PEQ020), P1HIS (PEQ030)	1 = 8th grade or below 2 = 9th to 12th grade 3 = High school diploma/equivalent 4 = Voc/Tech program 5 = Some college 6 = Bachelor's degree 7 = Graduate/professional school/no degree 8 = Master's degree 9 = Doctorate or professional degree
P1BMAFB	Family/HH	Nonresident biological mother's age when she first gave birth	P1OLDMOM (MHQ165)	Continuous
WKBMOMAR	Family/HH	Indicates whether the nonresident biological mother was married to the biological father at the time of the child's birth	P1LEGMAR (MHQ020), P1BIOLOG (MHQ060), P1BIOPAR (MHQ125), P1WHNMAR (MHQ025), P1BIOLO2 (MHQ065), P1BIOLO3 (MHQ065), P1MDWHN (MHQ130), P1MDWH2 (MHQ130), P2BIOLOG (SPQ150)	1 = Yes 2 = No

Table 7-6. Composite variables (continued)

Variable Name	Category	Description	Derived from	Values
P1ADADED	Family/HH	Education of adoptive nonresident father	P1HIG (PEQ020), P1HIS (PEQ030)	1 = 8th grade or below 2 = 9th to 12th grade 3 = High school diploma/equivalent 4 = Voc/Tech program 5 = Some college 6 = Bachelor’s degree 7 = Graduate/professional school/no degree 8 = Master’s degree 9 = Doctorate or professional degree
P1AMOMED	Family/HH	Education of adoptive nonresident mother	P1HIG (PEQ020), P1HIS (PEQ030)	1 = 8th grade or below 2 = 9th to 12th grade 3 = High school diploma/equivalent 4 = Voc/Tech program 5 = Some college 6 = Bachelor’s degree 7 = Graduate/professional school/no degree 8 = Master’s degree 9 = Doctorate or professional degree
B1AGE	Teacher	Teacher’s age	B1YRBORN	21-78
A1TQUEX	Teacher	Presence of round 1 teacher questionnaire A AM data		0 = False 1 = True
A2ACLASS	Teacher	Presence of round 2 teacher questionnaire A AM data	A2ANEW	0 = False 1 = True

Table 7-6. Composite variables (continued)

Variable Name	Category	Description	Derived from	Values
A2DCLASS	Teacher	Presence of round 2 teacher questionnaire A all day data	A2DNEW	0 = False 1 = True
A2PCLASS	Teacher	Presence of round 2 teacher questionnaire A PM data	A2PNEW	0 = False 1 = True
A2TQUEX	Teacher	Presence or absence of round 2 teacher questionnaire A data		0 = False 1 = True
B1TQUEX	Teacher	Presence of round 1 teacher questionnaire B data		0 = False 1 = True
B2TQUEX	Teacher	Presence of round 2 teacher questionnaire B data		0 = False 1 = True
A1ACLASS	Classroom	Presence of round 1 teacher questionnaire A AM data	A1ABLACK, A1ATOTRA	0 = False 1 = True
A1APBLK	Classroom	Percentage of blacks in class-AM	A1ABLACK, A1ATOTRA	Continuous
A1APHIS	Classroom	Percentage of Hispanics in class-AM	A1AHISP, A1ATOTRA	Continuous
A1APMIN	Classroom	Percentage of minorities in class-AM	A1AASIAN, A1AHISP, A1ABLACK, A1AAMRIN, A1ARACEO, A1ATOTRA	Continuous
A1DCLASS	Classroom	Presence of round 1 teacher questionnaire A all day (AD) data		0 = False 1 = True
A1DPBLK	Classroom	Percentage of blacks in class-AD	A1DBLACK, A1DTOTRA	Continuous
A1DPHIS	Classroom	Percentage of Hispanics in class-AD	A1DHISP, A1DTOTRA	Continuous
A1DPMIN	Classroom	Percentage of minorities in class-AD	A1DASIAN, A1DHISP, A1DBLACK, A1DAMRIN, A1DRACEO, A1DTOTRA	Continuous
A1PCLASS	Classroom	Presence of round 1 teacher questionnaire A PM data		0 = False 1 = True
A1PPBLK	Classroom	Percentage of blacks in Class-PM	A1PBLACK, A1PTOTRA	Continuous
A1PPHIS	Classroom	Percentage of Hispanics in Class-PM	A1PHISP, A1PTOTRA	Continuous
A1PPMIN	Classroom	Percentage of minorities in Class-PM	A1PASIAN, A1PHISP, A1PBLACK, A1PAMRIN, A1PRACEO, A1PTOTRA	Continuous

Table 7-6. Composite variables (continued)

Variable Name	Category	Description	Derived from	Values
CREGION	School	Indicates the geographic region in which the child lives	Sampling Frame	1 = Northeast: CT, ME, MA, NH, RI, VT, NJ, NY, PA 2 = Midwest: IL, IN, MI, OH, WI, IA, KS, MN, MO, NE, ND, SD 3 = South: DE, DC, FL, GA, MD, NC, SC, VA, WV, AL, KY, MS, TN, AR, LA, OK, TX 4 = West: AZ, CO, ID, MT, NV, NM, UT, WY, AK, CA, HA, OR, WA
KURBAN	School	Denoted urbanicity, from the sampling frame. Urbanicity is assigned by the Census Bureau's TIGER geographic information system.	Sampling Frame	1 = Large city - a central city of Consolidated Metropolitan Statistical Area (CMSA) with a pop. greater than or equal to 250,000 2 = Mid-size city - a central city of a CMSA or Metropolitan Statistical Area (MSA) with a pop. less than 250,000 3 = Urban fringe of large city – any incorporated place, Census Designated Place, or nonplace territory within a CMSA or MSA of a large city and defined as urban by the Census Bureau

Table 7-6. Composite variables (continued)

Variable Name	Category	Description	Derived from	Values
KURBAN	School	Denoted urbanicity, from the sampling frame. Urbanicity is assigned by the Census Bureau's TIGER geographic information system.	Frame	5 = Large town – an incorporated place or Census Designated Place with a pop. greater than or equal to 25,000 and located outside a CMSA or MSA 6 = Small town – an incorporated place or Census Designated Place with a pop. less than 25,000 and greater than 2,500 – located outside a CMSA or MSA 7 = Rural – any incorporated place, Census Designated Place, or nonplace territory designated as rural by the Census Bureau
S2KSCTYP	School	School type from the school administrator questionnaire	S2PUBLIC, S2CATHOL, S2OTHREL, S2OTHPRI, CS_TYPE2	1 = Catholic 2 = Other religious 3 = Other private 4 = Public
S2KPUPRI	School	Public or private school	S2KSCTYP	1 = Public 2 = Private
S2KENRLK	School	Total school kindergarten enrollment	S2HDCHDN, S2FDCHDN, S2KCCCHDN, S2TKCHDN, S2T1CHDN	Continuous
S2KENRLS	School	Total school enrollment	S2ANUMCH1, TOT_STD (from frame)	1 = 0-149 students 2 = 150-299 students 3 = 300-499 students 4 = 500-749 students 5 = 750 and above students
S2KLNGTH	School	School year length	S2NUMDAY	Continuous
S2KMINOR	School	Percentage of minority students in school	F2STUMIN (from frame), S2PCTHIS, S2INDPCT, S2ASNPCT, S2BLKPCT, S2PCFPCT, S2WHTPCT	Continuous
S2KGFTED	School	Percentage of gifted/talented students in school	S2GFTNBR, S2ANUMCH	Continuous
S2KFLNCH	School	Percentage of students eligible for free lunch in school	S2ELILNC, S2ANUMCH	Continuous

Table 7-6.   Composite variables (continued)

Variable Name	Category	Description	Derived from	Values
S2KRLNCH	School	Percent of students eligible for reduced price lunch in school	S2ELIRED, S2ANUMCH	Continuous
S2KSCLVL	School	School instructional level from school administrator questionnaire	S212TH, S211TH, S2TENTH, S2NINTH, S28TH, S27TH, S2SIXTH, S2FIFTH, S2FOURTH, S2THIRD, S2SECOND, S2GRADE1, S2PRE1, S2KINDER, S2TRANS, S2PRKNDR, S2TWEL12, S2ELEV11, S2TEN10, S2NIN9TH, S2ATE8TH, S2SEV7TH, S2SIX6TH, S2FIF5TH, S2FOR4TH, S2THI3RD, S2SEC2ND, S2FIR1ST, S2PRE1ST, S2KINGAR, S2REDDYN, S2PREK, HIGRADE (from frame)	1 = School high grade less than 1st grade 2 = School high grade is 1st, 2nd, or 3rd grade 3 = School high grade is 4th, 5th, or 6th grade 4 = School high grade is 7-12 grade
K2INFAC	School	Presence or absence of facilities checklist data		0 = ID not in facilities checklist 1 = ID in facilities checklist
S2INSAQ	School	Presence or absence of school administrator questionnaire (SAQ) data		0 = ID not in SAQ 1 = ID in SAQ

7.9 Masked Variables

All the variables from the ECLS-K restricted-use file are included in the same order on the ECLS-K public-use file. New variables created during the masking process are added to the end of the files. For some of the variables, certain categories were modified. The variable labels for those masked variables were updated from the restricted-use variables to reflect the new categories that were created during the masking process.

Variables on the restricted-use files were modified in different ways based on the disclosure analysis NCES conducted in order to protect the identity of the respondents and children. There are several types of modifications on the public-use files.

1. Outliers are top- or bottom- coded to prevent identification of unique schools, teachers, parents and children without affecting overall data quality.
2. Certain schools identified as at risk for disclosure have a 5 to 10 percent noise introduced in those variables that pose a risk for disclosure.
3. Variables with too few cases and a sparse distribution are suppressed in the public-use files.
4. Certain continuous variables are modified into categorical variables, and certain categorical variables have their categories collapsed in the public-use file. While this protects the cases from a disclosure risk, these variables can still be used in all different kinds of analysis such as regression analysis.

There is a comment field in the variable frequency distribution view screen of the electronic code book that displays a comment for each masked variable indicating whether the variable from the restricted-use file has been recoded or suppressed in the public-use file. Variables that were recoded in any way during the data masking process display the comment, "This data recoded for respondent confidentiality." Variables that were suppressed on the public-use file for protection of the respondent or child from identification display the comment, "This data suppressed for respondent confidentiality" and all values for the variable are set to equal –2 for that variable. New variables created during the masking process also display the comment, "This data recoded for respondent confidentiality."

The following tables present the list of the masked variables for each catalog. The tables display the variable name, variable label and the comment displayed in the electronic code book indicating if the variable was recoded or suppressed. A crosswalk of new variables created during masking with the restricted-use variables is presented to help the user identify the recoded variables. For example, A1KGTYP was created during masking and holds the values 1=regular kindergarten type and 2=other kindergarten type. Table 7-7 presents the masked variables for the child catalog. The table is sorted sequentially by the variable Field ID (see section 8.6.1 for how to use the variable Field ID for the child catalog.) Table 7-8 presents the masked variables for the teacher catalog. The table is sorted sequentially by the variable Field ID (see section 8.7.1 for how to use the variable Field ID for the teacher catalog.) Table 7-9 presents the masked variables for the school catalog. The table is sorted sequentially by the variable Field ID (see section 8.8.1 for how to use the variable Field ID for the school catalog.)

Table 7-7. Masked variables in the child catalog

FieldID	FieldName	FieldLabel	Comment
10	CS_TYPE2	TYPE OF SCHOOL IN SAMPLE FRAME	This data recoded for respondent confidentiality.
11	KURBAN	LOCATION TYPE IN SAMPLE FRAME	This data recoded for respondent confidentiality.
14	DOBYY	CHILD COMPOSITE DOB YEAR	This data recoded for respondent confidentiality.
17	R1_KAGE	R1 COMPOSITE CHILD ASSESSMENT AGE(MNTHS)	This data recoded for respondent confidentiality.
227	P1AGEENT	P1 AGE (MONTHS) AT KINDERGARTEN ENTRY	This data recoded for respondent confidentiality.
320	A1PBLK	A1 PERCENT OF BLACKS IN CLASS	This data recoded for respondent confidentiality.
323	B1AGE	B1 TEACHER'S AGE	This data recoded for respondent confidentiality.
324	B2AGE	B2 TEACHER'S AGE	This data recoded for respondent confidentiality.
326	S2KENRLK	S2 TOTAL SCHOOL K ENROLLMENT	This data recoded for respondent confidentiality.
327	S2KENRLS	S2 TOTAL SCHOOL ENROLLMENT	This data recoded for respondent confidentiality.
328	S2KMINOR	S2 PERCENT MINORITY STUDENTS	This data recoded for respondent confidentiality.
329	S2KFLNCH	S2 PCT FREE LUNCH ELIGIBLE STUDENTS	This data recoded for respondent confidentiality.
330	S2KRLNCH	S2 PCT REDUCED LUNCH ELIGIBLE STUDENTS	This data recoded for respondent confidentiality.
332	S2KLNGTH	S2 SCHOOL YEAR LENGTH	This data recoded for respondent confidentiality.
333	S2KGFTED	S2 PERCENT GIFTED/TALENTED STUDENTS	This data recoded for respondent confidentiality.
816	P1WHATDI	P1 CHQ125 1ST DIAGNOSIS-LEARNING ABILITY	This data suppressed for respondent confidentiality.
817	P1HOWOLD	P1 CHQ130 AGE AT 1ST DIAGNS-LRN ABLTY	This data suppressed for respondent confidentiality.
818	P1OLDUNT	P1 CHQ130 AGE 1ST DIAGNS-LRN ABLTY	This data suppressed for respondent confidentiality.
819	P1MMDIAG	P1 CHQ135 MNTH AT 1ST DIAGNS-LRN ABLTY	This data suppressed for respondent confidentiality.
820	P1YYDIAG	P1 CHQ135 YR AT 1ST DIAGNOSIS-LRN ABLTY	This data suppressed for respondent confidentiality.
824	P1PROFFD	P1 CHQ155 IF ACTIVITY PROBLEM DIAGNOSED	This data suppressed for respondent confidentiality.
825	P1WHATD3	P1 CHQ160 WHAT 1ST DIAGNOSIS - ACTIVITY	This data suppressed for respondent confidentiality.
826	P1HOWOL2	P1 CHQ165 AGE AT 1ST DIAGNOSIS-ACTIVITY	This data suppressed for respondent confidentiality.
827	P1OLDUN2	P1 CHQ165 AGE UNIT AT 1ST DIAGNS-ACTVITY	This data suppressed for respondent confidentiality.
828	P1MMDIA2	P1 CHQ170 MNTH AT 1ST DIAGNOSIS-ACTIVITY	This data suppressed for respondent confidentiality.
829	P1YYDIA2	P1 CHQ170 YR AT 1ST DIAGNOSIS-ACTIVITY	This data suppressed for respondent confidentiality.
833	P1HOWOL3	P1 CHQ190 AGE AT 1ST DIAGNOSIS-MOBILITY	This data suppressed for respondent confidentiality.
834	P1OLDUN3	P1 CHQ190 AGE UNIT AT 1ST DIAGNS-MOBLITY	This data suppressed for respondent confidentiality.
835	P1MMDIA3	P1 CHQ190 MONTH AT 1ST DIAGNS-MOBLITY	This data suppressed for respondent confidentiality.
836	P1YYDIA3	P1 CHQ195 YR AT 1ST DIAGNOSIS-MOBILITY	This data suppressed for respondent confidentiality.
841	P1HOWOL4	P1 CHQ220 AGE AT 1ST DIAGNOSIS-SPEECH	This data suppressed for respondent confidentiality.
842	P1OLDUN4	P1 CHQ220 AGE UNIT AT 1ST DIAGNS-SPEECH	This data suppressed for respondent confidentiality.

Table 7-7. Masked variables in the child catalog (continued)

FieldID	FieldName	FieldLabel	Comment
843	P1MMDIA4	P1 CHQ225 MONTH AT 1ST DIAGNOSIS-SPEECH	This data suppressed for respondent confidentiality.
844	P1YYDIA4	P1 CHQ225 YEAR AT 1ST DIAGNOSIS-SPEECH	This data suppressed for respondent confidentiality.
848	P1HOWOL5	P1 CHQ250 AGE AT 1ST DIAGNOSIS-HEARING	This data suppressed for respondent confidentiality.
849	P1OLDUN5	P1 CHQ250 AGE UNIT AT 1ST DIAGNS-HEARING	This data suppressed for respondent confidentiality.
850	P1MMDIA5	P1 CHQ255 MONTH AT 1ST DIAGNOSIS-HEARING	This data suppressed for respondent confidentiality.
851	P1YYDIA5	P1 CHQ255 YR AT 1ST DIAGNOSIS-HEARING	This data suppressed for respondent confidentiality.
852	P1HEARS	P1 CHQ260 DEGREE OF CHILD'S DEAFNESS	This data suppressed for respondent confidentiality.
853	P1HEARAI	P1 CHQ265 IF CHILD WEARS HEARING AID	This data suppressed for respondent confidentiality.
854	P1COCHLE	P1 CHQ270 IF CHILD HAS COCHLEAR IMPLANTS	This data suppressed for respondent confidentiality.
855	P1HEARS2	P1 CHQ280 DEVICE EFFECT ON CHD'S HEARING	This data suppressed for respondent confidentiality.
859	P1HOWOL6	P1 CHQ305 AGE AT 1ST DIAGNOSIS - VISION	This data suppressed for respondent confidentiality.
860	P1OLDUN6	P1 CHQ305 AGE UNIT AT 1ST DIAGNOS-VISION	This data suppressed for respondent confidentiality.
861	P1DIA6MM	P1 CHQ310 MONTH AT 1ST DIAGNOSIS-VISION	This data suppressed for respondent confidentiality.
862	P1DIA6YY	P1 CHQ310 YR AT 1ST DIAGNOSIS-VISION	This data suppressed for respondent confidentiality.
863	P1CORREC	P1 CHQ315 IF CHD'S VISION IS CORRECTABLE	This data suppressed for respondent confidentiality.
864	P1BESTEY	P1 CHQ320 WHAT CAN CHILD BEST SEE	This data suppressed for respondent confidentiality.
882	P1THER14	P1 CHQ375 AGE AT THERAPY BEGINNING	This data suppressed for respondent confidentiality.
883	P1THER15	P1 CHQ375 AGE UNIT AT THERAPY BEGINNING	This data suppressed for respondent confidentiality.
884	P1THER16	P1 CHQ380 MONTH AT THERAPY BEGINNING	This data suppressed for respondent confidentiality.
885	P1THER17	P1 CHQ380 YEAR AT THERAPY BEGINNING	This data suppressed for respondent confidentiality.
887	P1LASTMM	P1 CHQ390 LAST THERAPY MONTH	This data suppressed for respondent confidentiality.
888	P1LASTYY	P1 CHQ390 LAST THERAPY YEAR	This data suppressed for respondent confidentiality.
890	P1SPECIL	P1 CHQ410 IF CHD USES SPECIAL EQUIPMENT	This data suppressed for respondent confidentiality.
1025	P1PRMLNG	P1 PLQ060 WHAT PRIMARY LANGUAGE AT HOME	This data recoded for respondent confidentiality.
1482	P2CNTRYB	P2 INQ310 COUNTRY OF BIRTH	This data suppressed for respondent confidentiality.
1517	P2PRMLNG	P2 SPQ157 WHAT PRIMARY LANGUAGE AT HOME	This data recoded for respondent confidentiality.
1831	P2INCOME	P2 PAQ100 TOTAL HOUSEHOLD INCOME ($)	This data recoded for respondent confidentiality.
1836	P2HOWPAY	P2 PAQ137 HOW MUCH PAID IN TUITION ($)	This data recoded for respondent confidentiality.
1848	A1TPREK	A1 Q4A TCH PREKINDERGARTEN LEVELS	This data suppressed for respondent confidentiality.
1849	A1TTRNK	A1 Q4B TCH TRANSITIONAL KINDERGARTEN	This data suppressed for respondent confidentiality.
1850	A1TREGK	A1 Q4C TEACHES REGULAR KINDERGARTEN	This data suppressed for respondent confidentiality.
1851	A1TPRE1	A1 Q4D TCH PRE-1ST GRADE LEVEL	This data suppressed for respondent confidentiality.
1852	A1T1ST	A1 Q4E TCH 1ST GRADE LEVEL	This data suppressed for respondent confidentiality.
1853	A1T2ND	A1 Q4F TCH 2ND GRADE LEVEL	This data suppressed for respondent confidentiality.

Table 7-7. Masked variables in the child catalog (continued)

FieldID	FieldName	FieldLabel	Comment
1854	A1T3RD	A1 Q4G TCH 3RD GRADE OR HIGHER LEVEL	This data suppressed for respondent confidentiality.
1855	A13YROL	A1 Q5A HOW MANY 3-YEAR-OLDS IN CLASS	This data recoded for respondent confidentiality.
1856	A14YROL	A1 Q5B HOW MANY 4-YEAR-OLDS IN CLASS	This data recoded for respondent confidentiality.
1857	A15YROL	A1 Q5C HOW MANY 5-YEAR-OLDS IN CLASS	This data recoded for respondent confidentiality.
1858	A16YROL	A1 Q5D HOW MANY 6-YEAR-OLDS IN CLASS	This data recoded for respondent confidentiality.
1859	A17YROL	A1 Q5E HOW MANY 7-YEAR-OLDS IN CLASS	This data recoded for respondent confidentiality.
1860	A18YROL	A1 Q5F HOW MANY 8-YEAR-OLDS IN CLASS	This data recoded for respondent confidentiality.
1861	A19YROL	A1 Q5G HOW MANY 9-YEAR-OLDS IN CLASS	This data recoded for respondent confidentiality.
1862	A1TOTAG	A1 Q5H TOTAL CLASS ENROLLMENT (AGE)	This data recoded for respondent confidentiality.
1863	A1ASIAN	A1 Q6A # OF ASIAN/PACIFIC ISLANDERS	This data recoded for respondent confidentiality.
1864	A1HISP	A1 Q6B # OF HISPANICS (ALL RACES)	This data recoded for respondent confidentiality.
1865	A1BLACK	A1 Q6C # OF NON-HISPANIC BLACKS	This data recoded for respondent confidentiality.
1866	A1WHITE	A1 Q6D # OF NON-HISPANIC WHITES	This data recoded for respondent confidentiality.
1867	A1AMRIN	A1 Q6E # OF AMERICAN INDIANS	This data recoded for respondent confidentiality.
1868	A1RACEO	A1 Q6F # OF STUDENTS OF OTHER RACES	This data recoded for respondent confidentiality.
1869	A1TOTRA	A1 Q6 TOTAL CLASS ENROLLMENT (RACES)	This data recoded for respondent confidentiality.
1896	A1TVTNM	A1 Q18C TCHRS SPEAK VIETNAMESE	This data suppressed for respondent confidentiality.
1897	A1TCHNS	A1 Q18D TCHRS SPEAK CHINESE	This data suppressed for respondent confidentiality.
1898	A1TJPNS	A1 Q18E TCHRS SPEAK JAPANESE	This data suppressed for respondent confidentiality.
1899	A1TKRN	A1 Q18F TCHRS SPEAK KOREAN	This data suppressed for respondent confidentiality.
1900	A1TFLPN	A1 Q18G TCHRS SPEAK A FILIPINO LNG	This data suppressed for respondent confidentiality.
1901	A1TOTAS	A1 Q18H TCHRS SPEAK OTHER ASIAN LNG	This data suppressed for respondent confidentiality.
1902	A1TOTLG	A1 Q18I TCHRS SPEAK ANOTHER LANGUAGE	This data suppressed for respondent confidentiality.
1910	A2GIFT	A2 Q2A # CLASSIFIED AS GFTED/TALENTED	This data recoded for respondent confidentiality.
1911	A2PRTGF	A2 Q2B # TAKE PART IN GIFTED/TALENTED	This data recoded for respondent confidentiality.
1916	A2DISAB	A2 Q3A NUMBER WITH DISABILITIES	This data recoded for respondent confidentiality.
1920	A2RETAR	A2 Q4D MENTAL RETARDATION	This data recoded for respondent confidentiality.
1926	A2MULTI	A2 Q4J MULTIPLE DISABILITIES	This data suppressed for respondent confidentiality.
1927	A2AUTSM	A2 Q4K AUTISM	This data suppressed for respondent confidentiality.
1929	A2DEAF	A2 Q4M DEAFNESS AND BLINDNESS	This data suppressed for respondent confidentiality.
1938	A2CVTNM	A2 Q7C STUDENTS SPEAK VIETNAMESE	This data suppressed for respondent confidentiality.
1939	A2CCHNS	A2 Q7D STUDENTS SPEAK CHINESE	This data suppressed for respondent confidentiality.
1940	A2CJPNS	A2 Q7E STUDENTS SPEAK JAPANESE	This data suppressed for respondent confidentiality.

Table 7-7. Masked variables in the child catalog (continued)

FieldID	FieldName	FieldLabel	Comment
1941	A2CKRN	A2 Q7F STUDENTS SPEAK KOREAN	This data suppressed for respondent confidentiality.
1942	A2CFLPN	A2 Q7G STUDENTS SPEAK FILIPINO LNG	This data suppressed for respondent confidentiality.
1943	A2OTASN	A2 Q7H STUDENTS SPEAK OTHER ASIAN LNG	This data suppressed for respondent confidentiality.
1944	A2OTLNG	A2 Q7I STUDENTS SPEAK ANOTHER LNG	This data suppressed for respondent confidentiality.
1945	A2LNGOS	A2 Q7I LANGUAGE OF INSTRUCTION - OTHER	This data suppressed for respondent confidentiality.
2255	B1YRBORN	B1 Q16 TEACHER'S YEAR OF BIRTH	This data recoded for respondent confidentiality.
2257	B1RACE1	B1 Q18 NATIVE AMERICAN OR PACIF ISLANDER	This data recoded for respondent confidentiality.
2260	B1RACE4	B1 Q18 NATIVE HAWAIIAN OR OTHER PAC IS	This data suppressed for respondent confidentiality.
2262	B1YRSPRE	B1 Q19A YRS TEACHER TAUGHT PRESCHOOL	This data recoded for respondent confidentiality.
2263	B1YRSKIN	B1 Q19B YRS TEACHER TAUGHT KINDERGARTEN	This data recoded for respondent confidentiality.
2264	B1YRSFST	B1 Q19C YRS TEACHER TAUGHT FIRST GRADE	This data recoded for respondent confidentiality.
2265	B1YRS2T5	B1 Q19D YRS TEACHER TAUGHT 2 TO 5 GRADE	This data recoded for respondent confidentiality.
2266	B1YRS6PL	B1 Q19E YRS TEACHER TAUGHT 6 GRADE OR UP	This data recoded for respondent confidentiality.
2267	B1YRSESL	B1 Q19F YRS TEACHER TAUGHT ESL	This data recoded for respondent confidentiality.
2268	B1YRSBIL	B1 Q19G YRS TEACHER TAUGHT BILINGUAL ED	This data recoded for respondent confidentiality.
2269	B1YRSSPE	B1 Q19H YRS TEACHER TAUGHT SPECIAL ED	This data recoded for respondent confidentiality.
2270	B1YRSPE	B1 Q19I YRS TEACHER TAUGHT PHYSICAL ED	This data recoded for respondent confidentiality.
2271	B1YRSART	B1 Q19J YRS TEACHER TAUGHT ART OR MUSIC	This data recoded for respondent confidentiality.
2272	B1YRSCH	B1 Q20 YRS TEACHER TAUGHT AT THIS SCHOOL	This data recoded for respondent confidentiality.
2273	B1HGHSTD	B1 Q21 HIGHEST ED LEVEL TEACHER ACHIEVED	This data recoded for respondent confidentiality.
2364	B2YRBORN	B2 Q16 TEACHER'S YEAR OF BIRTH	This data recoded for respondent confidentiality.
2366	B2RACE1	B1 Q18 NATIVE AMERICAN OR PACIF ISLANDER	This data recoded for respondent confidentiality.
2369	B2RACE4	B2 Q18 NATIVE HAWAIIAN OR OTHER PAC IS	This data suppressed for respondent confidentiality.
2371	B2YRSPRE	B2 Q19A YRS TEACHER TAUGHT PRESCHOOL	This data recoded for respondent confidentiality.
2372	B2YRSKIN	B2 Q19B YRS TEACHER TAUGHT KINDERGARTEN	This data recoded for respondent confidentiality.
2373	B2YRSFST	B2 Q19C YRS TEACHER TAUGHT FIRST GRADE	This data recoded for respondent confidentiality.
2374	B2YRS2T5	B2 Q19D YRS TEACHER TAUGHT 2 TO 5 GRADE	This data recoded for respondent confidentiality.
2375	B2YRS6PL	B2 Q19E YRS TEACHER TAUGHT 6 GRADE OR UP	This data recoded for respondent confidentiality.
2376	B2YRSESL	B2 Q19F YRS TEACHER TAUGHT ESL	This data recoded for respondent confidentiality.
2377	B2YRSBIL	B2 Q19G YRS TEACHER TAUGHT BILINGUAL ED	This data recoded for respondent confidentiality.
2378	B2YRSSPE	B2 Q19H YRS TEACHER TAUGHT SPECIAL ED	This data recoded for respondent confidentiality.
2379	B2YRSPE	B2 Q19I YRS TEACHER TAUGHT PHYSICAL ED	This data recoded for respondent confidentiality.
2380	B2YRSART	B2 Q19J YRS TEACHER TAUGHT ART OR MUSIC	This data recoded for respondent confidentiality.
2381	B2YRSCH	B2 Q20 YRS TEACHER TAUGHT AT THIS SCHOOL	This data recoded for respondent confidentiality.

Table 7-7. Masked variables in the child catalog (continued)

FieldID	FieldName	FieldLabel	Comment
2382	B2HGHSTD	B2 Q21 HIGHEST ED LEVEL TEACHER ACHIEVED	This data recoded for respondent confidentiality.
2476	S2NUMDAY	S2 Q1 NUMBER OF DAYS MUST ATTEND	This data recoded for respondent confidentiality.
2477	S2ADA	S2 Q2 % AVERAGE DAILY ATTENDANCE FOR YR.	This data recoded for respondent confidentiality.
2478	S2ANUMCH	S2 Q3A # ENROLLED AROUND 10/1/1998	This data recoded for respondent confidentiality.
2479	S2BNUMCH	S2 Q3B # ENROLLED SINCE 10/1/1998	This data recoded for respondent confidentiality.
2480	S2CNUMCH	S2 Q3C # WHO LEFT - DIDN'T RETURN	This data recoded for respondent confidentiality.
2503	S2BIASKL	S2 Q6D IS IT A TRIBAL SCHOOL	This data suppressed for respondent confidentiality.
2504	S2SPEDSK	S2 Q6E IS IT A SPECIAL ED SCHOOL	This data suppressed for respondent confidentiality.
2509	S2PRIVRD	S2 Q7A4 IS IT A PRIVATE ORDER	This data suppressed for respondent confidentiality.
2511	S2NAISKL	S2 Q7C PRIVATE-ACCREDITED BY NAIS	This data suppressed for respondent confidentiality.
2513	S2PVTSPD	S2 Q7E IS IT SPECIAL EDUCATION	This data suppressed for respondent confidentiality.
2516	S2EMPHAS	S2 Q9 WHAT'S THE SCHOOL EMPHASIS	This data suppressed for respondent confidentiality.
2517	S2PREK	S2 Q10 PRE-K PARTICIPATES	This data suppressed for respondent confidentiality.
2518	S2REDDYN	S2 Q10 TRANSITIONAL PARTICIPATES	This data suppressed for respondent confidentiality.
2519	S2KINGAR	S2 Q10 KINDERGARTEN PARTICIPATES	This data suppressed for respondent confidentiality.
2520	S2PRE1ST	S2 Q10 PREFIRST PARTICIPATES	This data suppressed for respondent confidentiality.
2521	S2FIR1ST	S2 Q10 1ST GRADE PARTICIPATES	This data suppressed for respondent confidentiality.
2522	S2SEC2ND	S2 Q10 2ND GRADE PARTICIPATES	This data suppressed for respondent confidentiality.
2523	S2THI3RD	S2 Q10 3RD GRADE PARTICIPATES	This data suppressed for respondent confidentiality.
2524	S2FOR4TH	S2 Q10 4TH GRADE PARTICIPATES	This data suppressed for respondent confidentiality.
2525	S2FIF5TH	S2 Q10 5TH GRADE PARTICIPATES	This data suppressed for respondent confidentiality.
2526	S2SIX6TH	S2 Q10 6TH GRADE PARTICIPATES	This data suppressed for respondent confidentiality.
2527	S2SEV7TH	S2 Q10 7TH GRADE PARTICIPATES	This data suppressed for respondent confidentiality.
2528	S2ATE8TH	S2 Q10 8TH GRADE PARTICIPATES	This data suppressed for respondent confidentiality.
2529	S2NIN9TH	S2 Q10 9TH GRADE PARTICIPATES	This data suppressed for respondent confidentiality.
2530	S2TEN10	S2 Q10 10TH GRADE PARTICIPATES	This data suppressed for respondent confidentiality.
2531	S2ELEV11	S2 Q10 11TH GRADE PARTICIPATES	This data suppressed for respondent confidentiality.
2532	S2TWEL12	S2 Q10 12TH GRADE PARTICIPATES	This data suppressed for respondent confidentiality.
2533	S2SPENR	S2 Q11 NUMBER IN SPECIAL PROGRAM	This data suppressed for respondent confidentiality.
2550	S2PCTHSP	S2 Q14 PERCENT OF HISPANIC STUDENTS	This data recoded for respondent confidentiality.
2551	S2INDPCT	S2 Q15A PERCENT OF NATIVE AMERICANS	This data suppressed for respondent confidentiality.
2552	S2ASNPCT	S2 Q15B PERCENT OF ASIAN STUDENTS	This data suppressed for respondent confidentiality.
2553	S2BLKPCT	S2 Q15C PERCENT OF BLACK STUDENTS	This data recoded for respondent confidentiality.
2554	S2PCFPCT	S2 Q15D PERCENT OF HAWAIIAN STUDENT	This data suppressed for respondent confidentiality.

Table 7-7. Masked variables in the child catalog (continued)

FieldID	FieldName	FieldLabel	Comment
2555	S2WHTPCT	S2 Q15E PERCENT OF WHITE STUDENTS	This data suppressed for respondent confidentiality.
2556	S2ELIBRK	S2 Q16A1 ELIGIBLE FOR FREE BREAKFAST	This data suppressed for respondent confidentiality.
2557	S2PARBRK	S2 Q16A2 PARTICIPATES IN BREAKFAST	This data suppressed for respondent confidentiality.
2558	S2ELILNC	S2 Q16B1 ELIGIBLE FOR FREE LUNCH	This data suppressed for respondent confidentiality.
2559	S2PARLNC	S2 Q16B2 PARTICIPATES IN FREE LUNCH	This data suppressed for respondent confidentiality.
2560	S2ELIRED	S2 Q16C1 ELIGIBLE IN REDUCED-PRICE LUNCH	This data suppressed for respondent confidentiality.
2561	S2PARRED	S2 Q16C2 PARTICIPATES IN RED-PRICE LUNCH	This data suppressed for respondent confidentiality.
2571	S2CHSRVD	S2 Q20 NUMBER SERVED BY TITLE 1	This data recoded for respondent confidentiality.
2572	S2HDCHDN	S2 Q21A1 HALFDAY - NUMBER OF CHILDREN	This data recoded for respondent confidentiality.
2573	S2HDCLS	S2 Q21B1 HALFDAY - NUMBER OF CLASSES	This data recoded for respondent confidentiality.
2574	S2HDDAYS	S2 Q21C1 HALFDAY - NUMBER OF DAYS PER WK	This data suppressed for respondent confidentiality.
2576	S2FDCHDN	S2 Q21A2 FULLDAY - NUMBER OF CHILDREN	This data recoded for respondent confidentiality.
2577	S2FDCLS	S2 Q21B2 FULLDAY - NUMBER OF CLASSES	This data recoded for respondent confidentiality.
2578	S2FDDAYS	S2 Q21C2 FULLDAY - NUMBER OF DAYS PER WK	This data suppressed for respondent confidentiality.
2580	S2KCCHDN	S2 Q21A3 K COMBINATION - # OF CHILDREN	This data recoded for respondent confidentiality.
2581	S2KCCLS	S2 Q21B3 K COMBINATION - # OF CLASSES	This data suppressed for respondent confidentiality.
2582	S2KCDAYS	S2 Q21C3 K COMBINATION - # DAYS PER WK	This data suppressed for respondent confidentiality.
2584	S2TKCHDN	S2 Q21A4 TRANSITIONAL K - # OF CHILDREN	This data recoded for respondent confidentiality.
2585	S2TKCLS	S2 Q21B4 TRANSITIONAL K - # OF CLASSES	This data suppressed for respondent confidentiality.
2586	S2TKDAYS	S2 Q21C4 TRANSITIONAL K - # DAYS PER WK	This data suppressed for respondent confidentiality.
2588	S2T1CHDN	S2 Q21A5 TRANSITIONAL 1ST- # OF CHILDREN	This data recoded for respondent confidentiality.
2589	S2T1CLS	S2 Q21B5 TRANSITIONAL 1ST- # OF CLASSES	This data suppressed for respondent confidentiality.
2590	S2T1DAYS	S2 Q21C5 TRANSITIONAL 1ST - DAYS PER WK	This data suppressed for respondent confidentiality.
2595	S2NOCUTO	S2 Q22A NO CUTOFF DATE TO TURN FIVE	This data suppressed for respondent confidentiality.
2607	S2CHLDNM	S2 Q24 # OF CHILDREN SITE ACCOMMODATES	This data recoded for respondent confidentiality.
2608	S2RMNUM	S2 Q25 NUMBER OF INSTRUCTIONAL ROOMS	This data recoded for respondent confidentiality.
2619	S2INSTCM	S2 Q27A # INSTRUCTIONAL COMPUTERS	This data recoded for respondent confidentiality.
2620	S2ADMNCM	S2 Q27B # INSTRUCT/ADMIN COMPUTERS	This data recoded for respondent confidentiality.
2621	S2TOTCM	S2 Q27C TOTAL NUMBER OF COMPUTERS	This data recoded for respondent confidentiality.
2624	S2LANRMS	S2 Q28A3 TOTAL ROOMS W/LAN ACCESS	This data recoded for respondent confidentiality.
2627	S2CDRRMS	S2 Q28B3 TOTAL ROOMS W/CD-ROM ACCESS	This data recoded for respondent confidentiality.
2630	S2NETRMS	S2 Q28C3 TOTAL RMS W/INTERNET ACCESS	This data recoded for respondent confidentiality.
2634	S2CMNITY	S2 Q30 COMMUNITY SCHOOL IS LOCATED IN	This data recoded for respondent confidentiality.
2690	S2NMBKPT	S2 Q40 NUMBER RETAINED GRADE-K	This data recoded for respondent confidentiality.

Table 7-7. Masked variables in the child catalog (continued)

FieldID	FieldName	FieldLabel	Comment
2732	S2NMADMN	S2 Q47A2 NUMBER OF ADMINISTRATORS	This data recoded for respondent confidentiality.
2734	S2NMTECH	S2 Q47B2 NUMBER OF TEACHERS	This data recoded for respondent confidentiality.
2736	S2NMDSTR	S2 Q47C2 NUMBER FROM DISTRICT OFFICE	This data recoded for respondent confidentiality.
2738	S2NUMBRD	S2 Q47D2 NUMBER OF BOARD MEMBERS	This data recoded for respondent confidentiality.
2740	S2NMOMDD	S2 Q47E2 NUMBER OF PARENTS	This data recoded for respondent confidentiality.
2742	S2NMREPS	S2 Q47F2 NUMBER OF REPRESENTATIVES	This data recoded for respondent confidentiality.
2744	S2NUMOTH	S2 Q47G2 NUMBER OF OTHER GROUPS	This data recoded for respondent confidentiality.
2748	S2BILING	S2 Q50 BILINGUAL SERVICES PERCENT	This data suppressed for respondent confidentiality.
2749	S2ESLONL	S2 Q50 ESL SERVICES PERCENT	This data suppressed for respondent confidentiality.
2750	S2ESLBIL	S2 Q50 ESL AND BILINGUAL PERCENT	This data suppressed for respondent confidentiality.
2751	S2YRSESL	S2 Q51A YEARS FOR ESL SERVICES	This data recoded for respondent confidentiality.
2752	S2YRSBIL	S2 Q51B YEARS FOR BILINGUAL SERVICES	This data recoded for respondent confidentiality.
2753	S2SRVBTH	S2 Q51C YEARS FOR ESL & BILINGUAL SER	This data recoded for respondent confidentiality.
2777	S2GFTNBR	S2 Q57 NUMBER GIFTED/TALENTED	This data suppressed for respondent confidentiality.
2778	S2ADMFTE	S2 Q58A FTE ADMINISTRATIVE STAFF	This data recoded for respondent confidentiality.
2779	S2TCHFTE	S2 Q58B FTE CLASSROOM TEACHERS	This data recoded for respondent confidentiality.
2780	S2GYMFTE	S2 Q58C FTE GYM/DRAMA/ART TEACHERS	This data recoded for respondent confidentiality.
2781	S2CMPFTE	S2 Q58D FTE COMPUTER SPECIALISTS	This data recoded for respondent confidentiality.
2782	S2HDSFTE	S2 Q58E FTE HEAD START/PRE-K TEACHERS	This data recoded for respondent confidentiality.
2783	S2CRDFTE	S2 Q58F FTE COORDINATOR/SPECIALIST	This data recoded for respondent confidentiality.
2784	S2TT1FTE	S2 Q58G FTE TITLE 1 TEACHERS	This data recoded for respondent confidentiality.
2785	S2AIDFTE	S2 Q58H FTE TEACHER AIDES	This data recoded for respondent confidentiality.
2786	S2SPEFTE	S2 Q58I FTE SPECIAL ED TEACHERS	This data recoded for respondent confidentiality.
2787	S2SPAFTE	S2 Q58J FTE SPECIAL ED AIDES	This data recoded for respondent confidentiality.
2788	S2ESLFTE	S2 Q58K FTE ESL-BILINGUAL TEACHERS	This data recoded for respondent confidentiality.
2789	S2ESAFTE	S2 Q58L FTE ESL-BILINGUAL AIDES	This data recoded for respondent confidentiality.
2790	S2LIAFTE	S2 Q58M FTE PARENT LIAISONS	This data recoded for respondent confidentiality.
2791	S2COUFTE	S2 Q58N FTE SOCIAL WORKERS	This data recoded for respondent confidentiality.
2792	S2LIBFTE	S2 Q58O LIBRARY/MEDIA SPECIALISTS	This data recoded for respondent confidentiality.
2793	S2THRFTE	S2 Q58P SPEECH/PHYSICAL THERAPISTS	This data recoded for respondent confidentiality.
2794	S2NRSFTE	S2 Q58Q FTE SCHOOL NURSES	This data recoded for respondent confidentiality.
2795	S2NONFTE	S2 Q58R FTE NON-INSTRUCTIONAL STAFF	This data recoded for respondent confidentiality.
2796	S2REAFTE	S2 Q58S FTE READING RECOVERY TCHRS	This data suppressed for respondent confidentiality.
2797	S2MSCFTE	S2 Q58T FTE MATH/SCIENCE SPECIALISTS	This data recoded for respondent confidentiality.

Table 7-7. Masked variables in the child catalog (continued)

FieldID	FieldName	FieldLabel	Comment		
2798	S2FRLFTE	S2 Q58U FTE FOREIGN LNG SPECIALISTS	This data recoded for respondent confidentiality.		
2799	S2OTHFTE	S2 Q58V OTHER FTE	This data suppressed for respondent confidentiality.		
2800	S2FTETOT	S2 Q58W TOTAL FTE SCHOOL STAFF	This data recoded for respondent confidentiality.		
2803	S2ETHNIC	S2 Q61 % HISPANIC/LATINO TEACHERS	This data recoded for respondent confidentiality.		
2804	S2Q62IND	S2 Q62 % AMERICAN NATIVE TEACHERS	This data recoded for respondent confidentiality.		
2805	S2Q62ASN	S2 Q62 % ASIAN TEACHERS	This data recoded for respondent confidentiality.		
2806	S2Q62BLK	S2 Q62 % BLACK TEACHERS	This data recoded for respondent confidentiality.		
2807	S2Q62HAW	S2 Q62 % HAWAIIAN TEACHERS	This data suppressed for respondent confidentiality.		
2904	S2YSTCH	S2 Q74A NUMBER OF YRS TEACHING	This data recoded for respondent confidentiality.		
2905	S2TOTPRI	S2 Q74B NUMBER OF YRS AS PRINCIPAL	This data recoded for respondent confidentiality.		
2906	S2PRINHR	S2 Q74C NUMBER YRS A PRINCIPAL HERE	This data recoded for respondent confidentiality.		
2926	S2EDLVL	S2 Q77 HIGHEST LEVEL OF EDUCATION	This data recoded for respondent confidentiality.		
	New Variables Replacing Suppressed Variables			**Suppressed Variables**	
3989	P1LANGS1	P1 PLQ040 OTHER LANGUAGE - SPANISH 1=Yes 2=No -1=Not applicable -7=Refused -8=Don't know -9=Not ascertained	This data recoded for respondent confidentiality.	P1LANG12	P1 PLQ040 OTHER LANGUAGE - SPANISH
3990	P1LANGS2	P1 PLQ040 OTHER-OTHER EUROPEAN LANGS 1=Yes 2=No -1=Not applicable -7=Refused -8=Don't know -9=Not ascertained	This data recoded for respondent confidentiality.	P1LANG4 P1LANG5 P1LANG6 P1LANG7 P1LANG10 P1LANG11 P1LANG16 P1LANG20	P1 PLQ040 OTHER LANGUAGE - FRENCH P1 PLQ040 OTHER LANGUAGE - GERMAN P1 PLQ040 OTHER LANGUAGE - GREEK P1 PLQ040 OTHER LANGUAGE - ITALIAN P1 PLQ040 OTHER LANGUAGE - POLISH P1 PLQ040 OTHER LANGUAGE - PORTUGUESE P1 PLQ040 OTHER-EASTERN EUROPEAN LANGS P1 PLQ040 OTHER-WESTERN EUROPEAN LANGS

Table 7-7. Masked variables in the child catalog (continued)

	New Variables Replacing Suppressed Variables			Suppressed Variables	
3991	P1LANGS3	P1 PLQ040 OTHER-ASIAN LANGUAGES 1=Yes 2=No -1=Not applicable -7=Refused -8=Don't know -9=Not ascertained	This data recoded for respondent confidentiality.	P1LANG2 P1LANG3 P1LANG8 P1LANG9 P1LANG13 P1LANG17 P1LANG21 P1LANG22 P1LANG23	P1 PLQ040 OTHER LANGUAGE - CHINESE P1 PLQ040 OTHER LANGUAGE - FILIPINO P1 PLQ040 OTHER LANGUAGE - JAPANESE P1 PLQ040 OTHER LANGUAGE - KOREAN P1 PLQ040 OTHER LANGUAGE - VIETNAMESE P1 PLQ040 OTHER-NATIVE AMERICAN LANGS P1 PLQ040 OTHER-INDIAN SUBCONTINTL LANGS P1 PLQ040 OTHER-SOUTHEAST ASIAN LANGS P1 PLQ040 OTHER-PACIFIC ISLANDER LANGS
3992	P1LANGS4	P1 PLQ040 OTHER LANGUAGE - OTHER LANGS 1=Yes 2=No -1=Not applicable -7=Refused -8=Don't know -9=Not ascertained	This data recoded for respondent confidentiality.	P1LANG1 P1LANG14 P1LANG15 P1LANG18 P1LANG19	P1 PLQ040 OTHER LANGUAGE - ARABIC P1 PLQ040 OTHER LANGUAGE - SPECIFY P1 PLQ040 OTHER-AFRICAN LANGUAGES P1 PLQ040 OTHER-SIGN LANGUAGES P1 PLQ040 OTHER-MIDDLE EASTERN LANGS
3993	A1KGTYP	A1 Q3 TYPE OF KINDERGARTEN PROG TCH 1=Regular kindergarten only 2=Other kindergarten type	This data recoded for respondent confidentiality.	A1REGK A12YRK1 A12YRK2 A1TRNK A1PR1ST A1UNGR A1MULGR	A1 Q3A TCH REGULAR 1-YR KINDERGARTEN A1 Q3B TEACHES 1ST YR OF 2-YR K A1 Q3C TEACHES 2ND YR OF 2-YR K A1 Q3D TCH TRANSITIONAL KINDERGARTEN A1 Q3E TEACHES PRE-1ST GRADE AFTER K A1 Q3F TEACHES UNGRADED CLASS A1 Q3G TEACHES MULTIGRADE CLASS
3994	S2PRACE	S2 Q73 PRINCIPAL'S RACE 1=White, non-Hispanic 2=Black, or African-American, non-Hispanic 3=Other race -9=Not Ascertained	This data recoded for respondent confidentiality.	S2RACE1 S2RACE2 S2RACE3 S2RACE4 S2RACE5	S2 Q73A PRINCIPAL IS AMERICAN INDIAN S2 Q73B PRINCIPAL IS ASIAN S2 Q73C PRINCIPAL IS BLACK S2 Q73D PRINCIPAL IS HAWAIIAN OR PAC IS S2 Q73E PRINCIPAL IS WHITE

Table 7-8. Masked variables in the teacher catalog

FieldID	FieldName	FieldLabel	Comment
3	KURBAN	LOCATION TYPE IN SAMPLE FRAME	This data recoded for respondent confidentiality.
5	CS_TYPE2	TYPE OF SCHOOL IN SAMPLE FRAME	This data recoded for respondent confidentiality.
20	S2KENRLS	S2 TOTAL SCHOOL ENROLLMENT	This data recoded for respondent confidentiality.
22	S2KMINOR	S2 PERCENT MINORITY STUDENTS	This data recoded for respondent confidentiality.
24	A1APBLK	A1 PERCENT OF BLACKS IN CLASS-AM	This data recoded for respondent confidentiality.
27	A1PPBLK	A1 PERCENT OF BLACKS IN CLASS-PM	This data recoded for respondent confidentiality.
30	A1DPBLK	A1 PERCENT OF BLACKS IN CLASS-AD	This data recoded for respondent confidentiality.
33	B1AGE	B1 TEACHER'S AGE	This data recoded for respondent confidentiality.
43	A1ATPREK	A1 Q4A TCH PREKINDERGARTEN LEVELS-AM	This data suppressed for respondent confidentiality.
44	A1ATTRNK	A1 Q4B TCH TRANSITIONAL KINDERGARTEN-AM	This data suppressed for respondent confidentiality.
45	A1ATREGK	A1 Q4C TEACHES REGULAR KINDERGARTEN-AM	This data suppressed for respondent confidentiality.
46	A1ATPRE1	A1 Q4D TCH PRE-1ST GRADE LEVEL-AM	This data suppressed for respondent confidentiality.
47	A1AT1ST	A1 Q4E TCH 1ST GRADE LEVEL-AM	This data suppressed for respondent confidentiality.
48	A1AT2ND	A1 Q4F TCH 2ND GRADE LEVEL-AM	This data suppressed for respondent confidentiality.
49	A1AT3RD	A1 Q4G TCH 3RD GRADE OR HIGHER LEVEL-AM	This data suppressed for respondent confidentiality.
50	A1A3YROL	A1 Q5A HOW MANY 3-YEAR-OLDS IN CLASS-AM	This data recoded for respondent confidentiality.
51	A1A4YROL	A1 Q5B HOW MANY 4-YEAR-OLDS IN CLASS-AM	This data recoded for respondent confidentiality.
52	A1A5YROL	A1 Q5C HOW MANY 5-YEAR-OLDS IN CLASS-AM	This data recoded for respondent confidentiality.
53	A1A6YROL	A1 Q5D HOW MANY 6-YEAR-OLDS IN CLASS-AM	This data recoded for respondent confidentiality.
54	A1A7YROL	A1 Q5E HOW MANY 7-YEAR-OLDS IN CLASS-AM	This data recoded for respondent confidentiality.
55	A1A8YROL	A1 Q5F HOW MANY 8-YEAR-OLDS IN CLASS-AM	This data recoded for respondent confidentiality.
56	A1A9YROL	A1 Q5G HOW MANY 9-YEAR-OLDS IN CLASS-AM	This data recoded for respondent confidentiality.
57	A1ATOTAG	A1 Q5H TOTAL CLASS ENROLLMENT (AGE)-AM	This data recoded for respondent confidentiality.
58	A1AASIAN	A1 Q6A # OF ASIAN/PACIFIC ISLANDERS-AM	This data recoded for respondent confidentiality.
59	A1AHISP	A1 Q6B # OF HISPANICS (ALL RACES)-AM	This data recoded for respondent confidentiality.
60	A1ABLACK	A1 Q6C # OF NON-HISPANIC BLACKS-AM	This data recoded for respondent confidentiality.
61	A1AWHITE	A1 Q6D # OF NON-HISPANIC WHITES-AM	This data recoded for respondent confidentiality.
62	A1AAMRIN	A1 Q6E # OF AMERICAN INDIANS-AM	This data recoded for respondent confidentiality.
63	A1ARACEO	A1 Q6F # OF STUDENTS OF OTHER RACES-AM	This data recoded for respondent confidentiality.
64	A1ATOTRA	A1 Q6 TOTAL CLASS ENROLLMENT (RACES)-AM	This data recoded for respondent confidentiality.
91	A1ATVTNM	A1 Q18C TCHRS SPEAK VIETNAMESE-AM	This data suppressed for respondent confidentiality.
92	A1ATCHNS	A1 Q18D TCHRS SPEAK CHINESE-AM	This data suppressed for respondent confidentiality.
93	A1ATJPNS	A1 Q18E TCHRS SPEAK JAPANESE-AM	This data suppressed for respondent confidentiality.

Table 7-8. Masked variables in the teacher catalog (continued)

FieldID	FieldName	FieldLabel	Comment
94	A1ATKRN	A1 Q18F TCHRS SPEAK KOREAN-AM	This data suppressed for respondent confidentiality.
95	A1ATFLPN	A1 Q18G TCHRS SPEAK A FILIPINO LNG-AM	This data suppressed for respondent confidentiality.
96	A1ATOTAS	A1 Q18H TCHRS SPEAK OTHER ASIAN LNG-AM	This data suppressed for respondent confidentiality.
97	A1ATOTLG	A1 Q18I TCHRS SPEAK ANOTHER LANGUAGE-AM	This data suppressed for respondent confidentiality.
109	A1PTPREK	A1 Q4A TCH PREKINDERGARTEN LEVELS-PM	This data suppressed for respondent confidentiality.
110	A1PTTRNK	A1 Q4B TCH TRANSITIONAL KINDERGARTEN-PM	This data suppressed for respondent confidentiality.
111	A1PTREGK	A1 Q4C TEACHES REGULAR KINDERGARTEN-PM	This data suppressed for respondent confidentiality.
112	A1PTPRE1	A1 Q4D TCH PRE-1ST GRADE LEVEL-PM	This data suppressed for respondent confidentiality.
113	A1PT1ST	A1 Q4E TCH 1ST GRADE LEVEL-PM	This data suppressed for respondent confidentiality.
114	A1PT2ND	A1 Q4F TCH 2ND GRADE LEVEL-PM	This data suppressed for respondent confidentiality.
115	A1PT3RD	A1 Q4G TCH 3RD GRADE OR HIGHER LEVEL-PM	This data suppressed for respondent confidentiality.
116	A1P3YROL	A1 Q5A HOW MANY 3-YEAR-OLDS IN CLASS-PM	This data recoded for respondent confidentiality.
117	A1P4YROL	A1 Q5B HOW MANY 4-YEAR-OLDS IN CLASS-PM	This data recoded for respondent confidentiality.
118	A1P5YROL	A1 Q5C HOW MANY 5-YEAR-OLDS IN CLASS-PM	This data recoded for respondent confidentiality.
119	A1P6YROL	A1 Q5D HOW MANY 6-YEAR-OLDS IN CLASS-PM	This data recoded for respondent confidentiality.
120	A1P7YROL	A1 Q5E HOW MANY 7-YEAR-OLDS IN CLASS-PM	This data recoded for respondent confidentiality.
121	A1P8YROL	A1 Q5F HOW MANY 8-YEAR-OLDS IN CLASS-PM	This data recoded for respondent confidentiality.
122	A1P9YROL	A1 Q5G HOW MANY 9-YEAR-OLDS IN CLASS-PM	This data recoded for respondent confidentiality.
123	A1PTOTAG	A1 Q5H TOTAL CLASS ENROLLMENT (AGE)-PM	This data recoded for respondent confidentiality.
124	A1PASIAN	A1 Q6A # OF ASIAN/PACIFIC ISLANDERS-PM	This data recoded for respondent confidentiality.
125	A1PHISP	A1 Q6B # OF HISPANICS (ALL RACES)-PM	This data recoded for respondent confidentiality.
126	A1PBLACK	A1 Q6C # OF NON-HISPANIC BLACKS-PM	This data recoded for respondent confidentiality.
127	A1PWHITE	A1 Q6D # OF NON-HISPANIC WHITES-PM	This data recoded for respondent confidentiality.
128	A1PAMRIN	A1 Q6E # OF AMERICAN INDIANS-PM	This data recoded for respondent confidentiality.
129	A1PRACEO	A1 Q6F # OF STUDENTS OF OTHER RACES-PM	This data recoded for respondent confidentiality.
130	A1PTOTRA	A1 Q6 TOTAL CLASS ENROLLMENT (RACES)-PM	This data recoded for respondent confidentiality.
157	A1PTVTNM	A1 Q18C TCHRS SPEAK VIETNAMESE-PM	This data suppressed for respondent confidentiality.
158	A1PTCHNS	A1 Q18D TCHRS SPEAK CHINESE-PM	This data suppressed for respondent confidentiality.
159	A1PTJPNS	A1 Q18E TCHRS SPEAK JAPANESE-PM	This data suppressed for respondent confidentiality.
160	A1PTKRN	A1 Q18F TCHRS SPEAK KOREAN-PM	This data suppressed for respondent confidentiality.
161	A1PTFLPN	A1 Q18G TCHRS SPEAK A FILIPINO LNG-PM	This data suppressed for respondent confidentiality.
162	A1PTOTAS	A1 Q18H TCHRS SPEAK OTHER ASIAN LNG-PM	This data suppressed for respondent confidentiality.
163	A1PTOTLG	A1 Q18I TCHRS SPEAK ANOTHER LANGUAGE-PM	This data suppressed for respondent confidentiality.

Table 7-8. Masked variables in the teacher catalog (continued)

FieldID	FieldName	FieldLabel	Comment
175	A1DTPREK	A1 Q4A TCH PREKINDERGARTEN LEVELS-AD	This data suppressed for respondent confidentiality.
176	A1DTTRNK	A1 Q4B TCH TRANSITIONAL KINDERGARTEN-AD	This data suppressed for respondent confidentiality.
177	A1DTREGK	A1 Q4C TEACHES REGULAR KINDERGARTEN-AD	This data suppressed for respondent confidentiality.
178	A1DTPRE1	A1 Q4D TCH PRE-1ST GRADE LEVEL-AD	This data suppressed for respondent confidentiality.
179	A1DT1ST	A1 Q4E TCH 1ST GRADE LEVEL-AD	This data suppressed for respondent confidentiality.
180	A1DT2ND	A1 Q4F TCH 2ND GRADE LEVEL-AD	This data suppressed for respondent confidentiality.
181	A1DT3RD	A1 Q4G TCH 3RD GRADE OR HIGHER LEVEL-AD	This data suppressed for respondent confidentiality.
182	A1D3YROL	A1 Q5A HOW MANY 3-YEAR-OLDS IN CLASS-AD	This data recoded for respondent confidentiality.
183	A1D4YROL	A1 Q5B HOW MANY 4-YEAR-OLDS IN CLASS-AD	This data recoded for respondent confidentiality.
184	A1D5YROL	A1 Q5C HOW MANY 5-YEAR-OLDS IN CLASS-AD	This data recoded for respondent confidentiality.
185	A1D6YROL	A1 Q5D HOW MANY 6-YEAR-OLDS IN CLASS-AD	This data recoded for respondent confidentiality.
186	A1D7YROL	A1 Q5E HOW MANY 7-YEAR-OLDS IN CLASS-AD	This data recoded for respondent confidentiality.
187	A1D8YROL	A1 Q5F HOW MANY 8-YEAR-OLDS IN CLASS-AD	This data recoded for respondent confidentiality.
188	A1D9YROL	A1 Q5G HOW MANY 9-YEAR-OLDS IN CLASS-AD	This data recoded for respondent confidentiality.
189	A1DTOTAG	A1 Q5H TOTAL CLASS ENROLLMENT (AGE)-AD	This data recoded for respondent confidentiality.
190	A1DASIAN	A1 Q6A # OF ASIAN/PACIFIC ISLANDERS-AD	This data recoded for respondent confidentiality.
191	A1DHISP	A1 Q6B # OF HISPANICS (ALL RACES)-AD	This data recoded for respondent confidentiality.
192	A1DBLACK	A1 Q6C # OF NON-HISPANIC BLACKS-AD	This data recoded for respondent confidentiality.
193	A1DWHITE	A1 Q6D # OF NON-HISPANIC WHITES-AD	This data recoded for respondent confidentiality.
194	A1DAMRIN	A1 Q6E # OF AMERICAN INDIANS-AD	This data recoded for respondent confidentiality.
195	A1DRACEO	A1 Q6F # OF STUDENTS OF OTHER RACES-AD	This data recoded for respondent confidentiality.
196	A1DTOTRA	A1 Q6 TOTAL CLASS ENROLLMENT (RACES)-AD	This data recoded for respondent confidentiality.
223	A1DTVTNM	A1 Q18C TCHRS SPEAK VIETNAMESE-AD	This data suppressed for respondent confidentiality.
224	A1DTCHNS	A1 Q18D TCHRS SPEAK CHINESE-AD	This data suppressed for respondent confidentiality.
225	A1DTJPNS	A1 Q18E TCHRS SPEAK JAPANESE-AD	This data suppressed for respondent confidentiality.
226	A1DTKRN	A1 Q18F TCHRS SPEAK KOREAN-AD	This data suppressed for respondent confidentiality.
227	A1DTFLPN	A1 Q18G TCHRS SPEAK A FILIPINO LNG-AD	This data suppressed for respondent confidentiality.
228	A1DTOTAS	A1 Q18H TCHRS SPEAK OTHER ASIAN LNG-AD	This data suppressed for respondent confidentiality.
229	A1DTOTLG	A1 Q18I TCHRS SPEAK ANOTHER LANGUAGE-AD	This data suppressed for respondent confidentiality.
310	B1YRBORN	B1 Q16 TEACHER’S YEAR OF BIRTH	This data recoded for respondent confidentiality.
312	B1RACE1	B1 Q18 NATIVE AMERICAN OR PACIF ISLANDER	This data recoded for respondent confidentiality.
315	B1RACE4	B1 Q18 NATIVE HAWAIIAN OR OTHER PAC IS	This data suppressed for respondent confidentiality.
317	B1YRSPRE	B1 Q19A YRS TEACHER TAUGHT PRESCHOOL	This data recoded for respondent confidentiality.

Table 7-8. Masked variables in the teacher catalog (continued)

FieldID	FieldName	FieldLabel	Comment
318	B1YRSKIN	B1 Q19B YRS TEACHER TAUGHT KINDERGARTEN	This data recoded for respondent confidentiality.
319	B1YRSFST	B1 Q19C YRS TEACHER TAUGHT FIRST GRADE	This data recoded for respondent confidentiality.
320	B1YRS2T5	B1 Q19D YRS TEACHER TAUGHT 2 TO 5 GRADE	This data recoded for respondent confidentiality.
321	B1YRS6PL	B1 Q19E YRS TEACHER TAUGHT 6 GRADE OR UP	This data recoded for respondent confidentiality.
322	B1YRSESL	B1 Q19F YRS TEACHER TAUGHT ESL	This data recoded for respondent confidentiality.
323	B1YRSBIL	B1 Q19G YRS TEACHER TAUGHT BILINGUAL ED	This data recoded for respondent confidentiality.
324	B1YRSSPE	B1 Q19H YRS TEACHER TAUGHT SPECIAL ED	This data recoded for respondent confidentiality.
325	B1YRSPE	B1 Q19I YRS TEACHER TAUGHT PHYSICAL ED	This data recoded for respondent confidentiality.
326	B1YRSART	B1 Q19J YRS TEACHER TAUGHT ART OR MUSIC	This data recoded for respondent confidentiality.
327	B1YRSCH	B1 Q20 YRS TEACHER TAUGHT AT THIS SCHOOL	This data recoded for respondent confidentiality.
328	B1HGHSTD	B1 Q21 HIGHEST ED LEVEL TEACHER ACHIEVED	This data recoded for respondent confidentiality.
346	A2AGIFT	A2 Q2A # CLASSIFIED AS GFTED/TALENTED-AM	This data recoded for respondent confidentiality.
347	A2APRTGF	A2 Q2B # TAKE PART IN GIFTED/TALENTED-AM	This data recoded for respondent confidentiality.
352	A2ADISAB	A2 Q3A NUMBER WITH DISABILITIES-AM	This data recoded for respondent confidentiality.
356	A2ARETAR	A2 Q4D MENTAL RETARDATION-AM	This data recoded for respondent confidentiality.
362	A2AMULTI	A2 Q4J MULTIPLE DISABILITIES-AM	This data suppressed for respondent confidentiality.
363	A2AAUTSM	A2 Q4K AUTISM-AM	This data suppressed for respondent confidentiality.
365	A2ADEAF	A2 Q4M DEAFNESS AND BLINDNESS-AM	This data suppressed for respondent confidentiality.
374	A2ACVTNM	A2 Q7C STUDENTS SPEAK VIETNAMESE-AM	This data suppressed for respondent confidentiality.
375	A2ACCHNS	A2 Q7D STUDENTS SPEAK CHINESE-AM	This data suppressed for respondent confidentiality.
376	A2ACJPNS	A2 Q7E STUDENTS SPEAK JAPANESE-AM	This data suppressed for respondent confidentiality.
377	A2ACKRN	A2 Q7F STUDENTS SPEAK KOREAN-AM	This data suppressed for respondent confidentiality.
378	A2ACFLPN	A2 Q7G STUDENTS SPEAK FILIPINO LNG-AM	This data suppressed for respondent confidentiality.
379	A2AOTASN	A2 Q7H STUDENTS SPEAK OTHER ASIAN LNG-AM	This data suppressed for respondent confidentiality.
380	A2AOTLNG	A2 Q7I STUDENTS SPEAK ANOTHER LNG-AM	This data suppressed for respondent confidentiality.
381	A2ALNGOS	A2 Q7I LANGUAGE OF INSTRUCTION - OTHER-AM	This data suppressed for respondent confidentiality.
384	A2PGIFT	A2 Q2A # CLASSIFIED AS GFTED/TALENTED-PM	This data recoded for respondent confidentiality.
385	A2PPRTGF	A2 Q2B # TAKE PART IN GIFTED/TALENTED-PM	This data recoded for respondent confidentiality.
390	A2PDISAB	A2 Q3A NUMBER WITH DISABILITIES-PM	This data recoded for respondent confidentiality.
394	A2PRETAR	A2 Q4D MENTAL RETARDATION-PM	This data recoded for respondent confidentiality.
400	A2PMULTI	A2 Q4J MULTIPLE DISABILITIES-PM	This data suppressed for respondent confidentiality.
401	A2PAUTSM	A2 Q4K AUTISM-PM	This data suppressed for respondent confidentiality.
403	A2PDEAF	A2 Q4M DEAFNESS AND BLINDNESS-PM	This data suppressed for respondent confidentiality.

Table 7-8. Masked variables in the teacher catalog (continued)

FieldID	FieldName	FieldLabel	Comment		
412	A2PCVTNM	A2 Q7C STUDENTS SPEAK VIETNAMESE-PM	This data suppressed for respondent confidentiality.		
413	A2PCCHNS	A2 Q7D STUDENTS SPEAK CHINESE-PM	This data suppressed for respondent confidentiality.		
414	A2PCJPNS	A2 Q7E STUDENTS SPEAK JAPANESE-PM	This data suppressed for respondent confidentiality.		
415	A2PCKRN	A2 Q7F STUDENTS SPEAK KOREAN-PM	This data suppressed for respondent confidentiality.		
416	A2PCFLPN	A2 Q7G STUDENTS SPEAK FILIPINO LNG-PM	This data suppressed for respondent confidentiality.		
417	A2POTASN	A2 Q7H STUDENTS SPEAK OTHER ASIAN LNG-PM	This data suppressed for respondent confidentiality.		
418	A2POTLNG	A2 Q7I STUDENTS SPEAK ANOTHER LNG-PM	This data suppressed for respondent confidentiality.		
419	A2PLNGOS	A2 Q7I LANGUAGE OF INSTRUCTION - OTHER-PM	This data suppressed for respondent confidentiality.		
422	A2DGIFT	A2 Q2A # CLASSIFIED AS GFTED/TALENTED-AD	This data recoded for respondent confidentiality.		
423	A2DPRTGF	A2 Q2B # TAKE PART IN GIFTED/TALENTED-AD	This data recoded for respondent confidentiality.		
428	A2DDISAB	A2 Q3A NUMBER WITH DISABILITIES-AD	This data recoded for respondent confidentiality.		
432	A2DRETAR	A2 Q4D MENTAL RETARDATION-AD	This data recoded for respondent confidentiality.		
438	A2DMULTI	A2 Q4J MULTIPLE DISABILITIES-AD	This data suppressed for respondent confidentiality.		
439	A2DAUTSM	A2 Q4K AUTISM-AD	This data suppressed for respondent confidentiality.		
441	A2DDEAF	A2 Q4M DEAFNESS AND BLINDNESS-AD	This data suppressed for respondent confidentiality.		
450	A2DCVTNM	A2 Q7C STUDENTS SPEAK VIETNAMESE-AD	This data suppressed for respondent confidentiality.		
451	A2DCCHNS	A2 Q7D STUDENTS SPEAK CHINESE-AD	This data suppressed for respondent confidentiality.		
452	A2DCJPNS	A2 Q7E STUDENTS SPEAK JAPANESE-AD	This data suppressed for respondent confidentiality.		
453	A2DCKRN	A2 Q7F STUDENTS SPEAK KOREAN-AD	This data suppressed for respondent confidentiality.		
454	A2DCFLPN	A2 Q7G STUDENTS SPEAK FILIPINO LNG-AD	This data suppressed for respondent confidentiality.		
455	A2DOTASN	A2 Q7H STUDENTS SPEAK OTHER ASIAN LNG-AD	This data suppressed for respondent confidentiality.		
456	A2DOTLNG	A2 Q7I STUDENTS SPEAK ANOTHER LNG-AD	This data suppressed for respondent confidentiality.		
457	A2DLNGOS	A2 Q7I LANGUAGE OF INSTRUCTION - OTHER-AD	This data suppressed for respondent confidentiality.		
	New Variables Replacing Suppressed Variables			**Suppressed Variables**	
886	A1AKGTYP	A1 Q3 TYPE OF KINDERGARTEN PROG TCH-AM 1=Regular kindergarten only 2=Other kindergarten type	This data recoded for respondent confidentiality.	A1AREGK A1A2YRK1 A1A2YRK2 A1ATRNK A1APR1ST A1AUNGR A1AMULGR	A1 Q3A TCH REGULAR 1-YR KINDERGARTEN-AM A1 Q3B TEACHES 1ST YR OF 2-YR K-AM A1 Q3C TEACHES 2ND YR OF 2-YR K-AM A1 Q3D TCH TRANSITIONAL KINDERGARTEN-AM A1 Q3E TEACHES PRE-1ST GRADE AFTER K-AM A1 Q3F TEACHES UNGRADED CLASS-AM A1 Q3G TEACHES MULTIGRADE CLASS-AM

Table 7-8. Masked variables in the teacher catalog (continued)

	New Variables Replacing Suppressed Variables			**Suppressed Variables**	
887	A1PKGTYP	A1 Q3 TYPE OF KINDERGARTEN PROG TCH-PM	This data recoded for respondent confidentiality.	A1PREGK	A1 Q3A TCH REGULAR 1-YR KINDERGARTEN-PM
		1=Regular kindergarten only		A1P2YRK1	A1 Q3B TEACHES 1ST YR OF 2-YR K-PM
		2=Other kindergarten type		A1P2YRK2	A1 Q3C TEACHES 2ND YR OF 2-YR K-PM
				A1PTRNK	A1 Q3D TCH TRANSITIONAL KINDERGARTEN-PM
				A1PPR1ST	A1 Q3E TEACHES PRE-1ST GRADE AFTER K-PM
				A1PUNGR	A1 Q3F TEACHES UNGRADED CLASS-PM
				A1PMULGR	A1 Q3G TEACHES MULTIGRADE CLASS-PM
888	A1DKGTYP	A1 Q3 TYPE OF KINDERGARTEN PROG TCH-AD	This data recoded for respondent confidentiality.	A1DREGK	A1 Q3A TCH REGULAR 1-YR KINDERGARTEN-AD
		1=Regular kindergarten only		A1D2YRK1	A1 Q3B TEACHES 1ST YR OF 2-YR K-AD
		2=Other kindergarten type		A1D2YRK2	A1 Q3C TEACHES 2ND YR OF 2-YR K-AD
				A1DTRNK	A1 Q3D TCH TRANSITIONAL KINDERGARTEN-AD
				A1DPR1ST	A1 Q3E TEACHES PRE-1ST GRADE AFTER K-AD
				A1DUNGR	A1 Q3F TEACHES UNGRADED CLASS-AD
				A1DMULGR	A1 Q3G TEACHES MULTIGRADE CLASS-AD

Table 7-9. Masked variables in the school catalog

FieldID	FieldName	FieldLabel	Comment
3	CS_TYPE2	TYPE OF SCHOOL IN SAMPLE FRAME	This data recoded for respondent confidentiality.
4	KURBAN	LOCATION TYPE IN SAMPLE FRAME	This data recoded for respondent confidentiality.
9	S2KENRLK	S2 TOTAL SCHOOL K ENROLLMENT	This data recoded for respondent confidentiality.
10	S2KENRLS	S2 TOTAL SCHOOL ENROLLMENT	This data recoded for respondent confidentiality.
11	S2KMINOR	S2 PERCENT MINORITY STUDENTS	This data recoded for respondent confidentiality.
12	S2KFLNCH	S2 PCT FREE LUNCH ELIGIBLE STUDENTS	This data recoded for respondent confidentiality.
13	S2KRLNCH	S2 PCT REDUCED LUNCH ELIGIBLE STUDENTS	This data recoded for respondent confidentiality.
15	S2KLNGTH	S2 SCHOOL YEAR LENGTH	This data recoded for respondent confidentiality.
16	S2KGFTED	S2 PERCENT GIFTED/TALENTED STUDENTS	This data recoded for respondent confidentiality.
18	S2NUMDAY	S2 Q1 NUMBER OF DAYS MUST ATTEND	This data recoded for respondent confidentiality.
19	S2ADA	S2 Q2 % AVERAGE DAILY ATTENDANCE FOR YR.	This data recoded for respondent confidentiality.
20	S2ANUMCH	S2 Q3A # ENROLLED AROUND 10/1/1998	This data recoded for respondent confidentiality.
21	S2BNUMCH	S2 Q3B # ENROLLED SINCE 10/1/1998	This data recoded for respondent confidentiality.
22	S2CNUMCH	S2 Q3C # WHO LEFT - DIDN'T RETURN	This data recoded for respondent confidentiality.
45	S2BIASKL	S2 Q6D IS IT A TRIBAL SCHOOL	This data suppressed for respondent confidentiality.
46	S2SPEDSK	S2 Q6E IS IT A SPECIAL ED SCHOOL	This data suppressed for respondent confidentiality.
51	S2PRIVRD	S2 Q7A4 IS IT A PRIVATE ORDER	This data suppressed for respondent confidentiality.
53	S2NAISKL	S2 Q7C PRIVATE-ACCREDITED BY NAIS	This data suppressed for respondent confidentiality.
55	S2PVTSPD	S2 Q7E IS IT SPECIAL EDUCATION	This data suppressed for respondent confidentiality.
58	S2EMPHAS	S2 Q9 WHAT'S THE SCHOOL EMPHASIS	This data suppressed for respondent confidentiality.
59	S2PREK	S2 Q10 PRE-K PARTICIPATES	This data suppressed for respondent confidentiality.
60	S2REDDYN	S2 Q10 TRANSITIONAL PARTICIPATES	This data suppressed for respondent confidentiality.
61	S2KINGAR	S2 Q10 KINDERGARTEN PARTICIPATES	This data suppressed for respondent confidentiality.
62	S2PRE1ST	S2 Q10 PREFIRST PARTICIPATES	This data suppressed for respondent confidentiality.

Table 7-9. Masked variables in the school catalog

FieldID	FieldName	FieldLabel	Comment
63	S2FIR1ST	S2 Q10 1ST GRADE PARTICIPATES	This data suppressed for respondent confidentiality.
64	S2SEC2ND	S2 Q10 2ND GRADE PARTICIPATES	This data suppressed for respondent confidentiality.
65	S2THI3RD	S2 Q10 3RD GRADE PARTICIPATES	This data suppressed for respondent confidentiality.
66	S2FOR4TH	S2 Q10 4TH GRADE PARTICIPATES	This data suppressed for respondent confidentiality.
67	S2FIF5TH	S2 Q10 5TH GRADE PARTICIPATES	This data suppressed for respondent confidentiality.
68	S2SIX6TH	S2 Q10 6TH GRADE PARTICIPATES	This data suppressed for respondent confidentiality.
69	S2SEV7TH	S2 Q10 7TH GRADE PARTICIPATES	This data suppressed for respondent confidentiality.
70	S2ATE8TH	S2 Q10 8TH GRADE PARTICIPATES	This data suppressed for respondent confidentiality.
71	S2NIN9TH	S2 Q10 9TH GRADE PARTICIPATES	This data suppressed for respondent confidentiality.
72	S2TEN10	S2 Q10 10TH GRADE PARTICIPATES	This data suppressed for respondent confidentiality.
73	S2ELEV11	S2 Q10 11TH GRADE PARTICIPATES	This data suppressed for respondent confidentiality.
74	S2TWEL12	S2 Q10 12TH GRADE PARTICIPATES	This data suppressed for respondent confidentiality.
75	S2SPENR	S2 Q11 NUMBER IN SPECIAL PROGRAM	This data suppressed for respondent confidentiality.
92	S2PCTHSP	S2 Q14 PERCENT OF HISPANIC STUDENTS	This data recoded for respondent confidentiality.
93	S2INDPCT	S2 Q15A PERCENT OF NATIVE AMERICANS	This data suppressed for respondent confidentiality.
94	S2ASNPCT	S2 Q15B PERCENT OF ASIAN STUDENTS	This data suppressed for respondent confidentiality.
95	S2BLKPCT	S2 Q15C PERCENT OF BLACK STUDENTS	This data recoded for respondent confidentiality.
96	S2PCFPCT	S2 Q15D PERCENT OF HAWAIIAN STUDENT	This data suppressed for respondent confidentiality.
97	S2WHTPCT	S2 Q15E PERCENT OF WHITE STUDENTS	This data suppressed for respondent confidentiality.
98	S2ELIBRK	S2 Q16A1 ELIGIBLE FOR FREE BREAKFAST	This data suppressed for respondent confidentiality.
99	S2PARBRK	S2 Q16A2 PARTICIPATES IN BREAKFAST	This data suppressed for respondent confidentiality.
100	S2ELILNC	S2 Q16B1 ELIGIBLE FOR FREE LUNCH	This data suppressed for respondent confidentiality.
101	S2PARLNC	S2 Q16B2 PARTICIPATES IN FREE LUNCH	This data suppressed for respondent confidentiality.
102	S2ELIRED	S2 Q16C1 ELIGIBLE IN REDUCED-PRICE LUNCH	This data suppressed for respondent confidentiality.
103	S2PARRED	S2 Q16C2 PARTICIPATES IN RED-PRICE LUNCH	This data suppressed for respondent confidentiality.
113	S2CHSRVD	S2 Q20 NUMBER SERVED BY TITLE 1	This data recoded for respondent confidentiality.
114	S2HDCHDN	S2 Q21A1 HALFDAY - NUMBER OF CHILDREN	This data recoded for respondent confidentiality.
115	S2HDCLS	S2 Q21B1 HALFDAY - NUMBER OF CLASSES	This data recoded for respondent confidentiality.
116	S2HDDAYS	S2 Q21C1 HALFDAY - NUMBER OF DAYS PER WK	This data suppressed for respondent confidentiality.
118	S2FDCHDN	S2 Q21A2 FULLDAY - NUMBER OF CHILDREN	This data recoded for respondent confidentiality.
119	S2FDCLS	S2 Q21B2 FULLDAY - NUMBER OF CLASSES	This data recoded for respondent confidentiality.
120	S2FDDAYS	S2 Q21C2 FULLDAY - NUMBER OF DAYS PER WK	This data suppressed for respondent confidentiality.

Table 7-9. Masked variables in the school catalog (continued)

FieldID	FieldName	FieldLabel	Comment
122	S2KCCHDN	S2 Q21A3 K COMBINATION - # OF CHILDREN	This data recoded for respondent confidentiality.
123	S2KCCLS	S2 Q21B3 K COMBINATION - # OF CLASSES	This data suppressed for respondent confidentiality.
124	S2KCDAYS	S2 Q21C3 K COMBINATION - # DAYS PER WK	This data suppressed for respondent confidentiality.
126	S2TKCHDN	S2 Q21A4 TRANSITIONAL K - # OF CHILDREN	This data recoded for respondent confidentiality.
127	S2TKCLS	S2 Q21B4 TRANSITIONAL K - # OF CLASSES	This data suppressed for respondent confidentiality.
128	S2TKDAYS	S2 Q21C4 TRANSITIONAL K - # DAYS PER WK	This data suppressed for respondent confidentiality.
130	S2T1CHDN	S2 Q21A5 TRANSITIONAL 1ST- # OF CHILDREN	This data recoded for respondent confidentiality.
131	S2T1CLS	S2 Q21B5 TRANSITIONAL 1ST- # OF CLASSES	This data suppressed for respondent confidentiality.
132	S2T1DAYS	S2 Q21C5 TRANSITIONAL 1ST - DAYS PER WK	This data suppressed for respondent confidentiality.
137	S2NOCUTO	S2 Q22A NO CUTOFF DATE TO TURN FIVE	This data suppressed for respondent confidentiality.
149	S2CHLDNM	S2 Q24 # OF CHILDREN SITE ACCOMMODATES	This data recoded for respondent confidentiality.
150	S2RMNUM	S2 Q25 NUMBER OF INSTRUCTIONAL ROOMS	This data recoded for respondent confidentiality.
161	S2INSTCM	S2 Q27A # INSTRUCTIONAL COMPUTERS	This data recoded for respondent confidentiality.
162	S2ADMNCM	S2 Q27B # INSTRUCT/ADMIN COMPUTERS	This data recoded for respondent confidentiality.
163	S2TOTCM	S2 Q27C TOTAL NUMBER OF COMPUTERS	This data recoded for respondent confidentiality.
166	S2LANRMS	S2 Q28A3 TOTAL ROOMS W/LAN ACCESS	This data recoded for respondent confidentiality.
169	S2CDRRMS	S2 Q28B3 TOTAL ROOMS W/CD-ROM ACCESS	This data recoded for respondent confidentiality.
172	S2NETRMS	S2 Q28C3 TOTAL RMS W/INTERNET ACCESS	This data recoded for respondent confidentiality.
176	S2CMNITY	S2 Q30 COMMUNITY SCHOOL IS LOCATED IN	This data recoded for respondent confidentiality.
232	S2NMBKPT	S2 Q40 NUMBER RETAINED GRADE-K	This data recoded for respondent confidentiality.
274	S2NMADMN	S2 Q47A2 NUMBER OF ADMINISTRATORS	This data recoded for respondent confidentiality.
276	S2NMTECH	S2 Q47B2 NUMBER OF TEACHERS	This data recoded for respondent confidentiality.
278	S2NMDSTR	S2 Q47C2 NUMBER FROM DISTRICT OFFICE	This data recoded for respondent confidentiality.
280	S2NUMBRD	S2 Q47D2 NUMBER OF BOARD MEMBERS	This data recoded for respondent confidentiality.
282	S2NMOMDD	S2 Q47E2 NUMBER OF PARENTS	This data recoded for respondent confidentiality.
284	S2NMREPS	S2 Q47F2 NUMBER OF REPRESENTATIVES	This data recoded for respondent confidentiality.
286	S2NUMOTH	S2 Q47G2 NUMBER OF OTHER GROUPS	This data recoded for respondent confidentiality.
290	S2BILING	S2 Q50 BILINGUAL SERVICES PERCENT	This data suppressed for respondent confidentiality.
291	S2ESLONL	S2 Q50 ESL SERVICES PERCENT	This data suppressed for respondent confidentiality.
292	S2ESLBIL	S2 Q50 ESL AND BILINGUAL PERCENT	This data suppressed for respondent confidentiality.
293	S2YRSESL	S2 Q51A YEARS FOR ESL SERVICES	This data recoded for respondent confidentiality.
294	S2YRSBIL	S2 Q51B YEARS FOR BILINGUAL SERVICES	This data recoded for respondent confidentiality.

Table 7-9. Masked variables in the school catalog (continued)

FieldID	FieldName	FieldLabel	Comment
295	S2SRVBTH	S2 Q51C YEARS FOR ESL & BILINGUAL SER	This data recoded for respondent confidentiality.
319	S2GFTNBR	S2 Q57 NUMBER GIFTED/TALENTED	This data suppressed for respondent confidentiality.
320	S2ADMFTE	S2 Q58A FTE ADMINISTRATIVE STAFF	This data recoded for respondent confidentiality.
321	S2TCHFTE	S2 Q58B FTE CLASSROOM TEACHERS	This data recoded for respondent confidentiality.
322	S2GYMFTE	S2 Q58C FTE GYM/DRAMA/ART TEACHERS	This data recoded for respondent confidentiality.
323	S2CMPFTE	S2 Q58D FTE COMPUTER SPECIALISTS	This data recoded for respondent confidentiality.
324	S2HDSFTE	S2 Q58E FTE HEAD START/PRE-K TEACHERS	This data recoded for respondent confidentiality.
325	S2CRDFTE	S2 Q58F FTE COORDINATOR/SPECIALIST	This data recoded for respondent confidentiality.
326	S2TT1FTE	S2 Q58G FTE TITLE 1 TEACHERS	This data recoded for respondent confidentiality.
327	S2AIDFTE	S2 Q58H FTE TEACHER AIDES	This data recoded for respondent confidentiality.
328	S2SPEFTE	S2 Q58I FTE SPECIAL ED TEACHERS	This data recoded for respondent confidentiality.
329	S2SPAFTE	S2 Q58J FTE SPECIAL ED AIDES	This data recoded for respondent confidentiality.
330	S2ESLFTE	S2 Q58K FTE ESL-BILINGUAL TEACHERS	This data recoded for respondent confidentiality.
331	S2ESAFTE	S2 Q58L FTE ESL-BILINGUAL AIDES	This data recoded for respondent confidentiality.
332	S2LIAFTE	S2 Q58M FTE PARENT LIAISONS	This data recoded for respondent confidentiality.
333	S2COUFTE	S2 Q58N FTE SOCIAL WORKERS	This data recoded for respondent confidentiality.
334	S2LIBFTE	S2 Q58O LIBRARY/MEDIA SPECIALISTS	This data recoded for respondent confidentiality.
335	S2THRFTE	S2 Q58P SPEECH/PHYSICAL THERAPISTS	This data recoded for respondent confidentiality.
336	S2NRSFTE	S2 Q58Q FTE SCHOOL NURSES	This data recoded for respondent confidentiality.
337	S2NONFTE	S2 Q58R FTE NON-INSTRUCTIONAL STAFF	This data recoded for respondent confidentiality.
338	S2REAFTE	S2 Q58S FTE READING RECOVERY TCHRS	This data suppressed for respondent confidentiality.
339	S2MSCFTE	S2 Q58T FTE MATH/SCIENCE SPECIALISTS	This data recoded for respondent confidentiality.
340	S2FRLFTE	S2 Q58U FTE FOREIGN LNG SPECIALISTS	This data recoded for respondent confidentiality.
341	S2OTHFTE	S2 Q58V OTHER FTE	This data suppressed for respondent confidentiality.
342	S2FTETOT	S2 Q58W TOTAL FTE SCHOOL STAFF	This data recoded for respondent confidentiality.
345	S2ETHNIC	S2 Q61 % HISPANIC/LATINO TEACHERS	This data recoded for respondent confidentiality.
346	S2Q62IND	S2 Q62 % AMERICAN NATIVE TEACHERS	This data recoded for respondent confidentiality.
347	S2Q62ASN	S2 Q62 % ASIAN TEACHERS	This data recoded for respondent confidentiality.
348	S2Q62BLK	S2 Q62 % BLACK TEACHERS	This data recoded for respondent confidentiality.
349	S2Q62HAW	S2 Q62 % HAWAIIAN TEACHERS	This data suppressed for respondent confidentiality.
446	S2YSTCH	S2 Q74A NUMBER OF YRS TEACHING	This data recoded for respondent confidentiality.
447	S2TOTPRI	S2 Q74B NUMBER OF YRS AS PRINCIPAL	This data recoded for respondent confidentiality.
448	S2PRINHR	S2 Q74C NUMBER YRS A PRINCIPAL HERE	This data recoded for respondent confidentiality.
468	S2EDLVL	S2 Q77 HIGHEST LEVEL OF EDUCATION	This data recoded for respondent confidentiality.

Table 7-9. Masked variables in the school catalog (continued)

	New Variables Replacing Suppressed Variables			Suppressed Variables	
703	S2PRACE	S2 Q73 PRINCIPAL'S RACE	This data recoded for respondent confidentiality.	S2RACE1	S2 Q73A PRINCIPAL IS AMERICAN INDIAN
		1=White, non-Hispanic		S2RACE2	S2 Q73B PRINCIPAL IS ASIAN
		2=Black or African-American, non-Hispanic		S2RACE3	S2 Q73C PRINCIPAL IS BLACK
		3=Other race		S2RACE4	S2 Q73D PRINCIPAL IS HAWAIIAN OR PAC IS
		-9=Not ascertained		S2RACE5	S2 Q73E PRINCIPAL IS WHITE

NATIONAL CENTER FOR EDUCATION STATISTICS

Early Childhood Longitudinal Study
Kindergarten Class of 1998-99

USER'S MANUAL FOR THE ECLS-K FIRST GRADE PUBLIC-USE DATA FILES AND ELECTRONIC CODE BOOK
NCES 2002-135

Prepared by:

Westat
Rockville, Maryland
Karen Tourangeau
John Burke
Thanh Le
Siu Wan
Margaret Weant
Christine Nord
Nancy Vaden-Kiernan
Elizabeth Bissett
Richard Dulaney
Alison Fields
Lauren Byrne
Ismael Flores-Cervantes
Jean Fowler

Educational Testing Service
Princeton, New Jersey
Judith Pollack
Donald Rock

University of Michigan
School of Education
Ann Arbor, Michigan
Sally Atkins-Burnett
Samuel Meisels

The National Center for Education Statistics
U.S. Department of Education
Washington, DC
Jonaki Bose
Jerry West

Education Statistics Services Institute
Washington, DC
Kristin Denton
Amy Rathbun
Jill Walston

U.S. Department of Education
Office of Educational Research and Improvement   NCES 2002-135

U.S. Department of Education
Rod Paige
Secretary

Office of Educational Research and Improvement
Grover J. Whitehurst
Assistant Secretary

National Center for Education Statistics
Gary W. Phillips
Acting Commissioner

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

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

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
Office of Educational Research and Improvement
U.S. Department of Education
555 New Jersey Avenue, N.W.
Washington, DC 20208-5574

January 2002

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

TABLE OF CONTENTS

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)

TABLE OF CONTENTS (continued)

List of Exhibits

Exhibit

BASE YEAR ERRATA AND COMPOSITES

Since the release of the base year data files and documentation, several errata have been identified. These errata are listed here for the base year public-use data files (NCES 2001-029e, June 2001). Errors in the data file are listed first, followed by errors in the documentation. Corrected versions of these base year variables are available in appendix D for all children who participated in the base year. Users who wish to use the corrected base year variables should refer to appendix D. In addition, appendix D contains several base year composites that were created after the base year data files were released. These composites are described in chapter 7, section 7.4.6.

A. The following errors were identified in the ECLS-K Base Year Public-Use *Child* data files (child.dat)

1. WKPOVRTY

There are 383 cases that were coded as at or above the poverty level (WKPOVRTY=2) that should have been coded as below the poverty level (WKPOVRTY=1). There are two additional cases (that should have coded above the poverty level, but instead were coded below poverty level).

There were 1179 cases that had a round 1 interview, but no round 2 interview. For these cases the poverty composite was created by imputing for income and using round 1 household size. All 1179 of these cases were coded as at or above the poverty level, when 383 of the 1179 cases should have been coded as below the poverty level. Thus, in addition to the 3855 cases already coded as below the poverty level, 383 more cases should have been included in this category.

Two cases (0060013C and 1114006C) were coded as below poverty level instead of above due to an error in the income cut-off point in the program.

For a list of the 385 cases with an incorrect poverty status by child ID please refer to the file POVERTY_IDS.TXTon the NCES website at *http://nces.ed.gov/ecls/kindergarten/errata.htm*.

NCES has re-released the corrected base year poverty composite with the first grade data; the corrected composite is called WKPOV_R on the first grade data file.

2. KURBAN

In the public-use file, a seven-category locale code was collapsed to a three-category locale code (KURBAN). The labels of the three-category locale variable correctly read as follows:

1. Central City (Large City and Mid-Size City)

2. Urban Fringe and Large Town (Urban Fringe should include Large City Urban Fringe and Mid-Size City Urban Fringe.)

3. Small Town and Rural

However, the categories were erroneously collapsed as:

1. Large Central City and its urban fringe

2. Mid-Size Central City and its urban fringe and Large Town

3. Small Town and Rural

NCES has re-released the corrected base year school locale code with the first grade data.

3. C2ASMTST

12 children were incorrectly coded as nonrespondents instead of not assessed in the variable C2ASMTST (C2 CHILD ASSESSMENT STATUS). Instead of 4=CHILD W/ DISABILITY, NOT ASSESSED, they were coded 5=NONRESPONDENT. This error only affects this variable. These 12 children were all correctly assigned spring kindergarten (round 2) weights and they were included appropriately in table 5-12 in chapter 5 of the base year user's manual.

The child IDs (CHILDID) for these children are as follows:

0105005C
0105008C
0105009C
0105012C
0105013C
0105018C
0105019C
0105020C
0105021C
0105023C
0105024C
2121014C

4. Academic Rating Scale (ARS) Scores

An error was identified in the base year ARS scores. Specifically, the fall and spring base year ARS scores use slightly different metrics and, therefore, are not directly comparable. The specifics of the problem are described below.

The ARS scale scores are interval level scores and may be used in analyses requiring that level of measurement. However, *the use of* gain scores (subtracting the fall score from the spring score) *is* not recommended. In order to represent the student's score within a 1-5 range, an arithmetic transformation using the mean and standard deviation was applied to the scores. Fall and spring scores were analyzed separately; therefore, the metric is slightly different due to differences in the range of the scores at the two time points. In other words, an increase of 0.1 on the fall scale is not the same as an increase of 0.1 on the spring scale. Further, because some children performed above grade level, their ratings on the ARS may be at the maximum score. Consequently, it is not possible to estimate how far beyond the skills, knowledge, and behaviors assessed on the ARS these students might have achieved. In short, any estimate of fall to spring growth will be underestimated for the highest achieving students. Although gain scores are not recommended, covariance models may be used (with the caveat that there are some ceiling and floor effects).

In the first grade data files, corrected scores for the kindergartners are included. These scores were calibrated using a combined calibration of fall and spring kindergarten ratings. Therefore, the unit for the fall and spring kindergarten scores are the same. The problem with estimating growth for students at the ceiling remain. The standard error of measurement (SEM) for the scores is provided in the User's Manual. The first grade scores are based on different items and should not be used to compare growth.

5. IF_INC

Income was collected in the round 2 spring kindergarten data collection. For those households that did not provide this information, an income value was imputed. Imputed income values are indicated on a file using a variable IF_INC. There is an error in the variable IF_INC on the file. There are 3379 cases on the file coded as "0" (not imputed) that should have been coded as "1" (imputed).

The file IMPUTE_IDS.TXT contains the child IDs of the cases with the erroneous flag on the NCES website at http://nces.ed.gov/ecls/kindergarten/errata.htm.

6. Five children in the spring-kindergarten had their parent interviews conducted under the wrong parent identification number. As a result the child records for these children have incorrect parent data, although they are correct for the other components (child data, non-parent weights, etc.). None of the five children have fall-kindergarten parent data. Two of these children should not have had any parent data since they were round 2 nonrespondents, while three of the cases had their parent data stored under other children's identification numbers. The IDs are:

Case (Child ID)		Has parent data belonging too child with ID…	Should have parent data that is currently in child ID…
1.	0162001C	3056007C	No parent data (R2 nonrespondent)
2.	0192001C	0192003C	No parent data (R2 nonrespondent)
3.	0192003C	0192015C	0192001C
4.	0192015C	None	0192003C
5.	3056007C	None	0162001C

There are two recommended strategies for correcting this error. Users may go ahead and use the cases since the effect on any analysis will be trivial. The other option would be to discard these five cases during analysis. There would not be much benefit in reassigning the correct data to the appropriate child, as parent-level weights were not created for cases 0192015C and 3056007C for round 2 as they were considered as having missing parent data during the process of the creation of round 2 and base year longitudinal parent weights.

7. The format (value) labels for four variables P2AGREE1 (P2 NRQ264 AGREEMENT W/ BIOLOGICAL FATHER), P2AGREE2 (P2 NRQ264 AGREEMENT W/ BIOLOGICAL MOTHER), P2AGREE3 (P2 NRQ264 AGREEMENT W/ ADOPTIVE FATHER), and P4AGREE4 (P2 NRQ264 AGREEMENT W/ ADOPTIVE MOTHER) are incorrect.

   Currently, each of the variables has value labels ranging in valued from 1 to 4 with labels (1=biological father, 2=biological mother, 3=adoptive father, 4=adoptive mother). The variables should all have had the labels 1=yes and 2=no. All cases have a value of either 1 or 2 for each of these four variables; there are no cases with values of 3 or 4.

   For example, for the variable P2AGREE1 (P2 NRQ264 AGREEMENT W/ BIOLOGICAL FATHER) the frequency distribution is as follows:

Code in restricted-use file	Frequency (# of cases)	Correct Code
1 (biological father)	95	1 (yes)
2 (biological mother)	36	2 (no)
3 (adoptive father)	0	---
4 (adoptive mother)	0	---

The same situation applies to all the P2 NRQ264 variables.

8. The variable labels for the question series P2 PRRDP_* (P2 ROSTER ROUND DEPARTED - PERSON *) (where * ranges from 1 to 17) are in error. The labels currently read

   1 Joined Round 1

   2 Joined Round 2

   The labels should read:

   1 Departed Round 1

   2 Departed Round 2

1. INTRODUCTION

This manual provides guidance and documentation for users of the first 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 and use 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 and teachers, 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 the Survey Research Center and the School of Education at the University of Michigan and Educational Testing Service (ETS) in Princeton, New Jersey.

The ECLS-K is following a nationally representative cohort of children from kindergarten through fifth grade. The base year data were collected in the fall and spring of 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 table 1-1). It was a design enhancement whose goal was 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 moved

[1] The term "first grade" is used throughout this document to refer to the data collections that took place in the 1999-2000 school year, at which time most of the sampled children—but not all of them—were in first 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.

ahead to second grade. In addition, children who were not in kindergarten in the United States during the 1998-99 school year and, therefore, did not have a chance to be selected to participate in the base year of the ECLS-K were added to the spring-first grade sample.[4] Such children include immigrants, 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.

Table 1-1.—ECLS-K waves of data collection[1]

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[2]
Spring-first grade	Spring 2000	Full sample
Spring-third grade	Spring 2002	Full sample
Spring-fifth grade	Spring 2004	Full sample

[1] See section 1.3 for a description of the study components.

[2] 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.

The final two waves of data collection that are currently planned are scheduled for spring 2002 and spring 2004 when most of the study children will be in the third grade and fifth grade, respectively.

The ECLS-K has several major objectives and numerous potential applications. The ECLS-K combines elements of (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.

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

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) will follow a national sample of children, born in the year 2001, from birth through first grade. The ECLS-B will focus on the characteristics of children and their families that influence children's first experiences with the demands of formal school, as well as children's early health care and in- and out-of-home experiences. Together these cohorts will provide the range and breadth of data required to more fully describe and understand children's health and early learning, development, and education experiences.

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

1.1 Background

National policymakers and the public at large have increasingly recognized that the prosperity of the United States depends on the successful functioning of the American education system. There is also growing awareness that school reform efforts cannot focus solely on the secondary and postsecondary years but must pay attention to the elementary and preschool years as well. Increased policy interest in the early grades and the early childhood period is reflected in an intensified recent national policy aimed at ensuring that children are capable of reading by the third grade, providing college student and adult volunteer tutors for children who are having difficulty learning to read, and increasing the number of children from low-income families served by Head Start to one million by the year 2002.

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 affected by 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. This produces a more accurate measurement of these antecedent factors and enables stronger causal inferences to be made about their relationship to later academic progress.

The ECLS-K enables educational policy analysts to use an ecological perspective on early childhood education, using techniques such as multilevel modeling to study how school and classroom factors affect the progress of individual children. The data collected will enable analysts to examine how children's status at school entry and performance in school are jointly determined by an interaction of child characteristics and school and family environments.

Data collected during the kindergarten year serve as baseline measures to examine how schooling shapes later individual development. 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.

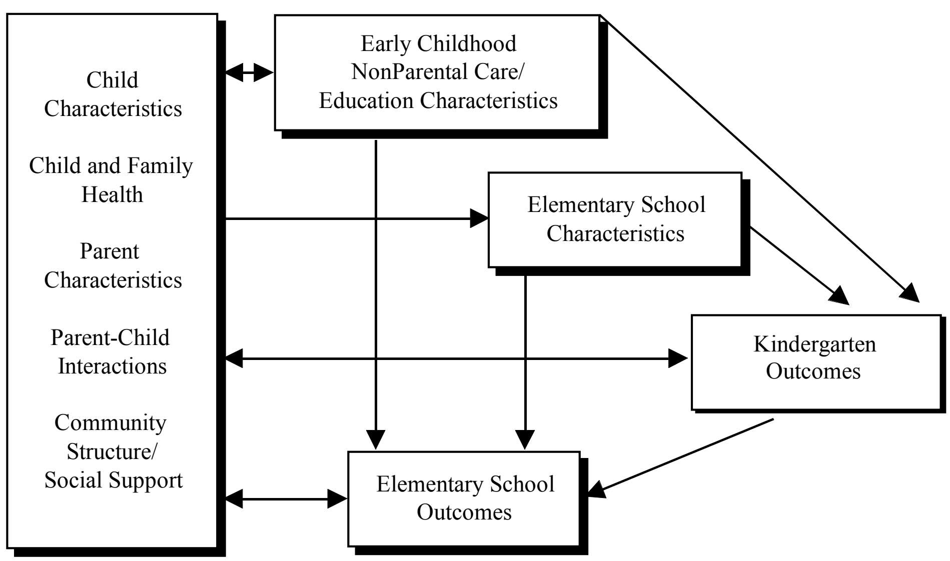
The first grade data collection has two distinct purposes. As noted previously, the fall-first grade data collection can be used to study the extent of children's summer learning loss and the factors contributing to differential loss, and to better disentangle school and home influences on children's learning. To this end, the fall-first grade data collection gathered detailed information about how the children spent their summers and about various types of cognitive stimulation that they received over the summer either from their parents or from other sources such as camps, summer school, tutoring, or enrichment programs. In addition, detailed information was gathered about the children's child care arrangements during the summer. Thus, the fall-first grade data collection can also be used to study how kindergartners' child care arrangements change during the summer.

The spring-first grade data collection can be used to describe the diversity of first graders and the classrooms and schools they attend. It can also be used to study children's academic gains in the year 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 and approaches 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 figure 1-1. 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 primary 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.

Figure 1-1.—ECLS-K conceptual model



1.3 Study Components

The emphasis that is being placed on the whole of the child's environments and development 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** are asked to 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 (e.g., general knowledge, literacy, and quantitative skills) and noncognitive (e.g., fine motor and gross motor coordination, 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 will report on their own experiences in and out of school. 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. They are asked to provide key information about their children, especially during the first years of the study. Parents are one of the important sources of 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 but 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. Teachers complete self-administered questionnaires each time children are assessed, with the exception of the fall-first grade data collection.

- **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. Information is collected from school administrators via self-administered questionnaires during 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 base-year and first-grade data files that are either currently available or will be available shortly.

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. This is in compliance with the Privacy Act of 1974 and the National Education Statistics Act of 1994, both of which mandate the protection of confidentiality of respondents. The process begins with a formal disclosure risk analysis. Variables identified as posing the greatest disclosure risk are altered, and in some instances, entirely suppressed, and in this way the public-use data files are created. Every effort is made to alter the files as little as possible, consistent with the requirement for confidentiality protection. After altering the variables, the disclosure risk analysis is repeated to verify that the disclosure risk has been reduced to acceptable levels.

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

- Outlier values are top- or bottom-coded;
- 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) to reduce the risk;
- Some continuous variables are modified into categorical variables, and certain categorical variables have their categories collapsed; and
- Certain variables with too few cases and a sparse distribution are suppressed altogether, rather than modified.

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 the disabled, or children with specific non-English home languages or countries of birth, for example, will find that the restricted-use files contain a few more variables. 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

Several different ECLS-K base-year and first-grade data files are available or will shortly be available for use by analysts:

- **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 User's Manual (NCES 2001-029), February 2001 or the ECLS-K Restricted-Use Base Year User's Manual (NCES 2000-097), August 2000.

  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 User's Manual for the ECLS-K Base Year Restricted-Use Salary and Benefits Data Files and Electronic Code Book (NCES 2001-014).

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 User's Manual for the ECLS-K Base Year Restricted-Use Special Education Data Files and Electronic Code Book (NCES 2001-015)

The **student record abstract file** contains information from school records about children's school enrollment and attendance; Individualized Education Plan (IEP) and disability status; and home and school language. The student record 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 User's Manual for the ECLS-K Base Year Restricted-Use Student Record Abstract Data Files and Electronic Code Book (NCES 2001-0016).

The purpose of the **Head Start Verification Study** was two-fold: (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. The Head Start Verification file contains information collected during the verification process. This file is a restricted-use file. For more information about this file, see the User's Manual for the ECLS-K Base Year Restricted-Use Head Start Data Files and Electronic Code Book (NCES 2001-025).

- **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. First grade teacher and school files are not being 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 not only can be used to analyze data collected in the first grade but also provides 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. Instructions on how to create a longitudinal file using the base year and first grade data are provided in chapter 9. 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 here.

- **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 who wish 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.

1.5 Contents of Manual

This manual provides documentation for users of the first grade public-use data of the ECLS-K. 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 how the public-use first grade data file is structured, provides definitions of composite variables (chapter 7), explains how to use the Electronic Code Book (chapter 8) and describes how to use and merge the base year and first grade files (chapter 9). The Electronic Code Book contains unweighted frequencies for all variables. Because this manual focuses on the first grade data collection, minimal information is provided about the base year data. Users who wish to learn more about the base year data collection should refer to the ECLS-K Base Year Public-Use User's Manual (NCES 2001-029), February 2001 or the ECLS-K Restricted-Use Base Year User's Manual (NCES 2000-097), August 2000. Additional information about the ECLS program can be found on the World Wide Web at *http://nces.ed.gov/ecls.*

2. DESCRIPTION OF DATA COLLECTION INSTRUMENTS

This chapter describes the survey instruments used during the first grade data collection of the Early Childhood Longitudinal Study-Kindergarten Class of 1998-99 (ECLS-K). Table 2-1 lists all the instruments used during the first grade data collection. The instrumentation for the base year is also shown. Similarities and differences between the first grade and base year instruments are highlighted throughout this chapter.

The ECLS-K first grade data collection occurred in the fall and spring of the 1999-2000 school year. In the fall of 1999, data were collected by computer-assisted interviews (CAIs) with parents and by direct child assessments. Unlike the base year and spring-first grade data collections, no teacher or other school questionnaires were administered in the fall. As noted in chapter 1, the fall-first grade data collection was a design enhancement intended to examine changes in children's learning during the summer months. The fall instrumentation reflects this special focus.[1] In the spring of 2000, data were collected by computer-assisted parent interviews, teacher questionnaires, and direct child assessments. In addition, data were gathered from school administrators, student record abstracts, and through a facilities checklist. Because about five percent of the sampled children were retained in kindergarten, separate teacher questionnaires were developed for kindergarten and first grade teachers. Teachers of sampled students promoted to second grade received the first grade teacher questionnaires.

The Head Start verification study and the salary and benefits questionnaire administered in the spring of the base year were not repeated during the first grade data collection.[2] A copy of the first grade data collection instruments, except for the direct child assessment, the social rating scale (SRS)[3] in the parent interview and teacher questionnaire, and the adaptive behavior scale (ABS) are available on the CD-ROM as appendix A.[4] These latter measures contain copyright protected materials and agreements with the test publishers that restrict their distribution.

---

[1] See section 2.1 for more information about the purposes of the fall first-grade data collection.

[2] Users wanting information about the Head Start verification study should refer to the ECLS-K Base Year Restricted-Use Head Start Data Files and Electronic Code Book (NCES 2001-025), which includes a user's manual with detailed information about the verification process. Those wanting information about the salary and benefits data files should refer to the User's Manual for the ECLS-K Base Year Restricted-Use Salary and Benefits Data Files and Electronic Code Book (NCES 2001-014).

[3] Adapted with permission from *Elementary Scale A ("How Often?")*, F.M. Gresham and S.N. Elliott. (1990). American Guidance Service, Inc.

[4] Lambert, N., Nihira, K., and Leland, H. *Adaptive Behavior Scale-Second Edition*. (1993). The American Association on Mental Retardation.

Table 2-1.—Instruments used in the ECLS-K base year and fall- and spring-first grade

1998-99 school year		1999-2000 school year	
Fall-kindergarten	Spring-kindergarten	Fall-first grade[1]	Spring-first grade
Parent Interview	Parent Interview	Parent Interview	Parent Interview
Child Assessment	Child Assessment	Child Assessment	Child Assessment
Teacher Questionnaire-Part A	Teacher Questionnaire-Part A		First Grade Teacher Questionnaire - Part A
Teacher Questionnaire-Part B	Teacher Questionnaire-Part B		First Grade Teacher Questionnaire – Part B
Teacher Questionnaire-Part C	Teacher Questionnaire-Part C		First Grade Teacher Questionnaire – Part C
			Kindergarten Teacher Questionnaire - Part A
			Kindergarten Teacher Questionnaire - Part B
			Kindergarten Teacher Questionnaire - Part C
	Special Education Teacher Questionnaire-Part A		Special Education Teacher Questionnaire-Part A
	Special Education Teacher Questionnaire-Part B		Special Education Teacher Questionnaire-Part B
	Adaptive Behavior Scale		Adaptive Behavior Scale
	School Administrator Questionnaire		School Administrator Questionnaire - New Schools
			School Administrator Questionnaire - Returning Schools
	Student Record Abstract		Student Record Abstract
	School Facilities Checklist		School Facilities Checklist
	Salary and Benefits[2] Questionnaire		
	Head Start Verification[3]		

[1] The fall-first grade data collection consisted of a 30 percent subsample of the study schools. See section 2.1 for information about the purposes of the fall-first grade data collection.

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

[3] The Head Start Verification Study verified 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.

2.1 Purposes of the Fall-First Grade Enhancement

The fall-first grade data collection was limited to a 30 percent subsample of schools. Approximately 27 percent of the base year children who were eligible to be interviewed in Year 2 attended these schools. Nearly 5,300 study children in these schools were assessed and their parents interviewed. The fall-first grade data collection was designed to study an important aspect of children's learning experience—summer learning. The existing research on summer learning shows conflicting findings about whether children experience decline or growth in learning over the summer as compared with learning during the school year (Entwisle and Alexander, 1992; Heyns, 1978). There are also conflicting findings regarding how children's learning of various subject areas varies differentially over the summer months.

Factors that have been related to learning rates over the summer include school poverty level, family socioeconomic status, year-round schooling, the child's attendance at summer school, the child's participation in structured summer programs, and reading materials and activities in the home (Entwisle and Alexander, 1992; Entwisle and Alexander, 1994; Cooper et al., 1996; Heyns, 1978). Although there are many common findings across studies (e.g., that learning loss over the summer is often greater for disadvantaged than for advantaged students), more research is needed to examine the effects of multiple factors on children's summer learning.

There are many other factors that have rarely been included or not included at all in the summer learning research. For example, although comparisons between black and white student achievement gains, controlling for other background factors, are frequently made in the literature, such comparisons do not appear to extend to students from other racial and ethnic groups, such as Asian Americans and Hispanic students. Limited English proficiency (LEP) students have also not been included in the extant literature. In addition, few studies have included children from rural or suburban areas.

A number of variables have been identified as possible contributors to children's rate of learning during the summer months. Some of the variables that have not been studied extensively in relation to summer learning include: family structure and composition; the use of tutoring and educational support from parents; the start and end dates of summer school programs, with the possibility that summer school programs ending closer to the start of the following school year may lead to higher achievement gains; and the use of home computers during the summer.

The ECLS-K fall-first grade data collection includes not only factors that have been studied in previous research but also those that have not yet been studied. It provides an opportunity to examine children's summer learning on a national sample of children with a wide range of characteristics represented.

2.2 Direct Child Assessments

One-on-one direct child assessments were administered using CAI in the fall and spring of the 1999-2000 school year. The children were assessed regardless of whether they were retained in kindergarten, promoted to first grade, moved ahead to second grade, or were new to the sample (freshened students). The assessments took about one hour to administer. Table 2-2 displays the major domains measured during the direct child assessments from all four rounds of data collection. As in the base year, the first grade assessments consisted of cognitive and physical components. The fall-first grade assessments, conducted with a subsample of study children, can be compared to the spring-kindergarten assessments obtained during the base year to determine children's growth over the summer months and can also be linked to the spring-first grade assessments to study the children's gain during the first grade academic year. The spring-first grade assessments can be linked to the base year assessments conducted in the fall of 1998 and the spring of 1999.

Language Screener. As in the base year, the assessment began by verifying the child's name and administering a set of warm-up exercises similar in form to the items used to administer the cognitive component. Prior to administering the cognitive assessment battery, a language-screening assessment (the OLDS), was administered to those children identified from their school records (or by their teacher, if no school records were available) as coming from a language minority background (meaning that their primary home language was not English) and who had not passed the OLDS in any of the previous rounds of interviews.[5] This screening test was used to determine if a child was able to understand and respond to the cognitive assessment items in English. If a child did not pass the language

---

[5] Students who were part of the freshened sample followed the same procedures: If their school records indicated that they were a language minority student, they received the language screener.

Table 2-2.—Direct child assessment, by round of data collection

Direct child assessment	1998-99 school year Fall-kindergarten	1998-99 school year Spring-kindergarten	1999-2000 school year Fall-first grade	1999-2000 school year Spring-first grade
Language screener (OLDS)	X	/	/	/
Reading (language and literacy)	X	X	X	X
Mathematical thinking	X	X	X	X
General knowledge (science and social studies)	X	X	X	X
Psychomotor	X			
Height and weight	X	X	X	X

NOTE: The columns to the right of each construct correspond to the round of administration. Rounds that included the construct are marked with an “X.” A “/” indicates that the OLDS was given to language minority students new in the spring, or who did not pass the cut score in the English version during the previous OLDS administration.

screener but spoke Spanish, he or she was administered a Spanish translated form of the mathematics assessment and an alternate form of the language screener, the Spanish version of the OLDS (Spanish OLDS). The Spanish OLDS that was administered is similar in content to the English OLDS and measures the same constructs. For further information on the language screener, please refer to the ECLS-K Base Year Public-Use User’s Manual (NCES 2001-029), February 2001, or the ECLS-K Restricted-Use Base Year User’s Manual (NCES 2000-097), August 2000.

Cognitive Components. The kindergarten-first grade cognitive assessment focused on three general areas of competence: (1) language use and literacy (reading); (2) mathematics; and (3) knowledge of the social and physical world, referred to as “general knowledge.” The assessment did not ask the children to write anything or to explain their reasoning; rather, they pointed to their answers or responded orally to complete the tasks. The assessment battery included the use of small easels with pictures, letters of the alphabet, words, short sentences, numbers, or number problems and a computer laptop for the assessor to enter children’s responses. The same two-stage cognitive assessment approach used in the base year was repeated in the first grade data collection. The purpose of the two-stage design was to maximize the accuracy of measurement and reduce administration time by using the children’s responses from the first stage to route the children to the appropriate level of difficulty in the second stage.[6] With the exception of the reading assessment domain, the same mathematics and general knowledge two-stage

[6] For details on the two-stage assessment design, see the ECLS-K Base Year Public-Use User's Manual (NCES 2001-029), February 2001, or the ECLS-K Restricted-Use Base Year User's Manual (NCES 2000-097), August 2000.

assessment batteries were re-administered in first grade. Analysis of the reading scores from spring-kindergarten showed a higher than expected number of respondents scoring near the ceiling. Therefore, to eliminate the possibility of ceiling effects, the number of reading items was increased by adding more difficult vocabulary words and text. The mathematical thinking and general knowledge assessments showed no such ceiling effects and so these assessments were not modified.

Language and Literacy. The language and literacy (reading) assessment included questions designed to measure basic skills (print familiarity, letter recognition, beginning and ending sounds, rhyming sounds, 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.

The reading assessment contained five proficiency 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 that comprised the earlier levels as well. These 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 words by sight; and (5) reading words in context.

Mathematical Thinking. The 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 for the appropriate parts of the assessment.

The items in the mathematics assessment could also be grouped into five proficiency levels, though the math clusters were less homogeneous in content than the reading clusters. The clusters of math 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 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.

General Knowledge. The general knowledge assessment battery consisted of items that measure knowledge in the natural sciences and social studies in a single scale. The science domain measures 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 domain included questions that measure children's knowledge in a wide range of disciplines such as history, government, culture, geography, economics, and law. The science domain 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. It 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 also measured.

The subject matter content of the general knowledge assessment domain was too diverse and the items insufficiently ranked or graded to permit the formation of a set of proficiency levels. A score was calculated to represent each child's breadth and depth of understanding and knowledge of the world around them.

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 was not included in subsequent rounds. The anthropometric component consisted of recording the children's height and weight in order to measure their physical growth and development. A Shorr Board (for measuring height) and a digital bathroom scale were used to obtain the height and weight measurements, which were recorded on a height and weight recording form.

2.3 Parent Interview

The first grade parent interviews were conducted using a computer-assisted interview (CAI). The parent interviews were conducted primarily in English, but provisions were made to interview parents who spoke other languages. Bilingual interviewers were trained to conduct the parent interview in either English or Spanish. In fall-first grade if the interview was conducted in Spanish, the interviewer used a hard-copy questionnaire and then entered the respondent's answers into the CAI program. In spring-first grade, a Spanish CAI instrument was used when needed. The spring-first grade interview was also translated into Hmong and Mandarin. Such interviews were conducted with paper and pencil rather than CAI.

Fall Parent Interview. Most families interviewed during fall-first grade had been interviewed in the fall and/or spring of the base year. Typically the respondent for the fall-first grade parent interview was the mother of the child; however, the respondent could be a father, stepparent, adoptive parent, foster parent, grandparent, another relative, or a nonrelative guardian. The respondent had to be knowledgeable about the child's care and education, be 18 years of age or older, and be living in the household with the child. Respondents for the parent interview were selected according to the following order of preference:

1. Respondent from the previous round;
2. The child's mother;
3. Another parent or guardian; and
4. Another household member.

The fall interview was shorter and more narrowly focused than the parent interviews conducted during the base year and during spring-first grade due to its focus on summer learning experiences. It included sections on children's summer activities, including vacations; attendance at summer school, school enrichment programs, and summer camp; receipt of tutoring; and participation in special activities such as music, dance, or swimming lessons. It also obtained information on the types of activities parents engaged in with their children during the summer; children's summer child care arrangements; and the availability of community resources such as recreation centers, community pools and parks, boys' or girls' clubs, and libraries.

Spring Parent Interview. The majority of parents participating in the spring-first grade data collection were interviewed in the fall and/or spring of the base year. However, the sample was freshened to include children who had not attended kindergarten in the United States during the 1998-99 school year (see chapter 5). The order of preference for the respondent to the parent interview was the same as described for the fall-first grade interview: (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 household member.

The parent interview for the spring-first grade data collection was extensive and asked questions covering first grade school experiences, child care, parent characteristics, and family health. Table 2-3 provides an overview of the topics covered in the base year and first grade data collections. As can be seen in the table, key topics such as family structure, parental involvement in school, and the child's home environment and cognitive activities are covered in most rounds. Other topics, such as parent income, employment, and education, are measured at least once in each school year. Although the general content areas are similar across the questionnaires, the items were updated, where appropriate, to reflect the first grade context. For example, in spring-first grade, the series of questions asking, "In a typical week, how often do you or any other family member do the following things with the child," the activity "practice reading, writing or working with numbers" was added to the list of activities. Similarly, in spring-first grade questions about whether the child was tutored on a regular basis and in what subjects were added to the questionnaire.

Parents or guardians of children added to the sample through freshening or who were nonrespondents in the prior rounds in spring-first grade were asked to complete a supplementary section that contained key items asked in previous rounds. The supplementary section included questions about the children's child care arrangements in the year prior to kindergarten, whether they had attended Head Start in the year before kindergarten, their birth weight, and whether they had been born prematurely. In addition, parents were asked whether the children had ever received therapy services or taken part in programs for children with disabilities and, if so, the types of services or programs they had participated in. Finally, information about the children's backgrounds were collected including questions about the primary language spoken in the home, the country in which the child was born, and whether the children's mothers had worked for pay outside the home between when the children were born and the start of kindergarten.

2.4 Spring-First Grade Teacher Questionnaires

During spring-first grade data collection, each teacher received a self-administered questionnaire consisting of three distinct parts. The first section, part A, asked about the teacher's class

Table 2-3.—ECLS-K parent interview by major content topics and round of data collection

Parent questionnaire	1998-99 school year		1999-2000 school year	
	Fall-kindergarten	Spring-kindergarten	Fall-first grade	Spring-first grade
Family structure	X	X	X	X
Demographics	X	X	X	X
Household roster	X	X	X	X
Marital status	X	X	X	X
Immigration status		X		X
Primary language(s) spoken in home	X	/	/	/
Parent's involvement with child's school		X	X	X
Child care	X		X	X
Current arrangements with relatives	X		X	X
Current arrangements with nonrelatives	X		X	X
Current arrangements with centers	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		/
Birth weight	X	/	/	/
Physical functioning	X	/	/	/
Services for children with special needs	X	/	/	/
Social skills rating	X	X		X
Home environment and cognitive activities	X	X	X	X
Frequency of literacy activities	X	X	X	X
Computer use		X	X	X
Television viewing		X	X	X
Summer activities and time use			X	
Parental educational expectations for child	X		X	X
Neighborhood		X	X	X
Safety		X		X
Resources (e.g., community center, library)			X	
Parent education	X	/	/	/
Parent employment	X			X
Parent income		X		X

Table 2-3.—ECLS-K parent interview by major content topics and round of data collection (continued)

Parent questionnaire	1998-99 school year Fall-kindergarten	1998-99 school year Spring-kindergarten	1999-2000 school year Fall-first grade	1999-2000 school year Spring-first grade
Welfare and other public assistance use	X	X		X
Parent/child interaction		X		X
Parent discipline		X		
Parent health and emotional well-being		X		
Relationships and social support	X	X		
Marital satisfaction		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	/	/	/

NOTE: The columns to the right of each construct correspond to the round of questionnaire administration. Rounds that included the construct are marked with an "X." Content areas asked only of new parent respondents in each round are marked with a "/."

and classroom characteristics. It was designed to collect data about the composition and demographics of the children in the class and was completed only by teachers of sampled children, unlike the base year when it was completed by all kindergarten teachers in the school, regardless of whether they taught a sampled child. Part B addressed more specific questions on class organization, typical class activities, and evaluation methods, as well as teacher views on school readiness, school environment, and overall school climate. Background questions about the teacher were also included in this section. Teachers were asked to complete one copy of part C for each of the sampled children in their classrooms; in this part, teachers were asked to respond to 21 questions about the child's academic performance. The academic rating scale (ARS) gathered data on each sampled child's skills in areas of language and literacy, general knowledge, and mathematical thinking. Part C also included questions from the social rating scale (SRS) that collected data on five areas of children's social skills. The ARS and SRS are described in more detail in sections 2.4.1 and 2.4.2, respectively.

Two different versions of the teacher questionnaire were available. The first was for teachers of children who had made the transition to the first grade or any higher elementary school grade, and the second was for teachers of children who were repeating or attending the second year of kindergarten. Any kindergarten-specific items were updated to first grade items in the questionnaire for first grade teachers. For example, in asking about the children's skills, first grade teachers were asked whether the children read first grade books independently with comprehension or read first grade books fluently. In kindergarten, the teachers were asked about the children reading "simple" books independently. Similarly, first grade teachers were asked whether the children compose stories with a clear beginning, middle, and end, while kindergarten teachers were asked whether the children compose simple stories. Teachers were asked about transition practices for children moving from kindergarten to first grade, while in the base year, the teachers were asked about transition practices to kindergarten. The questionnaire for kindergarten teachers combined elements from both the fall- and spring-base year teacher questionnaires.

Table 2-4 shows the overall structure of the base year and spring-first grade teacher questionnaires and the distribution of topics covered.

2.4.1 Content of the Academic Rating Scale

There are three scales of the ARS: language and literacy, general knowledge, and mathematics. Each of these is described below. 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 first grade based on the teacher's past observation and experience with the child.

- The **Language and Literacy** section of the ARS asks teachers to rate each child's proficiency in expressing ideas (one item), listening (one item), reading on grade level (four items), and writing (two items). In addition, teachers rate the child's computer literacy (one item).
- The **General Knowledge** section of the ARS asks teachers to rate each child's skills and knowledge in social studies (three items) and science (three items).
- In the **Mathematics** section, teachers rate each child on seven items that tap the following skills: understanding place values, making reasonable estimates of quantities, solving number problems, using various strategies, organizing and analyzing data (graphing), and measuring accurately.

See chapter 3, section 3.3 for scale scores, value ranges, means, and standard deviations for the ARS.

Table 2-4.—Teacher questionnaires

Teacher questionnaire	1998-99 school year		1999-2000 school year	
	Fall-kindergarten	Spring-kindergarten	Spring-first grade (First grade teacher)	Spring-first grade (Kindergarten teacher)
Parts A and B				
Description of class—age, race-ethnicity, and sex distribution	X[a]		X[a]	X[a]
Class organization				
Types of activity/interest areas	X[b]	/	X[a]	X[b]
Types of materials/resources		X[a]	X[a]	X[a]
Instructional time in different subjects		X[a]	X[a]	X[a]
Child vs. teacher initiated activities	X[b]	X[a]	X[a]	X[b]
Classroom characteristics				
Children with special needs		X[a]	X[a]	X[a]
Classroom aides		X[a]	X[a]	X[a]
Class assignment and grouping		X[a]	X[a]	X[a]
Behavior of children in classroom	X[a]	X[a]	X[a]	X[a]
Parent involvement		X[a]	X[a]	X[a]
Share progress information with parents		X[a]	X[a]	X[a]
Professional development		X[a]	X[b]	X[a]
Teachers' evaluation and grading practices	X[b]	/	X[a]	X[b]
Teachers' views on school readiness	X[b]	/	X[b]	X[b]
Perceptions about school climate	X[b]	/	X[b]	X[b]
Perception of personal influence on policies and classroom planning	X[b]	/	X[b]	X[b]
Teacher demographic information	X[b]	/	X[b]	X[b]
Teacher experience and education	X[b]	/	X[b]	X[b]
Job satisfaction	X[b]	/	X[b]	X[b]
Transition to school activities	X[b]	/	X[b]	X[b]

Table 2-4.—Teacher questionnaires (continued)

Teacher questionnaire	1998-99 school year		1999-2000 school year	
	Fall-kindergarten	Spring-kindergarten	Spring-first grade (First grade teacher)	Spring-first grade (Kindergarten teacher)
Part C				
Indirect child cognitive evaluation by teacher	X	X	X	X
Language and literacy, mathematics, general knowledge	X	X	X	X
Social skills	X	X	X	X
Additional information on sampled child		X	X	X
Participation in special services and programs		X	X	X
Overall academic skills and physical activity levels		X	X	X
Reading group participation		X	X	X

NOTE: The columns to the right of each construct correspond to the waves of questionnaire administration. Waves that included the construct are marked with an "X." Content areas asked only of new teacher participants are marked with a "/."

[a] Topic is in Teacher Questionnaire A.

[b] Topic is in Teacher Questionnaire B.

2.4.2 Teacher Social Rating Scale

Teachers rated individual students' social development on part C of the teacher questionnaire. These items are intended to measure approaches to learning, self-control, and interpersonal skills. The items were rated on a scale of one (Never) to four (Very often). Five scales are formed from these items. Three of the scales capture positive aspects of children's development, and two represent problem behaviors. See chapter 3, section 3.5.1 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.

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

- **Externalizing Problem Behaviors** Scale (Teacher SRS) include acting out behaviors. Five items on this scale rate the frequency with which a child argues, flights, gets angry, acts impulsively, and disturbs ongoing activities.

- 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 Elementary Scale A ("How Often?") (Gresham and Elliott, 1990).

2.5 Special Education Teacher Questionnaires

In the spring-first 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). Questionnaires were given to these special education teachers. If a child received special education services from more than one special education teacher, a field supervisor determined the child's primary special education teacher. The primary special education teacher 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 equipment.

The spring-first grade special education teacher questionnaires were very similar to the one used in the spring of the base year. The only difference was the updating of questions to refer to the transition from kindergarten into first grade rather than into kindergarten. Table 2-5 provides a summary of the content areas addressed in the special education teacher questionnaires in both the base year and spring-first grade. 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.

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(ren) was asked to complete a copy of part B for each sampled child she/he was responsible for overseeing.

2.6 Adaptive Behavior Scale

The ABS (Lambert, Nihira, and Leland, 1993) was completed for all sampled children who were identified in each of the data collection rounds as excluded from the direct child assessment due to a disability. A child with a disability was excluded from the assessment if he or she needed the assessment administered in Braille, enlarged print, or sign language, or if the child's IEP specifically prohibited the child from taking standardized assessments. This questionnaire was completed by the child's primary special education teacher and asked the teacher to provide ratings of the sampled child in three domains: independent functioning (domain I), language development (domain IV), and numbers and time (domain V).

Table 2-5.—Special education teacher questionnaire topics in spring-kindergarten and spring-first grade*

Teachers of sampled students with IEPs questionnaire	1998-99 school year spring-kindergarten	1999-2000 school year spring-first grade
Part A (Teacher Level)		
Teacher's sex	X	X
Teacher's age	X	X
Teacher's race-ethnicity	X	X
Teaching experience	X	X
Educational background	X	X
Special education teacher background	X	X
Location of service provision	X	X
Student load per week	X	X
Part B (Child Level)		
Disability category	X	X
IEP goals for the school year	X	X
Extent of services	X	X
Types of services provided for the year	X	X
Primary placement	X	X
Teaching practices, methods, and materials	X	X
Assistive technologies used by child	X	X
General education goals, expectations and assessments	X	X
Collaboration/communication with child's general education teacher	X	X
Frequency of communicating with child's parents	X	X
Receipt of formal evaluations in the past year	X	X

* Data collected only in the spring of each school year.

2.7 School Administrator Questionnaire

The school principal, administrator, or headmaster at the school attended by the sampled child, was asked to complete the school administrator questionnaire in the spring of 2000. This self-administered questionnaire was intended to gather information about the school, student body, teachers, school policies, and administrator characteristics. The questionnaire was divided into nine sections. The first seven 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.

Two versions of the questionnaire were used. One version was aimed at updating information from returning schools, that is, schools that were part of the base year. The second version of the questionnaire was given to school administrators of new schools. Both questionnaires incorporated questions from the spring-kindergarten school administrator's questionnaire. Table 2-6 shows the topics covered in the spring-kindergarten and spring-first grade questionnaires. The spring-first grade questionnaire for returning schools gathered less information than the questionnaire for new schools, as indicated in the table by "--".

For nonresponding and late-responding schools, interviewers were trained to visit the school and encourage the school administrators to complete the questionnaire. If necessary the interviewers were to sit down with the administrators to help them fill out the questionnaire. However, if the school administrators were still reluctant to complete the full questionnaire, the interviewers were instructed to obtain key information. This key information covered such topics as school sector and focus; the school environment, particularly, the safety of the school; school policies and practices; school programs for special populations; staffing and teacher characteristics; and principal characteristics.

2.8 School Facilities Checklist

ECLS-K supervisors completed the facilities checklist during their visits to the school in the spring of first grade. The facilities checklist collects information about the (1) availability and condition of selected school facilities such as classrooms, gymnasiums, toilets, etc., (2) presence and adequacy of security measures, (3) presence of environmental factors that may affect the learning environment, and (4) overall learning climate of the school. An additional set of questions on portable classrooms, that were not included in the spring-kindergarten facilities checklist, was added to the checklist completed in the spring-first grade data collection. (See chapter 5, section 5.5 for more detail on the collection of these data.)

Table 2-6.—School administrator questionnaire, spring-kindergarten and spring-first grade

School questionnaire	Spring-kindergarten	1999-2000 school year spring-first grade Returning schools	New schools
School characteristics	X	--	X
Type of school	X		X
Admission requirements	X		
School size	X	X	X
Student characteristics	X	X	X
Race-ethnicity of students	X	X	X
Children eligible for special services	X	X	X
Types of kindergarten programs	X		
School facilities and resources	X	--	X
Computer equipment	X	X	X
Community characteristics and school safety	X	X	X
Teaching and other school staff characteristics	X	X	X
Range of salary paid to teachers	X		X
Race-ethnicity of staff	X	X	X
School policies and programs	X	--	X
Assessments, testing, and retention	X	X	All grades
School-family-community connections	X	--	X
Programs and activities for families	X		X
Parent involvement and participation	X	X	X
Programs for special populations	X	X	X
ESL and bilingual education	X	X	X
Special education	X	--	X
Gifted and talented	X		X
Principal characteristics	X	X	X
Sex, race-ethnicity, age of principal	X	X	X
Experience and education	X	X	X
School governance and climate	X	X	X
Goals and objectives for teachers	X	X	X
School functioning and decisionmaking	X	X	X

NOTE: "--" indicates that fewer details on the topic were collected than for new schools.

2.9 School Records Abstract Form

School staff completed the student records abstract form for each sampled child in the spring of kindergarten and first grade. This instrument was used to obtain information about the child's attendance record, presence of and details on a child's IEP and information about the type of language or English proficiency screening that the school used. A copy of each child's report card was also obtained. The spring-first grade version 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 first grade followup. 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 detail on the collection of these forms.)

REFERENCES

Cooper, H., Nye, B., Charlton, K., Lindsay, J., and Greathouse, S. (1996). The effects of summer vacation on achievement test scores: A narrative and meta-analytic review. *Review of Educational Research* 66(3), 227-268.

Entwisle, D.R., and Alexander, K.L. (1992). Summer setback: Race, poverty, school composition, and mathematics achievement in the first two years of school. *American Sociological Review* 57(1), 72-84.

Entwisle, D.R., and Alexander, K.L. (1994). Winter setback: The racial composition of schools and learning to read. *American Sociological Review* 59(3), 446-460.

Gresham, F.M. and Elliott, S.N. (1990). *Elementary Scale A ("How Often?").* American Guidance Service, Inc.

Heyns, B. (1978). *Summer learning and the effects of schooling*. New York: Academic Press.

Lamert, N., Nihira, K., and Leland, H. (1993). *Adaptive Behavior Scale, Second Edition.* The American Association on Mental Retardation.

National Center for Education Statistics. (2001). *ECLS-K Base Year Public-Use User's Manual.* (NCES 2001-029).

National Center for Education Statistics. (2001). *ECLS-K Base Year Restricted-Use Head Start Data Files and Electronic Code Book.* (NCES 2001-025).

National Center for Education Statistics. (2001). *ECLS-K Base Year Restricted-Use Salary and Benefits Data Files and Electronic Code Book.* (NCES 2001-014).

National Center for Education Statistics. (2000). *ECLS-K Restricted-Use Base Year User's Manual.* (NCES 2000-097).

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

Several types of scores are used in the Early Childhood Longitudinal Study-Kindergarten Class of 1998-99 (ECLS-K) to describe children's cognitive and social development during kindergarten and first grade. These scores are for the direct cognitive assessment, the academic rating scale (ARS), and the social rating scale (SRS). 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 to use each cognitive test score are also provided in this chapter.

3.1 Direct Cognitive Assessment

The direct cognitive assessment contained items in reading, mathematics, and general knowledge in the fall and spring of kindergarten, and fall and spring of first grade.[2] In each subject area, children received a 12- to 20-item routing test. Performance on the routing items guided the selection and administration of one of several second-stage forms. The second-stage form contained items of appropriate difficulty for the level of ability indicated by the routing items.[3]

There are five different types of scores that can be used to describe children's performance on the direct cognitive assessment: (1) number-right scores and (2) Item Response Theory (IRT) scale scores, which measure children's performance on a set of test questions with a broad range of difficulty; (3) standardized scores (T-scores), which report children's performance relative to their peers; (4) criterion-referenced proficiency-level scores; and (5) proficiency probability scores, which evaluate children's performance with respect to subsets of test items that mark specific skills. Kindergarten and first grade test responses were pooled to stabilize the longitudinal estimates. As a result, the re-estimated kindergarten IRT scores, T-scores, and proficiency probability scores in this database differ very slightly from the scores in the base year kindergarten-only file previously released.[4] See chapter 2 for a description of the ECLS-K assessment battery.

---

[1] This user's manual is applicable to the data gathered during the 1999-2000 school year; information contained in this manual about data gathered during the 1998-99 school year (base year of the study) is provided primarily for background and comparison purposes.

[2] The fall-first grade sample was a 27 percent subsample of the children in 30 percent of the base-year sampled schools. See chapter 4 for more detail on the subsampling procedures.

[3] See chapter 2, section 2.2, for additional information on the two-stage process for the direct cognitive assessments.

[4] Please see ECLS-K, Base Year Public-Use Data File, Kindergarten Class of 1998-99: Data Files and Electronic Code Book (NCES 2001-029), February 2001.

Tables 3-1 through 3-6 show the types of scores, variable names, descriptions, ranges, means, and standard deviations 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), or 4 (spring-first grade).

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 tests that are the same for all children. However, when these scores are for tests that vary in average difficulty, they are not comparable to each other. For example, a student who took the middle difficulty mathematics second-stage form would probably have gotten more questions correct if he or she had taken the easier low form and fewer correct if the more difficult high form had been administered. For this reason, raw number-right scores are reported in the database only for the first-stage (routing) tests, which were the same for all children. Each routing test consisted of sets of items spanning a wide range of skills. For example, the reading routing test had four questions each on letter recognition, recognizing beginning sounds, recognizing ending sounds, reading simple sight words, and selecting words in the context of a sentence. An analyst might use the routing test number-right scores to report actual performance on this particular set of tasks.

See table 3-1 for the variable names, descriptions, ranges, weighted means, and standard deviations for the number-right scores.

3.1.2 Item Response Theory Scale Scores

Scores based on the full set of test items were calculated using IRT procedures. IRT made it possible to calculate scores that could be compared regardless of which second-stage form a child took. IRT uses the pattern of right, wrong, and omitted responses to the items actually administered in a test 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 test, plus a core set of items shared among the different second-stage forms, 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 test forms had been administered.

Table 3-1.—Direct cognitive assessment: number-right scores*

Variable	Description	Range of values	Weighted mean	Standard deviation
C1RROUNR	C1 Reading Routing Test, - Number Right	0 - 20	5.7	3.9
C1MROUNR	C1 Mathematics Routing Test, - Number Right	0 - 16	4.4	2.9
C1GROUNR	C1 General Knowledge Routing Test, - # Right	0 - 12	4.7	2.9
C2RROUNR	C2 Reading Routing Test, - Number Right	0 - 20	9.8	4.2
C2MROUNR	C2 Mathematics Routing Test, - Number Right	0 - 16	7.1	3.3
C2GROUNR	C2 General Knowledge Routing Test, - # Right	0 - 12	6.2	3.0
C3RROUNR	C3 Reading Routing Test, - Number Right	0 - 20	11.6	4.3
C3MROUNR	C3 Mathematics Routing Test, - Number Right	0 - 16	8.8	3.4
C3GROUNR	C3 General Knowledge Routing Test, - # Right	0 - 12	7.3	2.9
C4RROUNR	C4 Reading Routing Test, - Number Right	0 - 20	16.2	3.7
C4MROUNR	C4 Mathematics Routing Test, - Number Right	0 - 16	11.6	3.1
C4GROUNR	C4 General Knowledge Routing Test, - # Right	0 - 12	8.4	2.7

* See chapter 7, section 7.3 for variable naming conventions.

IRT has several other advantages over raw number-right scoring. By using the overall pattern of right and wrong responses 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, a correct difficult item is, in effect, assumed to have been guessed. 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, in effect, 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 test questions. Finally, IRT scoring makes possible longitudinal measurement of gain in achievement over time, even though the tests administered are not identical at each point. The common items present in the routing test and in overlapping second-stage forms allow the test scores to be placed on the same scale, even as the two-stage test design adapts to children's growth over time. As noted earlier, kindergarten and first-grade test responses were pooled to stabilize the longitudinal estimates. As a result, the re-estimated kindergarten IRT scores in this database differ very slightly from the IRT scores in the base year kindergarten-only file previously released. In addition, the maximum value of the reading scale score has been expanded from 72 to 92 to include the more difficult items administered in the first-grade tests.

The IRT scale scores in the database represent estimates of the number of items students would have answered correctly if they had taken all of the 92 questions in the first- and second-stage reading forms, the 64 questions in all of the mathematics forms, and the 51 general knowledge items. These scores are not integers because they are probabilities of correct answers, summed over all items in the pool. Gain scores may be obtained by subtracting the estimated number-right at fall-kindergarten from the estimated number-right at spring-kindergarten, spring-kindergarten from spring-first grade, etc. (Note that scores for different subject areas are not comparable to each other because they are based on different numbers of test questions, 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.)

See table 3-2 for variable names, descriptions, ranges, weighted means, and standard deviations for the IRT scale scores.

Table 3-2.—Direct cognitive assessment: Item Response Theory scale scores*

Variable	Description	Range of values	Weighted mean	Standard deviation
C1RSCALE	C1 Reading IRT Scale Score	0 - 92	22.7	8.6
C1MSCALE	C1 Mathematics IRT Scale Score	0 - 64	19.3	7.1
C1GSCALE	C1 General Knowledge IRT Scale Score	0 - 51	22.1	7.4
C2RSCALE	C2 Reading IRT Scale Score	0 - 92	32.5	10.9
C2MSCALE	C2 Mathematics IRT Scale Score	0 - 64	27.2	8.7
C2GSCALE	C2 General Knowledge IRT Scale Score	0 - 51	26.8	7.8
C3RSCALE	C3 Reading IRT Scale Score	0 - 92	38.0	12.7
C3MSCALE	C3 Mathematics IRT Scale Score	0 - 64	32.4	9.6
C3GSCALE	C3 General Knowledge IRT Scale Score	0 - 51	30.0	7.9
C4RSCALE	C4 Reading IRT Scale Score	0 - 92	54.8	14.2
C4MSCALE	C4 Mathematics IRT Scale Score	0 - 64	42.8	9.5
C4GSCALE	C4 General Knowledge IRT Scale Score	0 - 51	34.0	7.7

* See chapter 7, section 7.3 for variable naming conventions.

3.1.3 Standardized Scores (T-Scores)

T-scores provide norm-referenced measurements of achievement, that is, estimates of achievement level *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 mean that group members have mastered a particular set of skills, only that their 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, they provide information on *status compared to children's peers*, while the IRT scale scores and proficiency scores represent *status with respect to achievement on a particular criterion set of test items*. The T-scores can 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 reported in the database are transformations of the IRT theta (ability) estimates, rescaled to a mean of 50 and standard deviation of ten using cross-sectional sample weights for each wave of data. For example, a T-score of 55 (C1RTSCOR) represents a reading achievement level that is one-half of a standard deviation higher than the mean for the fall-kindergarten population represented by the tested sample of ECLS-K participants.

See table 3-3 for variable names, descriptions, ranges, weighted means, and standard deviations for the standardized (T) scores.

Table 3-3.—Direct cognitive assessment: standardized scores*

Variable	Description	Range of values	Weighted mean	Standard deviation
C1RTSCOR	C1 Reading T-Score	0 - 90	50.0	10.0
C1MTSCOR	C1 Mathematics T-Score	0 - 90	50.0	10.0
C1GTSCOR	C1 General Knowledge T-Score	0 - 90	50.0	10.0
C2RTSCOR	C2 Reading T-Score	0 - 90	50.0	10.0
C2MTSCOR	C2 Mathematics T-Score	0 - 90	50.0	10.0
C2GTSCOR	C2 General Knowledge T-Score	0 - 90	50.0	10.0
C3RTSCOR	C3 Reading T-Score	0 - 90	50.0	10.0
C3MTSCOR	C3 Mathematics T-Score	0 - 90	50.0	10.0
C3GTSCOR	C3 General Knowledge T-Score	0 - 90	50.0	10.0
C42RTSCOR	C4 Reading T-Score	0 - 90	50.0	10.0
C4MTSCOR	C4 Mathematics T-Score	0 - 90	50.0	10.0
C4GTSCOR	C4 General Knowledge T-Score	0 - 90	50.0	10.0

* See chapter 7, section 7.3 for variable naming conventions.

3.1.4 Proficiency Scores

Proficiency scores 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. Since the ECLS-K direct cognitive child assessment was a two-stage design (where not all children were administered all items), information on children's specific proficiencies are presented in two ways: proficiency scores (raw scores) and proficiency probability scores (IRT-based scores). In most situations, analysts use the proficiency probability scores in analyzing children's specific reading and mathematics knowledge and skills. Clusters of assessment questions having similar content and difficulty were included at several points along the score scale of the reading and mathematics assessments. No proficiency scores were computed for the general knowledge test because the questions did not follow a hierarchical pattern. The following proficiencies were identified in the reading and mathematics assessments.

Reading:

- **Letter recognition:** identifying upper- and lower-case letters by name
- **Beginning sounds:** associating letters with sounds at the beginning of words
- **Ending sounds:** associating letters with sounds at the end of words
- **Sight words:** recognizing common words by sight
- **Comprehension of words in context:** reading words in context

Mathematics:

- **Number and shape:** identifying some one-digit numerals, recognizing geometric shapes, and one-to-one counting of up to ten objects
- **Relative size:** reading all single-digit numerals, counting beyond ten, recognizing a sequence of patterns, and using nonstandard units of length to compare objects
- **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
- **Addition/subtraction:** solving simple addition and subtraction problems
- **Multiplication/division:** solving simple multiplication and division problems and recognizing more complex number patterns

Clusters of items provide a more reliable test of proficiency than do single items because of the possibility of guessing; it is very unlikely that a student who has not mastered a particular skill would be able to guess enough answers correctly to pass a four-item cluster. 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 have indicated nonmastery at higher levels. Only a very small percentage of students in kindergarten and first 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 four rounds of data collection, only about 6 percent of reading response patterns, and about 5 percent of math test 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 are due to children guessing.

Proficiency-level (dichotomous) scores and proficiency probability-level (continuous) scores are two types of proficiency scores used in the ECLS-K. The following is a description of these scores.

Proficiency-Level Scores (Dichotomous)

The proficiency-level scores reflect the children's raw ECLS-K direct cognitive assessment scores. These scores are intended for very distinct kinds of analysis. When using the dichotomous proficiency-level scores, it is necessary to take into account the sources of missing data: that scores may be missing because certain items were not administered to particular children. For this reason, most analysts prefer to use the proficiency probability scores, which already compensate for missing data.

For each proficiency level, a score of one was assigned to children who correctly answered at least three of the four items in the cluster, and a score of zero was given if at least two items were incorrect or don't know. If children did not answer enough items for pass or fail to be determined for a particular cluster, a pass/fail score was assigned only if the remaining proficiency scores indicated a level that was unambiguous. That is, a "fail" may be inferred if there were easier cluster(s) that had been failed and no higher cluster passed; and a "pass" may be assumed if harder cluster(s) were passed and no easier one failed. In the case of ambiguous (e.g., pass, blank, fail, where the blank could be either a pass or a fail) or contradictory (e.g., fail, blank, pass) patterns, no imputation is possible without reference to other information, such as the IRT-based scores.

Averaging the sample subgroup's zero and one scores for a particular proficiency cluster results in an estimate of the proportion of children in the subgroup *who answered that cluster* and had mastered the material at that level. The difference between this average at two points in time represents the proportion of these children who attained mastery during that time period. To the extent that there is non-imputable missing data for these levels, the scores are *not* designed to extrapolate to the entire population of kindergarten children. These scores simply show that *of the children who took the items represented by the level*, how many passed the level. In the reading routing test, for example, the harder items (levels four and five) were not administered to children who had difficulty with the easier tasks. Therefore, the missing data for the upper levels consists primarily of children who were not able to get to the more difficult material in the test. It would be incorrect to average the zero and one scores for reading level five and assume that the average represented a population proportion, since the missing cases would have lowered the average substantially. For example, the passing rate for reading level 4 in fall-kindergarten was .04, while the passing rate for level 5 was .26. This does *not* indicate that more children in the population would be likely to do well at level 5 than at level 4. It is a result of the test being discontinued for children who could not answer the items, so the .26 proportion is based only on the small sample of the very best readers: those who were able to complete the whole reading routing test.

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

Proficiency Probability Scores (Continuous)

The proficiency probability scores are based on the same clusters of items as the proficiency-level scores but differ from them in several ways. They are continuous rather than dichotomous and can take on any value between zero and one. They are estimates based on overall performance rather than counts of actual item responses. They are also estimates for all children with scorable test data, not only for the ones who were administered the test items in the cluster.

Table 3-4.—Direct cognitive assessment: proficiency-level scores*

Variable	Description	Range of values	Weighted mean	Standard deviation
C1RPROF1	C1 Prof 1 - Letter Recognition	0 - 1	0.65	0.48
C1RPROF2	C1 Prof 2 - Beginning Sounds	0 - 1	0.30	0.46
C1RPROF3	C1 Prof 3 - Ending Sounds	0 - 1	0.18	0.38
C1RPROF4	C1 Prof 4 - Sight Words	0 - 1	0.04	0.20
C1RPROF5	C1 Prof 5 - Word in Context	0 - 1	0.26	0.44
C1MPROF1	C1 Prof 1 - Count, Number, Shape	0 - 1	0.89	0.31
C1MPROF2	C1 Prof 2- Relative size	0 - 1	0.55	0.50
C1MPROF3	C1 Prof 3 - Ordinality, sequence	0 - 1	0.20	0.40
C1MPROF4	C1 Prof 4 - Add/Subtract	0 - 1	0.04	0.20
C1MPROF5	C1 Prof 5 - Multiply/Divide	0 - 1	0.02	0.13
C2RPROF1	C2 Prof 1 - Letter Recognition	0 - 1	0.92	0.28
C2RPROF2	C2 Prof 2 - Beginning Sounds	0 - 1	0.70	0.46
C2RPROF3	C2 Prof 3 - Ending Sounds	0 - 1	0.50	0.50
C2RPROF4	C2 Prof 4 - Sight Words	0 - 1	0.14	0.35
C2RPROF5	C2 Prof 5 - Word in Context	0 - 1	0.24	0.43
C2MPROF1	C2 Prof 1- Count, Number, Shape	0 - 1	0.95	0.21
C2MPROF2	C2 Prof 2 - Relative Size	0 - 1	0.82	0.38
C2MPROF3	C2 Prof 3 - Ordinality, Sequence	0 - 1	0.54	0.50
C2MPROF4	C2 Prof 4 - Add/Subtract	0 - 1	0.17	0.37
C2MPROF5	C2 Prof 5 - Multiply/Divide	0 - 1	0.04	0.19
C2RPROF1	C3 Prof 1 - Letter Recognition	0 - 1	0.95	0.22
C3RPROF2	C3 Prof 2 - Beginning Sounds	0 - 1	0.81	0.39
C3RPROF3	C3 Prof 3 - Ending Sounds	0 - 1	0.66	0.48
C3RPROF4	C3 Prof 4 - Sight Words	0 - 1	0.25	0.43
C3RPROF5	C3 Prof 5 - Word in Context	0 - 1	0.32	0.46
C3MPROF1	C3 Prof 1- Count, Number, Shape	0 - 1	0.94	0.24
C3MPROF2	C3 Prof 2 - Relative Size	0 - 1	0.89	0.31
C3MPROF3	C3 Prof 3 - Ordinality, Sequence	0 - 1	0.71	0.46
C3MPROF4	C3 Prof 4 - Add/Subtract	0 - 1	0.32	0.47
C3MPROF5	C3 Prof 5 - Multiply/Divide	0 - 1	0.08	0.27
C4RPROF1	C4 Prof 1 - Letter Recognition	0 - 1	0.99	0.11
C4RPROF2	C4 Prof 2 - Beginning Sounds	0 - 1	0.93	0.26
C4RPROF3	C4 Prof 3 - Ending Sounds	0 - 1	0.89	0.31
C4RPROF4	C4 Prof 4 - Sight Words	0 - 1	0.78	0.42
C4RPROF5	C4 Prof 5 - Word in Context	0 - 1	0.52	0.50
C4MPROF1	C4 Prof 1- Count, Number, Shape	0 - 1	0.94	0.24
C4MPROF2	C4 Prof 2 - Relative Size	0 - 1	0.96	0.19
C4MPROF3	C4 Prof 3 - Ordinality, Sequence	0 - 1	0.91	0.29
C4MPROF4	C4 Prof 4 - Add/Subtract	0 - 1	0.70	0.46
C4MPROF5	C4 Prof 5 - Multiply/Divide	0 - 1	0.25	0.43

* See chapter 7, section 7.3 for variable naming conventions.

Due to the two-stage format of the cognitive assessment battery, not all children received all items. An IRT model was employed to produce proficiency probability scores, which reflect the probability that a child would have passed a proficiency level. 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. Because the proficiency probabilities were based on overall performance, they could be calculated for all children who had scorable test data, not just those with relatively complete sets of responses to the necessary item clusters.

The proficiency probability scores can be averaged to produce estimates of mastery rates within population subgroups. These continuous measures can provide a closer 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 are making their largest gains in achievement during a particular time interval. Thus, students' school experiences can be related to improvements in specific skills.

Proficiency-level scores differ from proficiency probability scores. Proficiency-level scores are based on the items administered to each child. Since not all children received the same items because of the two-stage assessment design, these scores represent only those children who were administered the items. The use of proficiency-level scores to estimate the total population of children mastering a specific proficiency level is not recommended because stopping rules within the test resulted in missing data for the lower-achieving children. The proficiency probability scores are more suited for estimating the total population of children mastering specific proficiency scores.

See table 3-5 for variable names, descriptions, ranges, weighted means, and standard deviations for the proficiency probability scores.

Familiarity with Conventions of Print

Some items from the child assessment measured children's familiarity with conventions of print but were not part of the set of proficiency scores because they did not fit the hierarchical pattern.

Table 3-5.—Direct cognitive assessment: proficiency probability scores*

Variable	Description	Range of values	Weighted mean	Standard deviation
C1RPROB1	C1 Prob 1 - Letter Recognition	0 - 1	0.64	0.41
C1RPROB2	C1 Prob 2 - Beginning Sounds	0 - 1	0.29	0.33
C1RPROB3	C1 Prob 3 - Ending Sounds	0 - 1	0.16	0.26
C1RPROB4	C1 Prob 4 - Sight Words	0 - 1	0.02	0.13
C1RPROB5	C1 Prob 5 - Word in Context	0 - 1	0.01	0.08
C1MPROB1	C1 Prob 1 - Count, Number, Shape	0 - 1	0.92	0.18
C1MPROB2	C1 Prob 2 - Relative Size	0 - 1	0.53	0.35
C1MPROB3	C1 Prob 3 - Ordinality, Sequence	0 - 1	0.19	0.30
C1MPROB4	C1 Prob 4 - Add/Subtract	0 - 1	0.03	0.12
C1MPROB5	C1 Prob 5 - Multiply/Divide	0 - 1	0.00	0.03
C2RPROB1	C2 Prob 1 - Letter Recognition	0 - 1	0.92	0.22
C2RPROB2	C2 Prob 2 - Beginning Sounds	0 - 1	0.67	0.33
C2RPROB3	C2 Prob 3 - Ending Sounds	0 - 1	0.48	0.35
C2RPROB4	C2 Prob 4 - Sight Words	0 - 1	0.13	0.28
C2RPROB5	C2 Prob 5 - Word in Context	0 - 1	0.04	0.16
C2MPROB1	C2 Prob 1 - Count, Number, Shape	0 - 1	0.99	0.07
C2MPROB2	C2 Prob 2 - Relative Size	0 - 1	0.83	0.25
C2MPROB3	C2 Prob 3 - Ordinality, Sequence	0 - 1	0.52	0.39
C2MPROB4	C2 Prob 4 - Add/Subtract	0 - 1	0.16	0.26
C2MPROB5	C2 Prob 5 - Multiply/Divide	0 - 1	0.02	0.08
C3RPROB1	C3 Prob 1 - Letter Recognition	0 - 1	0.96	0.16
C3RPROB2	C3 Prob 2 - Beginning Sounds	0 - 1	0.81	0.27
C3RPROB3	C3 Prob 3 - Ending Sounds	0 - 1	0.65	0.32
C3RPROB4	C3 Prob 4 - Sight Words	0 - 1	0.24	0.37
C3RPROB5	C3 Prob 5 - Word in Context	0 - 1	0.09	0.25
C3MPROB1	C3 Prob 1 - Count, Number, Shape	0 - 1	0.99	0.04
C3MPROB2	C3 Prob 2 - Relative Size	0 - 1	0.91	0.19
C3MPROB3	C3 Prob 3 - Ordinality, Sequence	0 - 1	0.71	0.36
C3MPROB4	C3 Prob 4 - Add/Subtract	0 - 1	0.32	0.34
C3MPROB5	C3 Prob 5 - Multiply/Divide	0 - 1	0.05	0.16
C4RPROB1	C4 Prob 1 - Letter Recognition	0 - 1	0.99	0.07
C4RPROB2	C4 Prob 2 - Beginning Sounds	0 - 1	0.96	0.14
C4RPROB3	C4 Prob 3 - Ending Sounds	0 - 1	0.91	0.20
C4RPROB4	C4 Prob 4 - Sight Words	0 - 1	0.76	0.37
C4RPROB5	C4 Prob 5 - Word in Context	0 - 1	0.42	0.41
C4MPROB1	C4 Prob 1 - Count, Number, Shape	0 - 1	1.00	0.02
C4MPROB2	C4 Prob 2 - Relative Size	0 - 1	0.98	0.09
C4MPROB3	C4 Prob 3 - Ordinality, Sequence	0 - 1	0.93	0.21
C4MPROB4	C4 Prob 4 - Add/Subtract	0 - 1	0.71	0.34
C4MPROB5	C4 Prob 5 - Multiply/Divide	0 - 1	0.24	0.33

* See chapter 7, section 7.3 for variable naming conventions.

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 were not part of the hierarchical set of proficiency and proficiency probability scores because they did not fit the proficiency scoring pattern. The proficiency levels assume that mastery of a higher level usually means that the child has mastered lower levels. This was not the case with conventions of print. Some children scored high on conventions of print but could not recognize letters, while others had the reverse pattern. Thus, the score for familiarity with conventions of print is reported separately.

See table 3-6 for variable names, descriptions, ranges, weighted means, and standard deviations for the conventions of print scores.

Table 3-6.—Direct cognitive assessment: print familiarity scores[a]

Variable	Description	Range of values	Weighted mean	Standard deviation
C1RPRINT	C1 Print Familiarity	0 - 3[b]	1.8	1.1
C2RPRINT	C2 Print Familiarity	0 - 3[b]	2.3	0.9
C3RPRINT	C3 Print Familiarity	0 - 3[b]	2.6	0.8
C4RPRINT	C4 Print Familiarity	0 - 3[b]	2.8	0.6

[a] See chapter 7, section 7.3 for variable naming conventions.

[b] Ranges for discrete scores (routing number right, proficiency level right/wrong, print familiarity) are reported without a decimal point because they are always integers. Ranges for continuous scores (IRT estimates and probability scores) have decimal points because the scores themselves are not integers.

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.

Item Response Theory-Based Scores

The scores derived from the IRT model (IRT scale scores, T-scores, proficiency probabilities) are 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, are used to estimate a point on an ability continuum, and this ability estimate, theta, then provides the basis for criterion-referenced and norm-referenced scores.

- **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 are used as longitudinal measures of overall growth. Gains made at different points on the scale have qualitatively different interpretations. For example, children who make gains in recognizing letters and letter sounds are learning very different lessons from 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 gains 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 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 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 show 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 are making big gains in basic skills, but most have not yet made major gains in reading words and sentences. 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.

Non-Item Response Theory-Based Scores

The routing test number-right, proficiency level, and conventions of print scores do **not** depend on the assumptions of the IRT model. They are counts of actual number correct for specific sets of test items, rather than estimates based on patterns of overall performance.

- **Routing test number-right scores** for the reading, math, and general knowledge assessments are based on 20, 16, and 12 items 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 test items, which was the same for all children who took the test.

- **Proficiency-level scores** are based on the same sets of items as the proficiency probability scores but are dichotomous, rather than continuous, measures of proficiency. They simply report whether children were able to answer correctly on at least three out of four actual test items in a cluster.

  Users of the proficiency-level scores should be aware of possible bias due to missing data. Stopping and starting rules employed in the administration of the tests to minimize stress on low-performing children result in substantial missing scores. For example, low-performing children did not receive the items associated with the higher proficiencies, and therefore they are missing scores on the higher proficiency levels. And, higher performing children may not have received the items associated with the lower proficiency levels, and therefore are missing scores on the lower proficiency levels. Estimates based on variables with substantial amounts of missing data can be assumed to generalize to the whole sample *only* if "missing-ness" is unrelated to what the variable is measuring. This condition is called "MAR," or Missing-At-Random. The missing are *not* missing-at-random, they were not administered based on performance. Estimates based on proficiency-level scores, without adjustments for missing data, would overstate or understate the population performance.

Users of the proficiency-level scores will need to compensate for the missing data in order to use these scores appropriately. A simple approach would be to impute a "pass" score (=1) for missing data where the child has passed a more difficult level (and has not failed an easier one). For example, a score pattern for the 5 levels of blank-pass-pass-pass-fail could reasonably be imputed to be pass-pass-pass-pass-fail. Similarly, one might assume that a blank level following one or more failed levels could be interpreted as a potential failure at the higher level. A child who had not been administered reading level 5 because of poor performance on easier tasks might have a pattern of pass-fail-fail-fail-blank. Interpreting the missing score as a probable failure would not be unreasonable. Some score patterns do not have a clear indication of pass or fail (e.g., pass-pass-blank-fail-fail, where either a pass or a fail for level 3 would produce a consistent pattern; or a "reversal" pattern such as fail-pass-fail-blank-blank). These situations are a relatively small proportion of the missing scores. They may best be imputed by reference to the corresponding proficiency probability score, and imputing a "pass" if the probability is high, or a "fail" if it is low. This may be done either by rounding the probability to 0 or 1, or by generating a random 0 or 1, using the proficiency probability score as the probability of generating a 1.

- **Conventions of print scores**, like the proficiency level scores, are based on a count of the number correct for a particular set of items. Users may wish to relate this score to process variables to get a perspective that is somewhat different from that of the hierarchical levels of reading skills.

3.1.6 Reliabilities

Reliability statistics appropriate for each type of score were computed for each subject area, for fall- and spring-kindergarten and fall- and spring-first grade assessments. For the IRT-based scores, the reliability of the overall ability estimate, theta, is based on the variance of repeated estimates of theta. These reliabilities, ranging from 0.88 to 0.97, 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 0.78 to 0.88. Split half reliabilities were computed for the item clusters that made up the dichotomous proficiency-level scores and the conventions of print cluster. These reliabilities were higher for the reading clusters (0.60 to 0.83) than for the math levels (0.26 to 0.67). (As noted earlier, the general knowledge test did not contain hierarchical proficiency levels.) The difference in internal consistency statistics is due to the reading items being essentially replications of the same task, while the math items had a greater diversity of content.

Note that the split half reliabilities for the low-level item clusters decreased from fall- to spring-kindergarten and beyond, while the reliabilities for the clusters at the upper end tended to increase. This is a consequence of changes in the variance of the cluster scores as children progressed in their development of skills. After the initial round of testing, the vast majority of children had mastered the lowest proficiency levels, so the sample variance was low, resulting in lower reliability for subsequent rounds. Conversely, the sample variance for the difficult tasks was very low in the fall-kindergarten, when most children had *not* mastered these skills, and the reliability tended to rise as more and more children attained high-level proficiency in spring-kindergarten and first grade, increasing the total variance. This effect is more pronounced for the math than for the reading clusters for two reasons. First, the math item clusters were more heterogeneous than the reading, in terms of content and difficulty. Second, the reading item clusters were based entirely on items from the routing test, which was taken by all children, while the lowest math cluster employed items from the low level second-stage test as well. By spring-kindergarten, fewer than one-half of the test takers were routed to the low form (and only about seven percent in spring of first grade), and this constrained variance in ability resulted in a lower reliability for children who had this score. Tables 3-7, 3-8, and 3-9 present the reliability statistics for all of the test scores.

Table 3-7.—Reliability of Item Response Theory-based scores*

	IRT-based scores (reliability of theta)		IRT-based scores (reliability of theta)	
Category	Fall-kindergarten	Spring-kindergarten	Fall-first grade	Spring-first grade
Reading	0.93	0.95	0.96	0.97
Mathematics	0.92	0.94	0.94	0.94
General Knowledge	0.88	0.89	0.89	0.89

* Though the majority of base year children were in first grade during the 1999-2000 school year, about five percent of the sampled children were retained in kindergarten or in a second year of kindergarten and a handful of others were in second grade during the 1999-2000 school year.

Table 3-8.—Reliability of routing test number correct (alpha coefficient) *

Category	Routing test number correct (alpha coefficient)		Routing test number correct (alpha coefficient)	
	Fall-kindergarten	Spring-kindergarten	Fall-first grade	Spring-first grade
Reading	0.86	0.88	.88	.86
Mathematics	0.78	0.81	.83	.80
General Knowledge	0.79	0.79	.79	.78

* Though the majority of base year children were in first grade during the 1999-2000 school year, about five percent of the sampled children were retained in kindergarten and a handful of others were in second grade during the 1999-2000 school year.

Table 3-9.—Split half reliability of item-cluster-based scores (proficiency-level scores)*

Category	Fall-kindergarten	Spring-kindergarten	Fall-first grade	Spring-first grade
Reading level 1	0.83	0.79	0.77	0.78
Reading level 2	0.76	0.76	0.73	0.70
Reading level 3	0.72	0.76	0.76	0.68
Reading level 4	0.78	0.77	0.80	0.78
Reading level 5	0.60	0.69	0.73	0.73
Conventions of print	0.70	0.68	0.68	0.60
Mathematics level 1	0.41	0.27	0.26	0.26
Mathematics level 2	0.58	0.49	0.51	0.32
Mathematics level 3	0.63	0.66	0.67	0.59
Mathematics level 4	0.54	0.63	0.66	0.63
Mathematics level 5	0.46	0.53	0.61	0.65

* Though the majority of base year children were in first grade during the 1999-2000 school year, about five percent of the sampled children were retained in kindergarten and a handful of others were in second grade during the 1999-2000 school year.

3.2 Indirect Cognitive Assessment

The academic rating scale (ARS) was developed for the ECLS-K to measure teachers' evaluations of students' academic achievement in the three domains that are also directly assessed in the cognitive battery: language and literacy (reading), general knowledge (science and social studies), and mathematical thinking. Teachers rated the child's skills, knowledge, and behaviors on a scale from "Not

Yet" to "Proficient" (see table 3-10). If a skill, knowledge, or behavior had not been introduced into the classroom yet, the teacher coded that item as N/A (not applicable). The difference 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 2.4.1.

Table 3-10.—Academic rating scale response scale

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.

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 measures only the products of children's achievement. 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 math problems, or investigate a scientific phenomenon.

Another difference between the ARS and direct cognitive assessment is that the skills, knowledge, and behaviors on the ARS reflect a broader sampling of the most recent national curriculum standards and guidelines from early childhood professionals and researchers. The ARS items were not limited by the constraints of a standardized testing format as were the direct cognitive items. Therefore, the scope of curricular content represented in the indirect measures is broader than the content represented on the direct assessment battery. These criterion-referenced indirect measures 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.

3.3 Rasch Scores Available for the Academic Rating Scale

IRT was used to calculate scores for the ARS in order to compare performance of students on a hierarchy of skills, knowledge, and behavior. The Rasch Rating Scale model uses the pattern of ratings on the items to determine an estimate of the difficulty of each item and to place each student on a continuous ability scale (in this case 1-5). Rasch provides interval-level measurement. The Rasch analysis showed that the reliability of the estimates of child ability was very high for all domains (see table 3-11).

Table 3-11.—Person reliability for the Rasch-based score

Category	Spring-kindergarten	Spring-first grade
ARS Language and Literacy	.91	.94
ARS Mathematical Thinking	.94	.94
ARS General Knowledge	.95	.95

As mentioned, the ARS scores were scaled to have a low of one and a high of five to correspond to the five-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. Students who received maximum ratings on all the items or minimum ratings on all the items are assigned an estimated score using Bayesian techniques.

The variable names, descriptions, value ranges, weighted means, and standard deviations for the first grade (T4) ARS scores are shown in table 3-12. The description for each variable in the tables begins with a “T,” indicating that it is a teacher questionnaire child-level variable. The round of data collection is indicated by a “1” for fall-kindergarten, a “2” for spring-kindergarten, and a “4” for spring-first grade. As noted earlier, an error was identified in the base year ARS scores. Specifically, the fall and spring base year ARS scores used slightly different metrics and, therefore, were not directly comparable.

These scores have been recomputed to put them on the same metric. The corrected base year ARS scores are included on the first grade data files.

On the ARS teachers indicate "not applicable" when the knowledge, skill, or behavior has not been introduced to the classroom. Because some children might already have 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 is set to missing.

Table 3-12.—First grade academic rating scale: variable names, descriptions, ranges, weighted means, and standard deviations[1]

Variable name	Description[2]	Range of values	Weighted mean	Standard deviation
T1RARSLI	T1 REC Language and Literacy ARS Score	1 - 5	2.48	0.73
T1RARSMA	T1 REC Mathematical Thinking ARS Score	1 - 5	2.54	0.82
T1RARSGE	T1 REC General Knowledge ARS Score	1 - 5	2.62	0.98
T2RARSLI	T2 REC Literacy ARS Score	1 - 5	3.33	0.81
T2RARSMA	T2 REC Math ARS Score	1 - 5	3.50	0.86
T2RARSGE	T2 REC General Knowledge ARS Score	1 - 5	3.55	0.99
T4ARSLIT	T4 Literacy ARS Score	1 - 5	3.40	0.93
T4ARSMAT	T4 Math ARS Score	1 - 5	3.43	0.90
T4ARSGEN	T4 General Knowledge ARS Score	1 - 5	3.26	0.99

[1] See chapter 7, section 7.3 for variable naming conventions.

[2] In the table, the letters "REC" in the variable description indicate that the variable was recalibrated.

Tables 3-13 to 3-18 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 both kindergarten and first grade items.

Table 3-13.—Kindergarten academic rating scale language and literacy item difficulties (arranged in order of difficulty)

Item difficulty	Item number and abbreviated content
2.44	Q3. Easily and quickly names all upper- and lower- case letters of the alphabet
2.76	Q4. Produces rhyming words
2.79	Q2. Understands and interprets a story or other text read to him/her
2.83	Q1. Uses complex sentence structures
3.20	Q8. Demonstrates an understanding of some of the conventions of print
3.21	Q5. Reads simple books independently
3.30	Q6. Uses different strategies to read unfamiliar words
3.45	Q7. Composes simple stories

Table 3-14.—Kindergarten academic rating scale mathematical thinking item difficulties (arranged in order of difficulty)

Item difficulty	Item number and abbreviated content
2.72	Sorts, classifies, and compares math materials by rules and attributes
2.79	Demonstrates an understanding of graphing activities
2.82	Orders a group of objects
2.90	Shows an understanding of the relationship between quantities
3.08	Solves problems involving numbers using concrete objects
3.22	Uses a variety of strategies to solve math problems
3.45	Uses instruments accurately for measuring

Table 3-15.—Kindergarten academic rating scale general knowledge item difficulties (arranged in order of difficulty)

Item difficulty	Item number and abbreviated content
2.78	Classifies and compares living and non-living things in different ways
2.95	Uses his/her senses to explore and observe
2.99	Recognizes distinct differences in group habits and living patterns
3.11	Recognizes some ways in which people rely on each other for goods and services
3.12	Forms explanations based on observations and explorations

Table 3-16.—Spring-first grade academic rating scale language and literacy item difficulties (arranged in order of difficulty)

Item difficulty	Item number and abbreviated content
2.73	Q3. Reads words with regular vowel sounds
2.85	Q1. Contributes relevant information to classroom discussions
2.85	Q5. Reads first grade books independently with comprehension
2.89	Q2. Understands and interprets a story or other text read to him/her
2.97	Q6. Reads first grade books fluently
3.14	Q8. Demonstrates an understanding of some of the conventions of print
3.17	Q4. Reads words with irregular vowel sounds,
3.30	Q7. Composes a story with a clear beginning, middle, and end

Table 3-17.—Spring-first grade academic rating scale mathematical thinking item difficulties (arranged in order of difficulty)

Item difficulty	Item number and abbreviated content
2.39	Models, reads, writes, and compares whole numbers
2.61	Demonstrates an understanding of place value
2.70	Surveys, collects, and organizes data into simple graphs
2.73	Makes reasonable estimates of quantities
2.81	Counts change with two different types of coins
2.90	Measures to the nearest whole number using common instruments
2.90	Uses a variety of strategies to solve math problems

Table 3-18.—Spring-first grade academic rating scale general knowledge item difficulties (arranged in order of difficulty)

Item difficulty	Item number and abbreviated content
2.82	Classifies and compares living and non-living things in different ways
2.88	Identifies similarities and differences in group habits and living patterns
2.96	Shows a beginning understanding that maps represent actual places
3.02	Forms explanations and conclusions based on observation and investigation
3.02	Recognizes some ways in which people rely on each other for goods and services
3.02	Makes logical predictions

The scale was designed to provide information on children's abilities at a given point in time, not necessarily over time. Even though a common calibration of the fall and spring scores was used for the kindergarten ratings, it is not recommended that users compute a change score. In addition, although the item stems are similar across kindergarten and first grade, the actual items include performance criteria that increase in difficulty from one grade to the next. Moreover, the kindergarten and first grade ARS scores are placed on different metrics. Therefore, change scores should not be used between kindergarten and first grade. Covariance models may be used to compare teacher's ratings of performance in kindergarten and first grade. 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 3-19 to 3-24 provide standard errors (SE) for each of the Rasch scores for kindergarten and first 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 3-19.—Kindergarten academic rating scale language and literacy standard errors

Score	Measure	S.E.	Score	Measure	S.E.	Score	Measure	S.E.
8	1.00E	.60	19	2.63	.16	30	3.43	.16
9	1.42	.34	20	2.70	.15	31	3.52	.16
10	1.69	.26	21	2.78	.15	32	3.60	.17
11	1.86	.22	22	2.85	.15	33	3.69	.17
12	2.00	.20	23	2.92	.15	34	3.78	.18
13	2.11	.19	24	2.99	.15	35	3.89	.19
14	2.22	.18	25	3.06	.15	36	4.00	.20
15	2.31	.17	26	3.13	.15	37	4.14	.22
16	2.40	.16	27	3.21	.15	38	4.32	.26
17	2.48	.16	28	3.28	.16	39	4.58	.34
18	2.56	.16	29	3.36	.16	40	5.00E	.60

Table 3-20.—Kindergarten academic rating scale mathematical thinking standard errors

Score	Measure	S.E.	Score	Measure	S.E.	Score	Measure	S.E.
7	1.00E	.46	17	2.56	.16	27	3.66	.16
8	1.33	.27	18	2.67	.16	28	3.77	.16
9	1.55	.21	19	2.79	.17	29	3.88	.16
10	1.71	.19	20	2.90	.16	30	3.99	.17
11	1.86	.18	21	3.01	.16	31	4.11	.18
12	1.99	.18	22	3.12	.16	32	4.25	.19
13	2.11	.17	23	3.23	.16	33	4.42	.22
14	2.23	.17	24	3.33	.16	34	4.66	.28
15	2.34	.16	25	3.44	.16	35	5.00E	.46
16	2.45	.16	26	3.55	.16			

Table 3-21.—Kindergarten academic rating scale general knowledge standard errors

Score	Measure	S.E.	Score	Measure	S.E.	Score	Measure	S.E.
5	1.00E	.38	12	2.41	.18	19	3.73	.20
6	1.28	.23	13	2.57	.18	20	3.95	.22
7	1.49	.19	14	2.75	.20	21	4.17	.20
8	1.66	.18	15	2.98	.22	22	4.35	.18
9	1.83	.19	16	3.21	.20	23	4.52	.19
10	2.03	.21	17	3.39	.18	24	4.72	.23
11	2.24	.19	18	3.55	.18	25	5.00E	.37

Table 3-22.—Spring-first grade academic rating scale language and literacy standard errors

Score	Measure	S.E.	Score	Measure	S.E.	Score	Measure	S.E.
8	1.02E	.50	19	2.51	.15	30	3.53	.16
9	1.37	.29	20	2.60	.15	31	3.63	.16
10	1.60	.22	21	2.69	.16	32	3.72	.16
11	1.75	.19	22	2.78	.16	33	3.82	.16
12	1.87	.17	23	2.87	.16	34	3.91	.16
13	1.97	.16	24	2.97	.16	35	4.02	.17
14	2.07	.16	25	3.06	.16	36	4.13	.18
15	2.16	.15	26	3.16	.16	37	4.26	.19
16	2.25	.15	27	3.25	.16	38	4.41	.22
17	2.34	.15	28	3.35	.16	39	4.65	.29
18	2.42	.15	29	3.44	.16	40	5.00E	.50

Table 3-23.—Spring-first grade academic rating scale mathematical thinking standard errors

Score	Measure	S.E.	Score	Measure	S.E.	Score	Measure	S.E.
7	1.00E	.46	17	2.54	.16	27	3.68	.17
8	1.33	.27	18	2.65	.16	28	3.80	.17
9	1.55	.21	19	2.76	.17	29	3.92	.17
10	1.71	.18	20	2.87	.17	30	4.05	.17
11	1.84	.17	21	2.99	.17	31	4.17	.18
12	1.96	.17	22	3.11	.17	32	4.30	.18
13	2.08	.17	23	3.23	.17	33	4.46	.21
14	2.19	.17	24	3.34	.17	34	4.68	.27
15	2.31	.17	25	3.45	.17	35	5.00E	.46
16	2.42	.17	26	3.57	.17			

Table 3-24.—Spring-first grade academic rating scale general knowledge standard errors

Score	Measure	S.E.	Score	Measure	S.E.	Score	Measure	S.E.
6	1.000E	.394	15	2.446	.163	24	3.867	.214
7	1.282	.231	16	2.577	.169	25	4.071	.196
8	1.476	.182	17	2.725	.185	26	4.234	.176
9	1.619	.167	18	2.902	.197	27	4.376	.171
10	1.751	.167	19	3.080	.186	28	4.523	.183
11	1.888	.173	20	3.232	.172	29	4.718	.231
12	2.036	.178	21	3.367	.167	30	5.001E	.394
13	2.183	.173	22	3.504	.174			
14	2.319	.165	23	3.664	.194			

The majority of teachers rated more than one student on the ARS. The number of students rated by each teacher ranged from one to 23.

3.4 Oral Language Development Scale

The language assessment scores (OLDS scores) for language minority children are located with the other child scores on the file. There are a total of 16 OLDS scores—four English and four Spanish for each of the two rounds (i.e., fall-first grade and spring-first grade). Children in households speaking languages other than English, who had not been administered the English ECLS-K cognitive assessments in the prior rounds of the study, were first given the English OLDS. Of that group, those scoring below the cut point of the English OLDS were administered the Spanish OLDS if the child's home language was noted as Spanish by the school. (See chapter 2, section 2.2 for more detail on the content of the OLDS items.)

The variable names, descriptions, value ranges, weighted means, and standard deviations for the OLDS are shown in table 3-25. The 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 3 (fall-first grade) or 4 (spring-first grade).

Table 3-25.—Oral language development scale: variable names, descriptions, ranges, means, and standard deviations*

Variable	Description	Range of values	Weighted mean	Standard deviation
C3SCTOT	C3 AIQ400 Child's Total OLDS Score	0 - 60	36.8	17.6
C3SCORD	C3 AIQ400 Simon Says Child Score	0 - 10	8.3	2.6
C3SCART	C3 AIQ400 Art Show Child Score	0 - 10	7.2	3.0
C3SCSTO	C3 AIQ400 Tell Stories Child Score	0 - 40	21.3	13.2
C3SSCTOT	C3 SAIQ400 Spanish Total OLDS Scores	0 - 42	24.5	6.7
C3SSCORD	C3 SAIQ400 Spanish Simon Says Child Score	0 - 10	9.8	0.7
C3SSCART	C3 SAIQ400 Spanish Art Show Child Score	0 - 10	8.5	2.2
C3SSCSTO	C3 SAIQ400 Spanish Tell Stories Score	0 - 24	6.2	6.0
C4SCTOT	C4 AIQ400 Child's Total OLDS Score	0 - 60	38.7	15.6
C4SCORD	C4 AIQ400 Simon Says Child Score	0 - 10	8.8	2.2
C4SCART	C4 AIQ400 Art Show Child Score	0 - 10	7.6	2.7
C4SCSTO	C4 AIQ400 Tell Stories Child Score	0 - 40	22.3	11.9
C4SSCTOT	C4 SAIQ400 Spanish Total OLDS Scores	0 - 40	23.9	6.7
C4SSCORD	C4 SAIQ400 Spanish Simon Says Child Score	0 - 10	9.7	1.5
C4SSCART	C4 SAIQ400 Spanish Art Show Child Score	0 - 10	8.7	1.9
C4SSCSTO	C4 SAIQ400 Spanish Tell Stories Score	0 - 20	5.5	5.7

* See chapter 7, section 7.3 for variable naming conventions.

3.5 Social Rating Scale

The SRS asked both teachers and parents to tell how often a student exhibited certain social skills and behaviors. Teachers and parents used a frequency scale (see table 3-26) to report on how often the student demonstrated the behavior described. See chapter 2, section 2.3 and 2.4 for additional information on the parent and teacher SRS instruments. The scale scores on all SRS scales are 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 same items were administered in fall and spring-kindergarten, and in spring-first grade, so change scores may be computed by subtracting time 1 (fall- kindergarten) from time two (spring-kindergarten) scores, etc. The reliability for the teacher SRS scales is high (see table 3-27). The reliability is lower for the parent scales (see table 3-28).

Table 3-26.—Social rating scale response scale

Answer		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.

Table 3-27.—Split half reliability for the teacher social rating scale scores

Category	Fall-kindergarten	Spring-kindergarten	Spring-first grade
Approaches to Learning	0.89	0.89	.89
Self-control	0.79	0.80	.80
Interpersonal	0.89	0.89	.89
Externalizing Problem Behaviors	0.90	0.90	.86
Internalizing Problem Behaviors	0.80	0.78	.77

Table 3-28.—Split half reliability for the parent social rating scale scores

Category	Fall-kindergarten	Spring-kindergarten	Spring-first grade
Approaches to Learning	0.68	0.69	.69
Self-Control	0.74	0.75	.75
Social Interaction	0.70	0.68	.69
Impulsive/Overactive	0.46	0.47	.48
Sad/Lonely	0.60	0.61	.63

3.5.1 Teacher Social Rating Scale

Teachers rated individual students as part of a self-administered questionnaire. The five social skill teacher scales are as follows: approaches to learning, self-control, interpersonal skills, externalizing problem behaviors, and internalizing problem behaviors.

Variable names for the teacher scores, descriptions, ranges, weighted means, and standard deviations for these scales are shown in table 3-29. Numbers in the table for round 4 are for first graders, with kindergarten repeaters' scores shown in parentheses. The description for each variable in the tables begin with a "T," indicating that it is a teacher variable, and a data collection round number, either 1 (fall-kindergarten), 2 (spring-kindergarten), or 4 (spring-first grade).

Care should be taken when entering these scales into the same analysis due to problems of multicollinearity. The intercorrelations among the five SRS factors are high. The factor intercorrelations with the internalizing problem behaviors are the lowest. The correlations between the teacher SRS factors range from 0.25 to 0.78 in fall-kindergarten and from 0.30 to 0.80 in spring-kindergarten (absolute values). In round 4, the correlations ranged from .31 to .81 for first graders, and .29 to .80 for kindergarten repeaters.

Table 3-29.—Teacher social rating scores: variable names, descriptions, ranges, weighted means, and standard deviations*

Variable	Description	Range of values	Weighted mean	Standard deviation
T1LEARN	T1 Approaches to Learning	1 - 4	3.0	0.7
T1CONTRO	T1 Self-Control	1 - 4	3.1	0.6
T1INTERP	T1 Interpersonal	1 - 4	3.0	0.6
T1EXTERN	T1 Externalizing Problem Behaviors	1 - 4	1.6	0.6
T1INTERN	T1 Internalizing Problem Behaviors	1 - 4	1.6	0.5
T2LEARN	T2 Approaches to Learning	1 - 4	3.1	0.7
T2CONTRO	T2 Self-Control	1 - 4	3.2	0.6
T2INTERP	T2 Interpersonal	1 - 4	3.1	0.6
T2EXTERN	T2 Externalizing Problem Behaviors	1 - 4	1.7	0.7
T2INTERN	T2 Internalizing Problem Behaviors	1 - 4	1.6	0.5
T4LEARN	T4 Approaches to Learning	1 - 4	3.0 (2.9)	0.7 (0.7)
T4CONTRO	T4 Self-Control	1 - 4	3.2 (3.0)	0.6 (0.6)
T4INTERP	T4 Interpersonal	1 - 4	3.1 (3.0)	0.6 (0.7)
T4EXTERN	T4 Externalizing Problem Behaviors	1 - 4	1.7 (1.8)	0.6 (0.7)
T4INTERN	T4 Internalizing Problem Behaviors	1 - 4	1.6 (1.7)	0.5 (0.6)

* See chapter 7, section 7.3 for variable naming conventions.

3.5.2 Parent Social Rating Scale

The items on the parent SRS were administered as part of a longer telephone or in-person survey. (See chapter 2, section 2.3 for a description of the parent scales.) The factors on the parent SRS are similar to the teacher SRS; however, the items in the parent SRS are geared to the home environment and thus are not the same items. It is also important to keep in mind that parents and teachers observe the children in very different environments. The five social skill parent scales are as follows: approaches to learning, self-control, social interaction, impulsive/overactive, and sad/lonely. The correlations between the parent SRS factors were not as high as the teacher SRS factors. They ranged from 0.05 to 0.45 in fall-kindergarten, from 0.08 to 0.45 in spring-kindergarten, and from .11 to .45 in spring-first grade (absolute values).

Variable names for the parent scores, descriptions, ranges, weighted means, and standard deviations for these scales are shown in table 3-30. The description for each variable in the tables begin with a "P," indicating that it is a parent variable, and a data collection round number, either 1 (fall-kindergarten), 2 (spring-kindergarten), or 4 (spring-first grade).

Table 3-30.—Parent social rating scores: variable names, descriptions, ranges, weighted means, and standard deviations*

Variable	Description	Range of values	Weighted mean	Standard deviation
P1LEARN	P1 Approaches to Learning	1 - 4	3.1	0.5
P1CONTRO	P1 Self-Control	1 - 4	2.8	0.5
P1SOCIAL	P1 Social Interaction	1 - 4	3.3	0.6
P1SADLON	P1 Sad/Lonely	1 - 4	1.5	0.4
P1IMPULS	P1 Impulsive/Overactive	1 - 4	2.0	0.7
P2LEARN	P2 Approaches to Learning	1 - 4	3.1	0.5
P2CONTRO	P2 Self-Control	1 - 4	2.9	0.5
P2SOCIAL	P2 Social Interaction	1 - 4	3.4	0.5
P2SADLON	P2 Sad/Lonely	1 - 4	1.6	0.4
P2IMPULS	P2 Impulsive/Overactive	1 - 4	2.0	0.7
P4LEARN	P4 Approaches to Learning	1 - 4	3.1	0.5
P4CONTRO	P4 Self-Control	1 - 4	3.0	0.5
P4SOCIAL	P4 Social Interaction	1 - 4	3.4	0.5
P4SADLON	P4 Sad/Lonely	1 - 4	1.5	0.4
P4IMPULS	P4 Impulsive/Overactive	1 - 4	1.9	0.7

* See chapter 7, section 7.3 for variable naming conventions.

4. SAMPLE DESIGN AND IMPLEMENTATION

The Early Childhood Longitudinal Study-Kindergarten Class of 1998-99 (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 is considered responding if there was a completed child assessment or parent interview in fall- or spring-kindergarten. While a full-scale data collection was mounted in spring-first grade, the effort for 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.

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/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 measure of size. An initial cross-classification of census region with metropolitan statistical area (MSA) status created eight superstrata. These were further subdivided by percent minority, size, and 1988 per capita income. From each non-SR stratum, two PSUs were selected PPS without replacement using Durbin's Method.[1]

[1] Durbin, J. (1967). Design of multi-stage surveys for the estimation of sampling errors. *Journal of the Royal Statistical Society* C, 16, 152-164.

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

Table 4-1.—Distribution of the ECLS-K primary sampling unit sample by self-representing status, metropolitan statistical area status, and census region

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

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[2] (CCD) and the 1995-96 Private School Universe Survey[3] (PSS). 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. A procedure was implemented to identify kindergarten programs that would be operational at the time of ECLS-K's Base Year data collection, but that were not included in the frame just described. These were newly opened schools that were not listed in CCD and PSS, and schools that were in CCD and 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.

---

[2] U.S. Department of Education, National Center for Education Statistics, Common Core of Data, Public School Universe Survey, 1995-96.

[3] U.S. Department of Education, National Center for Education Statistics. *Private School Universe Survey, 1995-96*, NCES 98-229, by Stephen P. Broughman and Lenore A. Colaciello. Washington, DC: 1998.

The characteristics of the ECLS-K school sample are presented in table 4-2. Schools that were discovered to be ineligible during recruitment have been omitted from the tabulation.

Table 4-2.—Characteristics of the ECLS-K base year school sample

School type	Census region				
	Northeast	Midwest	South	West	Total
Public	161	210	306	237	914
Private*	82	88	112	81	363
Total	243	298	418	318	1277

*120 Catholic, 149 other religious, 94 nonsectarian

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 which 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. 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.

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.

4.2 Fall-First Grade Subsample

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

Base year schools in the 62 fall-first grade sampled PSUs were stratified by frame source (original public, original private, freshened, etc.) 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 4-3.

Table 4-3.—Characteristics of base year cooperating schools selected for fall-first grade

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

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. By eligible we mean a base year respondent (i.e., a child who had either a fall- or spring-kindergarten child assessment or parent interview). Base year nonrepondents would be adjusted for during weighting. Because of the additional burden of school recruiting, the cost of collecting data for a child who transferred from the school in which he or she was originally sampled greatly exceeds that for a child who stayed enrolled. To contain these costs, 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.

Prior to subsampling with equal probability, children were stratified into groups of nonmovers, movers with information identifying their new schools, and movers without such identifying information. A flag was created for each child indicating whether the child had been sampled to be followed.

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

Table 4-4.—Characteristics of children subsampled for fall-first grade*

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

* School characteristics are of the original sampled schools.

Table 4-4.—Characteristics of children subsampled for fall-first grade* (continued)

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

* School characteristics are of the original sampled schools.

4.3 Spring-First Grade Sample

As mentioned in the introduction to this chapter, the ECLS-K spring-first grade data collection targeted all base year respondents. 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.

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 follow-up, the three primary strata were SR PSUs, NSR PSUs that had been selected for fall-first grade, and NSR PSUs that had not been selected for fall-first grade. Within these major strata, schools were grouped by frame source (original public, original private, new from Catholic dioceses, new from local governments, etc.). Finally within each frame source, schools were stratified by response status, and 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 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 4-5 shows the characteristics of the children in the spring-first grade sample, excluding freshened students. Region, locale, religious affiliation, and type of school describe the school at which the child attended kindergarten.

Table 4-5.—Characteristics of children in spring-first grade sample, excluding freshened students *

Characteristic	Public	Private	Total
Total	14,248	3,836	18,084
Region			
Northeast	2,434	905	3,339
Midwest	3,474	1,104	4,578
South	5,029	1,021	6,050
West	3,311	806	4,117
Type of locale			
Large city	2,575	884	3,459
Midsize city	2,797	964	3,761
Urban fringe of large city	3,991	1,149	5,140
Urban fringe of midsize city	1,126	162	1,288
Large town	466	110	576
Small town	1,215	363	1,578
Rural	2,078	204	2,282
Religious affiliation			
Catholic	-	2,091	2,091
Other religious	-	1,139	1,139
Nonreligious, private	-	606	606
Type of school			
Regular	13,971	3,306	17,277
Ungraded	24	16	40
No grade beyond K	235	185	420
Unknown	18	329	347
Composite child race			
White	7,472	2,736	10,208
Black	2,289	308	2,597
Hispanic, with race	1,220	240	1,460
Hispanic, without race	1,456	192	1,648
Asian	939	210	1,149
Pacific Islander	186	16	202
Native American	294	38	332
Multirace	347	87	434
Unknown	45	9	54

* School characteristics are of the original sampled schools.

Table 4-5.—Characteristics of children in spring-first grade sample, excluding freshened students (continued)

Characteristic	Public	Private	Total
Highest parent level of education			
Less than high school	1,491	38	1,529
High school graduate	3,356	423	3,779
Vocational/technical	926	152	1,078
Some college	3,313	898	4,211
College graduate	2,194	1,154	3,348
Masters	719	472	1,191
Ph.D./professional	395	354	749
Unknown	1,854	345	2,199

* School characteristics are of the original sampled schools.

4.3.2 Student Freshening

The spring-first grade student freshening used a half-open interval sampling procedure.[4] 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 alphabetic 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 was linked to the base year sampled student (i.e., was assigned that student's probability of selection); (2) the record search procedure was repeated for the next listed child, etc. When the record search revealed that a child had been enrolled in kindergarten the previous year, (1) that child was not considered part of the freshened sample and (2) the procedure was begun all over again with the second base year sampled student name etc. 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 first graders in the United States by about 2.6 percent.

[4] Kish, L. (1965). *Survey Sampling*. John Wiley & Sons, New York., p 56

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 fall 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.

4.4 Calculation and Use of Sample Weights

As in the base year, 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 second year of the ECLS-K, which covers fall 1999 (fall-first grade) and spring 2000 (spring-first grade), only child-level weights were computed. The use of these weights is essential to produce estimates that are representative of the population of first grade children in the school year 1999-2000 and of kindergartners in 1998-99 one year later. The sample of teachers in the first grade year only represents teachers of ECLS-K children eligible for the first grade survey. The sample of teachers is not representative of the population of first grade teachers in the country. The same applies to schools. Therefore, teacher- and school-level weights are not provided.

Several sets of weights were computed for each of the two rounds of data collection (fall- and spring-first grade). First grade longitudinal weights were also computed for children with complete data from both rounds (fall and spring) of first grade. As in the base year, there are several survey instruments administered to sampled children and their parents, teachers and schools: cognitive and physical assessments for children in both fall- and spring-first grade; parent instruments in both fall- and spring-first grade; several types of teacher instruments in spring-first grade only; and school instruments, also in spring-first grade only. The stages of base year sampling in conjunction with the nonresponse at each stage and the diversity of survey instruments require that multiple sampling weights be computed for use in analyzing the ECLS-K data. Kindergarten through first grade longitudinal weights for the analysis of both kindergarten and first grade data were also created. Details on these longitudinal weights are available in chapter 9, section 9.4. This section describes the cross-sectional first grade weights and the within first grade longitudinal weights.

This section describes the different types of sample weights computed for the ECLS-K, how they were calculated, how they should be used, and their statistical characteristics.

4.4.1 Types of Sample Weights

Several sets of weights were computed for each round of data collection in the first grade year and for children with complete data from both rounds. 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.

Tables 4-6 and 4-7 summarize the different types of cross-sectional weights and how they should be used. Cross-sectional weights provide an accurate estimate for the specific round of data collection. Table 4-6 describes weights for fall-first grade estimates, and table 4-7 describes weights for spring-first grade estimates. Table 4-8 describes weights for first grade longitudinal estimates.

Table 4-6.—ECLS-K fall-first grade cross-sectional weights

Weight	to be used for analysis of ...
C3CW0	fall-first grade direct child assessment data alone or in conjunction with a) a limited set of child characteristics (e.g. age, sex, race-ethnicity), and b) school or teacher data collected in spring-first grade.
C3PW0	fall-first grade parent interview data alone or in combination with a) fall-first grade child assessment data, and b) school or teacher data collected in spring-first grade.

Table 4-7.—ECLS-K spring-first grade cross-sectional weights

Weight	to be used for analysis of ...
C4CW0	spring-first 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) any spring-first grade teacher questionnaire A, B or C data, and c) data from the school administrator questionnaire or facilities checklist
C4PW0	spring-first grade parent interview data alone or in combination with a) spring-first grade child assessment data, b) spring-first grade teacher questionnaire A, B, or C data, and c) data from the school administrator questionnaire or facilities checklist. *Exception*: If data from the parent AND child assessment AND teacher questionnaire A or B (not C) are used then C4CPTW0 should be used
C4CPTW0	spring-first grade direct child assessment data combined with spring-first grade parent interview data AND spring-first grade teacher data alone or in conjunction with data from the school administrator or facilities checklist

These tables are designed to help users choose appropriate weights for their analysis. Answers to the following questions can help in the selection of the correct weight.

1. Is the analysis concerned with one point in time or two? If the analysis pertains to a) fall-first grade (single point in time) then table 4-6 guides the selection of weights, b) spring-first grade (single point in time) then go to table 4-7, and c) both fall- and spring-first grade (two points in time) then go to table 4-8.

2. What instruments do the data to be used in the analysis come from? There are several options when deciding on which weights to use, and the source of the data affect which weight to use. In each of the tables, details under "to be used in the analysis of . . ." provide guidance based on whether the data were collected through the child assessments, parent interviews or teacher questionnaires.

Weight C3CW0 is used to estimate child-level characteristics or assessment scores for fall-first grade, and C4CW0 is for spring-first grade. Examples of such estimates are the percent of first grade children who are male, the percent of spring-kindergarten children who are API, the percent of fall-first grade children who are seven when they begin first grade, and the mean reading score of children in fall-first grade. 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 or because they were not proficient in English due to a non-English or non-Spanish home language (LM/not Spanish). These children were not administered the ECLS-K direct cognitive battery, but their background characteristics such as age, gender, race-ethnicity, and characteristics of their parents, teachers, classrooms, and schools are available from the parent interviews, the teacher questionnaires, and the school administrator questionnaire. The social rating scores (see chapter 3, section 3.5) from parents and teachers are also available for LM/not Spanish children and children with disabilities, regardless of whether they completed the direct child assessment.

When analyzing spring child assessment data in conjunction with teacher data collected in spring-first grade, weights C4CW0 (for spring-first grade) should be used. An example for the use of C4CW0 is in the analysis of the relationship between children's approaches to learning as rated by their teachers, the teacher's type of teaching certification, and the children's cognitive skills and knowledge. 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. However, different weights should be used for analysis of child data in conjunction with both parent and teacher data (C4CPTW0).

C3PW0 (for fall-first grade) and C4PW0 (for spring-first grade) are 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 are read to at least every day. 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, when exploring the relationship between home literacy behaviors and children's reading skills.

C4CPTW0 (for spring-first grade) is used when child direct assessment *and* teacher *and* parent data 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. These weights should be not be used for estimates using only direct child assessment data or only parent interview data.

Careful consideration should be given to which set of weights is appropriate for the desired analysis. Using the wrong weights will result in more biased or inefficient estimates. For example, if C4CPTW0 were used in an analysis of child and teacher/classroom data only, then the resulting estimates will be inefficient compared to estimates using C4CW0. The lower parent response causes C4CPTW0 to result in lower sample size with positive weights. There may be combinations of data from a different source for which no weights were developed, but most analyses are possible from the weights provided. No child-teacher weights were computed for analyzing child data in conjunction with teacher data because the response rates for the teachers are high; for the analysis of child assessment data in conjunction with teacher data, the child-level weights based solely on the presence of direct child assessment data should be used.

The longitudinal or panel weights (table 4-8) are used for estimates of differences at two points in time. Examples of analysis using longitudinal weights include the following:

- First grade fall-spring difference in mean child assessment scores (C34CW0);
- First grade fall-spring difference in mean social skills as rated by children's parents (C34PW0); and
- The relationship between the gains children make in their reading knowledge and skills, how often their parents read to them, how often their parents take them to the library (C34PW0). This weight is used when the analysis includes data from all four components—fall and spring child assessment, and parent data.

Table 4-8.—ECLS-K first grade longitudinal (panel) weights

Child-level weight	to be used for analysis of ...
C34CW0	child direct assessment data from BOTH fall-first grade and spring-first grade, alone or in conjunction with any combination of a limited set of child characteristics (e.g., age, sex, race-ethnicity)
C34PW0	parent interview data from BOTH fall-first grade and spring-first grade alone or in combination with fall- and/or spring-first grade child assessment data.

While not all combinations of how the data will be used are presented here, the first-grade longitudinal weights described in table 4-8 can be used in the analysis of data coming from different sources and different times. For example, what weight should be used in the analysis of fall- and spring-first grade child direct assessment data in conjunction with spring-kindergarten parent data? The fact that fall- and spring-first grade data are to be analyzed together determines that a first-grade longitudinal weight should be used. When parent data are added, then the type of weight to be used with parent data is appropriate. In the example just cited, C34PW0 is the appropriate weight for this type of analysis. If parent data from fall-first grade are to be analyzed with fall- and spring-first grade assessment data, then C34PW0 is also the weight to be used.

The distribution of schools by number of sampled students with nonzero first grade weights and the mean number of sampled students with nonzero weights per school are useful in analysis using hierarchical linear modeling. These are given in table 4-9. For the first grade year, there are a large number of schools with one to five ECLS-K students. For this reason, schools are classified on the basis of the number of students who did not transfer schools between the base year and the rounds of data collection. For example, for C3CW0, counts in table 4-9 include only fall-first grade students who have not transferred out of their original sample schools; for C34CW0, the counts include only students who are both in fall-first grade and spring-kindergarten who have not transferred out of their original sample schools.

Table 4-9.—Distribution of originally sampled schools by number of children with nonzero weights by first grade sample weights

Sample	Number of cases					Mean cases per school
	1 – 5	6 – 10	11 – 15	16 – 20	21 – 27	
Fall-first grade						
C3CW0	11	28	59	144	53	16
C3PW0	13	34	75	138	35	15
Spring-first grade						
C4CW0	55	70	269	436	129	16
C4PW0	77	128	314	363	92	14
C4CPTW0	80	150	301	318	65	14
Longitudinal						
C34CW0	9	35	85	132	32	15
C34PW0	17	47	114	97	19	14

4.4.2 Weighting Procedures

In fall-first grade, a sample of about one-third of the base year schools was selected, and all base year responding children in the sample schools were included in the study. A base year responding child is 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. In addition to subsampling schools, children who moved out of their original sample schools between fall-kindergarten and fall-first grade were subsampled to be followed into their new schools. The mover subsampling rate was 0.3 for children who moved between fall- and spring-kindergarten, and 0.5 otherwise.

The spring-first grade sample included all base year respondents as defined earlier, and a supplemental sample of first graders brought in through a sample freshening procedure. As in fall-first grade, only a subsample of children who moved from the schools they were attending when they were sampled originally were followed into their new schools. The freshening and mover followup activities targeted a 50 percent subsample of base year schools. To preserve the fall-first grade data series, the 30 percent of schools that constituted the fall-first grade sample were automatically included as part of the spring-first grade 50 percent sample.

Since base year responding children are the basis for all subsequent ECLS-K data collections, the weights that are common to both fall- and spring-first grade were as follows:

- Base year school weight adjusted for base year school nonresponse, and
- Base year child weight that is the product of the base year school nonresponse adjusted weight and the inverse of the within school child selection probability.

In subsequent stages, the base year child weights were then adjusted for:

- Subsampling of schools for the study (fall-first grade only),
- Subsampling of schools for freshening (spring-first grade only),
- School freshening nonresponse (spring-first grade only),
- Subsampling of movers (fall- and spring-first grade),

- Child freshening nonresponse (spring-first grade only), and
- Child nonresponse (fall- and spring-first grade).

The final stage of weighting was to rake the final adjusted weights to sample control totals, except for the first graders brought in through sample freshening (see section 4.4.5 on the computation of weights of children sampled in first grade). The computation of the base year child weights, common to both fall- and spring-first grade is described in section 4.4.3. The subsequent weight adjustments are described separately for fall- and spring-first grade in sections 4.4.4 and 4.4.5. Section 4.4.6 describes the different types of weights computed for fall- and spring-first grade.

In general, in each adjustment to the weight, the adjustment factor is multiplied by the weight in the step before 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.

4.4.3 Computation of Base Year Child Weights

Base Year Nonresponse Adjusted School Weights

The base year nonresponse adjusted school weight was computed as the school base weight adjusted for 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 but for spring-kindergarten. Schools that did not meet this condition are nonrespondents and their weights distributed (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 type of school (public, Catholic private, non-Catholic private, or nonsectarian 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 nonresponse cells were determined, the nonresponse adjustment factors are the reciprocals of the response rates within the selected nonresponse cells.

Base Year Child Weights

Only base year child respondents were fielded in fall- and spring-first grade. A base year child 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 type of school, locale, region, school enrollment class, and child characteristics such as age group, gender and race-ethnicity. These weighting classes are similar to those used for the child weights in the base year. For a description of the computation of child weights in the base year, see chapter 4, section 4.3.4 of the ECLS-K Base Year Public-Use User's Manual (NCES 2001-029), February 2001 or the ECLS-K Restricted-Use Base Year User's Manual (NCES 2000-097), August 2000.

4.4.4 Computation of Fall-First Grade Child Weights

In the first step of the computation of the fall-first grade child weight, an initial weight was created for every child using the nonresponse adjusted base year child weight and incorporating the school subsampling factor appropriate for fall-first grade. The weight was then trimmed to reduce the weight of all the children in one private school that had a large school weight.

Next, the weight was adjusted to reflect the subsampling of movers. A follow flag was created for every child fielded in fall-first grade. If the child moved out of the original sample school and the value of the flag was 1 then the child was followed into his or her new school. If the value of the flag was 0, then the child was not followed and no assessment was attempted. The adjustment factor for subsampling movers 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, 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 had value 1.

For the third category, the adjustment factor was computed within mover cells. Two mover cells were created; the first one included children identified as movers in spring-kindergarten (subsampled at a rate of 0.3), and the second cell included children identified as movers in fall-first grade (subsampled at a rate of 0.5).

After the adjustment for subsampling movers, the child weights were adjusted for nonresponse. 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. 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 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 the type of school the child attended when he or she was originally sampled (public, Catholic, non-Catholic private, and nonsectarian private).

To remove the variability due to the subsampling of schools and movers, the child weights were then raked to sample-based control totals[5] computed using the base year child weights adjusted for nonresponse. A record for every responding eligible child and every ineligible child in the base year is included in this process. In the previous steps, the weights of the nonresponding children have been distributed to the responding children while the weights of the ineligible children have not been 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 fall-first grade data. The reason for including them in the raking step is that these children were eligible in the base year and hence are in the estimates used as 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 fall-kindergarten and/or spring-kindergarten: type of school, region, urbanicity, gender, age, race-ethnicity and socioeconomic status (SES). To reflect the variation of the control totals, each replicate was raked to the corresponding replicated-based control totals. For a discussion of the replicates, see section 4.4.7.

4.4.5 Computation of Spring-First Grade Child Weights

The computation of the spring-first grade child weights was done separately for children sampled in the base year (referred to as children sampled in kindergarten) and children brought in through a sample freshening procedure (referred to as children sampled in first grade). For children sampled in kindergarten, the weighting steps are the same as for fall-first grade, except that the initial child weights did not include the adjustment for school subsample, applicable only to fall-first grade. Children sampled in first grade through freshening were first linked to children sampled in kindergarten in order to create a child base weight for each of them, then the different stages of adjustments applied.

---

[5] These are called sample-based control totals because the numbers used in the numerator of the adjustments are sample estimates subject to sampling errors of roughly the same order as the sampling errors of the estimates from the fall-first grade data. When the numbers used in the numerators are known population totals, the sampling error for estimates such as totals are often substantially reduced. Since the numerators are sample estimates special procedures are needed to reflect this fact in variance estimation.

Weights of Children Sampled in Kindergarten

The spring-first grade initial child weights are the nonresponse adjusted base year child weights described in section 4.4.3. When these weights were examined together with the first grade and kindergarten through first grade longitudinal weights (see chapter 9 for a description of the kindergarten through first grade longitudinal weights), the longitudinal weights that included a fall-first grade component were very large for all children in one particular school. This was a private school with an unusual combination of school and child weights (especially for fall-first grade that only included a subsample of schools) that caused all the children in the school to have large weights. To reduce the effect of the sample from this school on the overall longitudinal estimates and the variances of the estimates, all the weights of the sampled children in the school were trimmed in half; this applies to both cross-sectional and longitudinal weights. No other adjustment was done to compensate for the reduction in weights because of the raking procedure that came later.

Next, the initial child weights were adjusted to reflect the subsampling of movers. As in fall-first grade, follow flags were created for all children in the sample. Children who have moved out of their original sample school were followed in the random 50 percent of schools where the follow flag was set to 1. The adjustment factor for subsampling movers 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, 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 mover cells. Mover cells were created using the type of school of the original sample school and the region where the original sample school is located. Three children with large weights had their weights trimmed. However, the weights were not redistributed because the total sum of weights was re-established in the raking procedure that came later.

After the adjustment for subsampling movers, the child weights were adjusted for nonresponse. As in fall-first 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. 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 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 the type of household collected in the base year from the parent interviews.

To remove the variability due to the subsampling of schools and movers, the child weights were then raked to sample-based control totals computed using the base year child weights adjusted for nonresponse. A record for every responding eligible child and every ineligible child in the base year is included in this process. 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-first grade data. The reason for including them in the raking step is that these children were eligible in the base year and hence are in the estimates used as 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 fall-kindergarten/spring-kindergarten: type of school, region, urbanicity, gender, age, race-ethnicity, and SES. To reflect the variation of the control totals, each replicate was raked to the corresponding replicated-based control totals. For a discussion of the replicates, see section 4.4.7.

Weights of Children Sampled in First Grade

Since each student brought in through sample freshening was linked to a child sampled in kindergarten, the first step of the weighting procedure for children sampled in first grade was to create a weight using 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 linked to the freshened child and further adjusted for nonresponse due to not obtaining the data from the sample of freshened children.

First the base year nonresponse adjusted school weight (as computed in section 4.4.3) 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. The school freshening subsampling adjustment factor was computed as:

- 0 if the school was not in the set of schools subsampled for freshening, 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.

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 schools 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 weights will be used for child level estimates, not school level estimates. The nonresponse cells for this adjustment were created using the same variables used to create the response cells for the base year school nonresponse adjustment, namely school type, region, locale, and school enrollment class.

Next, 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 are the same as those discussed in section 4.4.3, namely, school characteristics such as type of school, locale, region, school enrollment class, and child characteristics such as age group, gender, and race-ethnicity.

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 of nonmovers in first grade or respondents were adjusted for the movers in first grade or nonrespondents. Each child was classified as a (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),
- The sum of the nonresponse adjusted base year child weights (computed in the step before) 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).

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 moved into a new school. The adjustment factor 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 step 1 of all first graders (group a or d) over the sum of the weights adjusted in step 1 of nonmovers in first grade (group d), if the child was a nonmover in first grade (group d).

Only two weighting cells were used for these adjustments: white children and nonwhite children.

The weights thus created for children sampled in kindergarten were then linked to the children that they brought into the sample in first grade through sample freshening. The children sampled in first grade were then separated for a final adjustment that did not concern the children sampled in kindergarten. This final adjustment was for children sampled in first grade who did not cooperate in the assessment. The nonresponse adjustment factor was computed as:

- 0, if the freshened child was a nonrespondent, and
- The sum of the weights adjusted for freshening of all freshened children over the sum of the weights adjusted for freshening of responding freshened children. if the child was a respondent.

This nonresponse adjustment was not done separately by nonresponse cells because of the small number of children brought in through freshening. Two records with large weights were trimmed by half, and the trimmed parts were distributed to the remaining records.

Note that no adjustment was made for unknown eligibility since every freshened child was assumed to be eligible as indicated by the school unless the parents explicitly stated otherwise. Also, the final weights were not raked because there were no control totals to which they could be raked.

4.4.6 Types of Weights

The different types of cross-sectional and longitudinal weights were described in tables 4-6 to 4-8. They were all created as described in sections 4.4.3 to 4.4.5, but the definition of which children were eligible respondents varied.

Weights to Be Used with Direct Child Assessment Data (C3CW0, C4CW0, C34CW0)

In fall-first grade, responding children for this type of weight were eligible children who had fall-first grade scorable direct child cognitive assessment data, or LM/not Spanish children who did not score at or above the OLDS cut score, or children with disabilities who according to specifications in their IEP could not participate in the assessments. A child was eligible if he or she was a base year respondent. Children who transferred to schools and were not flagged to be followed, who moved out of the country or were deceased were considered ineligible. In spring-first grade, responding children were classified using rules similar to those used in fall-first grade. For the longitudinal weights C34CW0, a respondent is defined as a child for whom both C3CW0 and C4CW0 were nonzero.

Table 4-10 shows the number of children who were not assessed due to the following special situations: children who were LM/not Spanish, children with disabilities, children who moved out of their original sample schools and were not flagged to be followed, children who moved and were flagged to be followed but could not be located or moved into a school in a nonsampled county, and children who moved outside of the country or who were deceased. Only the LM/not Spanish and children with disabilities had weights.

Table 4-10.—Number of children who were not assessed due to special situations

Sample	Number of children	
	Unweighted	Weighted
Fall-first grade		
Children with disabilities*	28	23,373
LM/not Spanish children	33	8,583
Moved from original sample schools		
Subsampled, not to be followed	779	465,761
Nonlocatable or moved to nonsampled PSU	121	75,023
To be followed but became ineligible in fall	4	1,992
Spring-first grade		
Children with disabilities	47	14,421
LM/not Spanish children	39	4,789
Moved from original sample schools		
Subsampled, not to be followed	2,850	543,651
Nonlocatable or moved to nonsampled PSU	719	136,056
To be followed but became ineligible in spring	48	7,852

* These children's IEPs specifically prohibited assessments.

Weights to Be Used with Parent Data (C3PW0, C4PW0, C34PW0)

The weights C3PW0 (fall-first grade) and C4PW0 (spring-first grade) are to be used with parent interview data. In both fall- and spring-first 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. For the longitudinal weights C34PW0, a respondent is defined as a child for whom both C3PW0 and C4PW0 are nonzero. Note that these weights are at the child level even though the data were collected from the parents; they sum to all first grade children, not to the parents of first grade children.

Weights to Be Used with a Combination of Child Direct Assessment Data and Parent Interview Data and Teacher Data (C4CPTW0)

The weight C4CPTW0 (spring-first grade) is to be used for analysis involving child, parent, and teacher data. A respondent for this type of weight was defined as a child who had scorable cognitive assessment data for spring-first grade (or LM/not Spanish children or children with disabilities), whose parent completed the FSQ section of the parent interview for spring-first grade, and whose teacher completed part B of the teacher questionnaire. Note that this weight was not computed for fall-first grade because teacher data were not collected in this round.

4.4.7 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. Any adjustments done to the full sample weights were repeated for the replicate weights.

For fall-first grade, there are 40 replicate weights. For spring-first 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 C3CW0 are C3CW1 through C3CW40; the replicate weights for C4CW0 are C4CW1 through C4CW90. 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 4.5.

4.4.8 Characteristics of Sample Weights

The statistical characteristics of the sample weights are presented in table 4-11. For each type of weight, the number of cases with nonzero weights is presented together with the mean weight, the standard deviation, the coefficient variation (i.e., the standard deviation as a percentage of the mean weight), the minimum weight, the maximum weight, the skewness, the kurtosis, and the sum of weights.

Table 4-11.—Characteristics of the first grade child-level weights

Sample	Number of cases	Mean	Standard deviation	CV (× 100)	Minimum	Maximum	Skewness	Kurtosis	Sum
Fall-first grade									
C3CW0	5,291	729.33	554.89	76.08	71.10	6374.63	3.78	19.37	3,858,882
C3PW0	5,071	760.96	484.52	63.67	76.35	5246.83	2.84	11.81	3,858,850
Spring-first grade									
C4CW0	16,727	235.46	207.19	88.00	1.76	3517.71	4.31	32.38	3,938,490
C4PW0	15,626	251.96	203.49	80.76	1.83	3271.78	3.98	28.56	3,937,097
C4CPTW0	13,491	291.74	316.85	108.61	2.21	3849.49	4.35	26.07	3,935,870
First grade longitudinal									
C34CW0	5,047	762.96	571.61	74.92	71.81	6225.66	3.63	18.85	3,850,650
C34PW0	4,682	822.17	526.93	64.09	81.12	5657.06	2.61	10.65	3,849,405

The difference in the estimate of the population of students (sum of weights) between rounds of data collection and types of weight is due a combination of factors, among them: (1) the number of base year respondents who became ineligible (due to death, leaving the country, or being a nonsampled mover) after the base year, and (2) the adjustment of the weights for the children of unknown eligibility. The larger sums of weights in spring-first grade is due to the freshening of the sample that brought in a small number of first graders. Otherwise, the population of inference for all weights is the same.

4.5 Variance Estimation

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

4.5.1 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:

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

where

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

The variance estimates of selected survey items presented in section 4.5 were produced using WesVar and JK2.

Replicate Weights

Replicate weights were created to be used in the calculation of replicate 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.[6] The Durbin method selects two first-stage units per stratum without replacement, with probability proportional to size and a known joint probability of inclusion.

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 and were treated as SR strata for variance

[6] Durbin, J. (1967). Design of Multi-Stage Surveys for the Estimation of Sampling Errors. *Journal of the Royal Statistical Society C*, 16, 152-164.

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, type of school (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 also numbered sequentially from 1 to 63, and so on until the list was exhausted, thus forming the desired 63 variance strata.

In strata with two units, a unit being a PSU in the case of NSR PSUs and a school in the case of SR PSUs, the base weight of the first unit was doubled to form the replicate weight, while the base weight of the second unit was multiplied by zero. In strata with three units, two variance strata were created: in the first variance stratum, the base weight of two of the three units was multiplied by 1.5 to form the replicate weight and the base weight of the last unit was multiplied by zero; in the second variance stratum, the base weight of a different group of two units was multiplied by 1.5, and the base weight of the third unit was multiplied by zero. Any adjustments done 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.

This replicate scheme was used for all of the spring-first grade cross-sectional weights. However, a new feature was added to take into account the freshening process. 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 4.4.5), 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.

For fall-first grade cross-sectional weights and the longitudinal weights that include the fall-first grade component, a set of 40 replicates was created using the paired jackknife method. The smaller number of replicates was due to the fact that only 30 percent of the full sample of schools was included in the fall-first grade subsample. The fall-first grade weights do not account for the Durbin method of sampling PSUs, since it no longer applied. Rather, they reflect the fact that only one of the two sampled PSUs in the NSR strata was kept in the subsample. To account for this feature, pairs of similar NSR PSUs were collapsed into 19 variance strata. The SR PSUs account the remaining 21 variance strata. Replicates were formed following the original scheme for the full sample described earlier.

4.5.2 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. 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 and first grade longitudinal weights. For each type of weights, the sample size was examined, 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. For fall-first grade, the stratum and first-stage unit identifiers reflect the special subsampling design.

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 and STATA. They are described in table 4-12.

Table 4-12.—ECLS-K Taylor Series stratum and first-stage unit identifiers

Variable name	Description
C3TCWSTR	Sampling stratum – fall-first grade C-weights
C3TCWPSU	First-stage sampling unit within stratum – fall-first grade C-weights
C3TPWSTR	Sampling stratum – fall-first grade P-weights
C3TPWPSU	First-stage sampling unit within stratum – fall-first grade P-weights
C4TCWSTR	Sampling stratum – spring-first grade C-weights
C4TCWPSU	First-stage sampling unit within stratum – spring-first grade C-weights
C4TPWSTR	Sampling stratum – spring-first grade P-weights
C4TPWPSU	First-stage sampling unit within stratum – spring-first grade P-weights
C4CPTSTR	Sampling stratum – spring-first grade CPT-weights
C4CPTPSU	First-stage sampling unit within stratum – spring-first grade CPT-weights
C34CSTR	Sampling stratum – fall-first grade/spring-first grade longitudinal C-weights
C34CPSU	First-stage sampling unit within stratum – fall-first grade/spring-first grade longitudinal C-weights
C34PSTR	Sampling stratum – fall-first grade/spring-first grade longitudinal P-weights
C34PPSU	First-stage sampling unit within stratum – fall-first grade/spring-first grade longitudinal P-weights

4.5.3 Specifications for Computing Standard Errors

Specifications for computing standard errors are given in table 4-13. For each type of analysis described in the table, users can choose the replication method or the Taylor Series method for computing standard errors.

Table 4-13.—Specifications for computing standard errors

Type of analysis	Full sample weight	Computing standard errors: Replication method (WesVarPC): ID	Replicate weights	Jackknife method	Taylor Series method (SUDAAN & STATA): Sample design	Nesting variables	Approximating sampling errors: DEFT (Average root design effect)
Fall-first grade cross-sectional							1.947
	C3CW0	CHILDID	C3CW1 – C3CW40	JK2	WR*	C3TCWSTR C3TCWPSU	
	C3PW0	PARENTID	C3PW1 – C3PW40	JK2		C3TPWSTR C3TPWPSU	
Spring-first grade cross-sectional							2.007
	C4CW0	CHILDID	C4CW1 – C4CW90	JK2	WR	C4TCWSTR C4TCWPSU	
	C4PW0	PARENTID	C4PW1 – C4PW90	JK2	WR	C4TPWSTR C4TPWPSU	
	C4CPTW0	CHILDID	C4CPTW1 – C4CPTW90	JK2	WR	C4CPTSTR C4CPTPSU	
First grade longitudinal	C34CW0	CHILDID	C34CW1 – C34CW40	JK2	WR	C34CSTR C34CPSU	1.663
	C34PW0	PARENTID	C34PW1 – C34PW40	JK2	WR	C34PSTR C34PPSU	

* WR = with replacement

For the replication method using WesVar, 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 fall-first grade child level estimates (e.g., mean reading scores) and their standard errors, users need to specify CHILDID in the ID box of the WesVar data file screen, C3CW0 as the full sample weight, C3CW1 to C3CW40 as the replicate weights, and JK2 as the method of replication. Note that for the fall-first grade weights, there are 40 replicate weights instead of 90 replicate weights for all base year and spring-first grade weights, as explained in section 4.4.7.

For the Taylor Series method using either SUDAAN or STATA, the full sample weight, the sample design, the nesting stratum and PSU variables are required. For the same example above, the full sample weight (C3CW0), the with replacement sample design (WR), the stratum variable (C3TCWSTR) and the PSU variable (C3TCWPSU) must be specified.

The last column in table 4-13 gives the average root design effect that can be used to approximate the standard errors for each type of analysis. For a discussion of the use of design effects, see section 4.6.1.

4.6 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.

4.6.1 Use of Design Effects

One method of computing standard errors for the ECLS-K is JK2, as described in section 4.4, using programs designed specifically for analyzing complex survey data such as WesVar. Another approach, Taylor Series linearization (and software designed for it), is also discussed in the same section. If a statistical analysis software package such as SPSS (Statistical Program for the Social Sciences) or SAS (Statistical Analysis System) is used, the standard errors should be corrected using *DEFT*, since these programs calculate standard errors, assuming the data were collected with a simple random sample. The standard error 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.

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

4.6.2 Average 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. One way to produce design effects for analysts' use is to calculate them for a number of variables and average them. The averaging can be done overall and for selected subgroups. The tables that follow show estimates, standard errors, and design effects for selected means and proportions based on the ECLS-K first grade child, parent, teacher, and school data. For each survey item, the tables present 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 estimate, the variable name as it appears in the ECLS-K first grade Electronic Code Book is also provided in the table. If multiple variables were combined to arrive at the estimate, then the names of all the variables used are provided. For example, the estimate of the mean number of days fall-first grade children spent on vacation was computed using two different survey items, P3SUMVAC (parent questionnaire item HEQ100, whether the child went on summer vacation with his or her family) and P3NMDVAC (parent questionnaire item HEQ130, number of days spent on vacation). 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.

Standard errors and design effects for the child-level items are presented in tables 4-14 to 4-18 for fall-first grade sample, spring-first grade sample, and for children in both the fall-first grade and spring-first grade samples. The survey items were selected so that there was a mix of items common to both fall- and spring-first grade and items that were specific to each round of data collection. For fall- and spring-first grade, the items include the different scores from the direct child assessment, the social rating scores as provided by parents and teachers (spring-first grade only), characteristics of the parents, and characteristics of the students as reported by the parents. For a small number of estimates, the data were subset to cases where the estimate is applicable; for example, the number of days spent on vacation is only for children who took summer vacation, the type of primary child care is only for children who had regular scheduled child care; the number of hours that the mothers work is only for women in the labor force. For the first grade student panel, design effects were calculated for some spring-first grade items from the parent data.

Table 4-14 presents standard errors and design effects for the fall-first grade sample, with a median design effect of 3.9. Table 4-15 presents design effect for the spring-first grade sample, with a median design effect of 4.1. For the panel of students common to both fall- and spring-first grade, the median design effect is 2.5, as shown in table 4-16. This lower median design effect is due to the smaller cluster size—or number of children sampled per school—in the panel.

Table 4-17 presents the median design effects for subgroups based on school type, child's gender and race-ethnicity, geographic region, level of urbanicity, and the socioeconomic scale (SES quintiles) of the parents. For fall-first grade, the median design effects vary from 1.3 (children of other race-ethnicity subgroups) to 11.4 (American Indians). For spring-first grade, the median design effect varies from 1.3 (Pacific Islanders) to 4.9 (children in small towns and rural areas). For the fall-first grade/spring-first grade panel, the range of variability of the median design effects is similar to that for fall-first grade. The variation in the design effects is largely a function of the sample size as well as the homogeneity of the children within schools. In fall-first grade, the samples of Pacific Islanders and American Indians are very clustered as reflected in the large design effects both for fall-first grade and for the fall-first grade/spring-kindergarten panel.

In spring-first grade, design effects are not computed for items from the teacher and school administrator's questionnaires since there are no teacher or school weights computed for the first-grade year. Although standard errors 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.

Table 4-14.—ECLS-K, fall-first grade: standard errors and design effects for the full sample

Survey item	Variable name	Number of cases	Estimate	Design SE	SRS SE	DEFT	DEFF
Child scores (mean)							
Reading score	C3RRSCAL	5,053	38.01	0.410	0.178	2.300	5.291
Math score	C3RMSCAL	5,226	32.41	0.280	0.133	2.109	4.448
General knowledge score	C3RGSCAL	5,044	30.05	0.288	0.111	2.587	6.694
Child characteristics (percent)							
Expected to graduate from college	P3EXPECT	5,036	60.06	1.443	0.690	2.091	4.371
Received summer booklist	P3BKLIST	4,981	42.10	1.686	0.700	2.410	5.809
Attended summer school	P3SUMSCH	5,046	10.89	1.204	0.439	2.745	7.534
Took summer vacation	P3SUMVAC	5,045	74.91	1.285	0.610	2.105	4.432
Read book to child everyday	P3RDBKTC	5,023	43.79	0.868	0.700	1.240	1.537
Watched children program	P3CHLPRM	4,968	90.14	0.626	0.423	1.481	2.193
Used computer for education 1-2 times/week	P3COMEDU	5,023	25.20	0.584	0.613	0.953	0.908
Visited museum	P3ARTSCI	5,039	38.89	1.288	0.687	1.876	3.519
Have library in neighborhood	P3LIBRAR	5,040	79.11	1.569	0.573	2.740	7.506
Attended back to school night	P3BTSNGT P3ATTBTS	3,586	70.48	1.195	0.762	1.569	2.463
Child characteristics (mean)							
Age of child in months	R3AGE	5,261	80.01	0.142	0.062	2.298	5.281
BMI	C3BMI	5,044	16.65	0.065	0.036	1.782	3.176
Number of hours in summer school	P3SUMSCH P3SUMSH	511	108.34	5.617	3.480	1.614	2.606
Number of days spend on vacation	P3SUMVAC P3NMDVAC	3,846	11.72	0.319	0.175	1.821	3.315
Number of times visited library during summer	P3VISLIB	5,020	4.19	0.115	0.087	1.326	1.758
Watched TV between breakfast/dinner (hours)	P3TVBRDH	4,919	1.44	0.034	0.020	1.717	2.948
Watched TV after dinner (hours)	P3TVAFDH	4,942	1.03	0.031	0.015	2.125	4.517
Median						1.984	3.945
Mean						1.944	4.015
Standard deviation						0.497	1.921
Coefficient of variation						0.255	0.478
Minimum						0.953	0.908
Maximum						2.745	7.534

Table 4-15.—ECLS-K, spring-first grade: standard errors and design effects for the full sample – Child and parent data

Survey item	Variable name	Number of cases	Estimate	Design SE	SRS SE	DEFT	DEFF
Child scores (mean)							
Reading score	C4RRSCAL	16,336	54.77	0.288	0.111	2.595	6.733
Math score	C4RMSCAL	16,639	42.79	0.196	0.074	2.661	7.083
General knowledge score	C4RGSCAL	16,324	34.00	0.158	0.061	2.604	6.781
Approaches to learning-Parent	P4LEARN	15,574	3.08	0.007	0.004	1.695	2.872
Self-control-Parent	P4CONTRO	15,573	2.93	0.009	0.004	2.113	4.466
Social interaction-Parent	P4SOCIAL	15,581	3.38	0.008	0.005	1.719	2.956
Withdrawn-Parent	P4SADLON	15,569	1.55	0.006	0.003	1.826	3.333
Impulsive/overactive-Parent	P4IMPULS	15,491	1.90	0.011	0.005	2.009	4.035
Approaches to learning-Teacher	T4LEARN	14,986	3.00	0.010	0.006	1.806	3.263
Self-control-Teacher	T4CONTRO	14,871	3.16	0.009	0.005	1.831	3.351
Interpersonal-Teacher	T4INTERP	14,829	3.09	0.009	0.005	1.684	2.836
Externalizing problems-Teacher	T4EXTERN	14,895	1.67	0.009	0.005	1.688	2.849
Internalizing problems-Teacher	T4INTERN	14,809	1.61	0.009	0.005	1.966	3.866
Child characteristics (percent)							
Lived in single parent family	P4HFAMIL	15,624	24.23	0.585	0.343	1.707	2.913
Lived in two-parent family	P4HFAMIL	15,624	73.49	0.634	0.353	1.794	3.219
Mom worked 35 hours+/week	P4HMEMP	11,002	66.37	0.654	0.450	1.452	2.107
Primary case is center-based	P4PRIMNW	6,402	35.60	1.359	0.598	2.271	5.158
Primary case is home-based	P4PRIMNW	6,402	64.40	1.359	0.598	2.271	5.158
Parents had high school or less	WKPARED	15,626	37.69	0.911	0.388	2.349	5.516
Parents attended PTA	P4ATTENP	15,605	40.75	0.905	0.393	2.301	5.294
Did homework 3-4 times per week	P4HMWORK	15,612	39.83	0.927	0.392	2.367	5.604
Parents helped with homework 3-4 times/week	P4HLPHWK	15,100	37.99	0.733	0.395	1.857	3.448
Practiced reading, writing, numbers daily	P4RDWRNM	15,605	52.13	0.764	0.400	1.911	3.651
Visited library	P4LIBRAR	15,597	45.18	0.806	0.399	2.022	4.089
Used computer 1-2 times per week	P4COMPWK	10,389	45.34	0.712	0.488	1.458	2.127
Have family rule for TV	P4TVRULE	15,467	90.31	0.332	0.238	1.397	1.953
HH received foodstamp in last 12 months	P4FSTAMP	15,545	15.93	0.652	0.294	2.221	4.932
Child characteristics (mean)							
Age of child in months	R4_AGE	16,675	87.17	0.078	0.037	2.132	4.547
Child's household size	P4HTOTAL	15,624	4.55	0.024	0.011	2.155	4.646
Number of children <18 in child's HH	P4LESS18	15,624	2.56	0.022	0.010	2.288	5.233
Number of siblings in HH	P4NUMSIB	15,624	1.52	0.019	0.009	2.043	4.173
Number of hours watched TV after dinner	P4TVAFDH	15,445	0.77	0.013	0.006	2.030	4.122
Median						2.016	4.062
Mean						2.007	4.135
Standard deviation						0.332	1.352
Coefficient of variation						0.165	0.327
Minimum						1.397	1.953
Maximum						2.661	7.083

Table 4-16.—ECLS-K, panel: standard errors and design effects for the full sample

Survey item	Variable name	Number of cases	Estimate	Design SE	SRS SE	DEFT	DEFF
Child characteristics (percent)							
Lived in single parent family	P4HFAMIL	4,681	23.31	0.834	0.618	1.350	1.823
Lived in two-parent family	P4HFAMIL	4,681	74.21	0.852	0.639	1.333	1.776
Mom worked 35 hours+/week	P4HMEMP	3,314	65.77	1.057	0.824	1.283	1.645
Primary case is center-based	P4PRIMNW	1,932	36.83	1.587	1.097	1.447	2.093
Primary case is home-based	P4PRIMNW	1,932	63.17	1.587	1.097	1.447	2.093
Parents had high school or less	WKPARED	4,682	36.99	1.017	0.705	1.442	2.079
Parents attended PTA	P4ATTENP	4,677	42.64	1.640	0.723	2.268	5.145
Did homework 3-4 times per week	P4HMWORK	4,676	38.68	1.366	0.712	1.918	3.678
Parents helped with homework 3-4 times/week	P4HLPHWK	4,533	36.81	1.126	0.716	1.572	2.471
Practiced reading, writing, numbers daily	P4RDWRNM	4,676	53.02	1.123	0.730	1.539	2.368
Visited library	P4LIBRAR	4,678	45.76	1.518	0.728	2.084	4.343
Used computer 1-2 times per week	P4COMPWK	3,109	45.55	1.116	0.894	1.249	1.561
Have family rule for TV	P4TVRULE	4,647	90.89	0.433	0.422	1.026	1.052
HH received foodstamp in last 12 months	P4FSTAMP	4,664	17.00	1.184	0.550	2.153	4.634
Child characteristics (mean)							
Age of child in months	R4_AGE	5,023	87.00	0.131	0.063	2.074	4.301
Child's household size	P4HTOTAL	4,681	4.61	0.040	0.021	1.940	3.763
Number of children <18 in child's HH	P4LESS18	4,681	2.58	0.033	0.017	1.944	3.780
Number of siblings in HH	P4NUMSIB	4,681	1.54	0.031	0.017	1.866	3.481
Number of hours watched TV after dinner	P4TVAFDH	4,643	0.79	0.020	0.012	1.663	2.765
Median						1.572	2.471
Mean						1.663	2.887
Standard deviation						0.358	1.207
Coefficient of variation						0.215	0.418
Minimum						1.026	1.052
Maximum						2.268	5.145

Table 4-17.—ECLS-K: median design effects for subgroups – child and parent data

Subgroups	Fall-first grade[1]		Spring-first grade[2]		Panel[3]	
	DEFT	DEFF	DEFT	DEFF	DEFT	DEFF
All students	1.984	3.945	2.016	4.062	1.572	2.471
Type of school						
Public	1.847	3.414	1.866	3.482	1.451	2.104
Private	1.628	2.654	1.961	3.845	1.797	3.231
Catholic private	1.611	2.593	1.751	3.065	1.418	2.012
Other private	1.586	2.515	1.781	3.173	1.751	3.065
Gender						
Male	1.585	2.513	1.718	2.951	1.385	1.919
Female	1.735	3.011	1.577	2.487	1.509	2.276
Race-ethnicity						
White	1.812	3.284	1.834	3.364	1.594	2.541
Black	1.275	1.626	1.683	2.831	1.361	1.851
Hispanic	1.389	1.928	1.417	2.006	1.152	1.327
Asian	1.518	2.304	1.570	2.466	1.491	2.224
Pacific Islander	2.605	6.785	1.140	1.300	3.022	9.132
American Indian	3.366	11.402	1.283	1.647	3.543	12.550
Other	1.161	1.349	1.364	1.859	1.217	1.480
Region						
Northeast	1.750	3.062	1.688	2.850	1.971	3.883
Midwest	1.887	3.562	2.133	4.550	1.622	2.630
South	1.879	3.553	2.062	4.253	1.767	3.122
West	1.734	3.008	1.707	2.911	1.497	2.242
Urbanicity						
Central city	1.772	3.141	1.839	3.381	1.472	2.166
Urban fringe and large town	1.694	2.868	1.732	3.000	1.691	2.859
Small town and rural area	2.088	4.371	2.210	4.884	1.949	3.800
SES quintiles						
First	1.332	1.774	1.456	2.119	1.283	1.646
Second	1.320	1.742	1.477	2.182	1.155	1.335
Third	1.350	1.822	1.422	2.022	1.361	1.852
Fourth	1.286	1.654	1.435	2.058	1.176	1.383
Fifth	1.406	1.977	1.424	2.027	1.286	1.653

[1] Each median is based on 20 items.

[2] Each median is based on 32 items.

[3] Each median is based on 19 items.

5. DATA COLLECTION METHODS AND RESPONSE RATES

5.1 Data Collection Methods

The following sections discuss the procedures used in the first grade data collection phase of the Early Childhood Longitudinal Study-Kindergarten Class of 1998-99 (ECLS-K).

The ECLS-K first grade data collection was conducted in the fall and spring of the 1999-2000 school year. While a full-scale data collection was mounted in spring-first grade, the effort for fall-first grade was limited to a 30 percent subsample of schools. Spring data collection included the direct child assessments, parent interview, teacher and school questionnaires, student record abstract and the facilities checklist. Fall data collection included just the direct child assessment and parent interview. The content and timeline of first grade data collections are shown in exhibit 5-1. Computer-assisted personal interviewing (CAPI) was the mode of data collection for the child assessment, and computer-assisted interviewing (both in-person and telephone) (CAI) was the mode of data collection for the parent interview; self-administered questionnaires were used to gather information from teachers, school administrators, and student records.

Exhibit 5-1.—Timeline of first grade data collection

First grade		
1999	2000	
Fall	Winter	Spring
Child assessments Parent interviews conducted (30 percent subsample)		Child assessments Parent interviews conducted Teacher information collected School and school administrator data Student record data

As in previous rounds of data collection, the field staff was organized into work areas, each with a data collection team consisting of one field supervisor and three 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, distributed and collected all school and teacher questionnaires, and completed school facilities checklists. Field staff members used in first grade were either continuing from previous rounds of data collection or were new to the project. Training approaches varied depending on the project experience of the field staff—continuing staff members were trained with a self-paced home study training package and new staff members were trained in person in a classroom setting.

5.2 Field Staff Training

All field staff members assigned to fall-first grade were continuing from previous rounds of ECLS-K data collection, so the training for fall-first grade data collection was accomplished using a home study training package. Two different training modes were employed for spring-first grade: one for continuing staff and one for new staff. Continuing staff members were trained via a home study package and new staff were trained in classroom sessions.

5.2.1 Fall-First Grade Field Staff Training

Field supervisors and assessors were trained for the fall-first grade data collection in September 1999, using a home study training program. The purpose of the home study training program was to introduce changes to the instruments between the kindergarten and first grade data collection and provide sufficient review and practice with the instruments and procedures to conduct child assessments and parent interviews. In total, 39 field supervisors and 102 assessors completed this training.

Home Study Training Program. As in the base year training, field supervisor 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, providing them with hands-on experience with the FMS and all field procedures prior to beginning data collection. The field supervisor home study program was 24 hours long. Topics included conducting the preassessment activities, identifying and locating children who moved from the school they attended in the previous round of data collection, identifying the regular teachers of ECLS-K children and linking them to those children, and the conduct of quality control observations. The field supervisors followed role plays and completed exercises that involved entering information specifically designed for training purposes into the FMS during the home study training exercises.

Assessors and field supervisors were trained on both the parent interview and child assessment. Child assessment home study materials included written exercises and self-study role plays. Parent interview home study materials included written exercises and practice on the parent interviews in pairs on the telephone with partners assigned by either field supervisors or field managers. Assessors entered information specifically designed for training purposes into the CAPI system on laptop computers using training scripts to complete the home study exercises. The assessor home study program was 16 hours long. Field supervisors ensured that the assessors on their teams were proficient in conducting the direct child assessment by observing a practice assessment with a child of appropriate age. They ensured that the assessors on their teams were proficient in conducting the parent interview by conducting a parent interview role play over the phone at the end of training.

5.2.2 Spring-First Grade Training

Field supervisors and assessors were trained for the spring-first grade data collection over a three-week period in March 2000. The majority of the field staff members were trained primarily using a home study training program. The purpose of the home study training program was to introduce changes to the instruments since the last time these returning staff had participated in ECLS-K (either spring-kindergarten or fall-first grade), provide sufficient practice with the instruments, and provide review and practice of the procedures that were used to conduct the child assessments and parent interviews. An in-person training session was conducted for staff new to the study in the spring. This training session incorporated any changes that were made in the fall-first grade training program using the revised instruments for the spring data collection. In total, 103 field supervisors and 317 assessors completed training.

Home Study Training Program. Field supervisor training was again conducted using the FMS. The field supervisor home study program was 24 hours long. The topics covered in the field supervisor home study training program included conducting the preassessment activities such as freshening the student sample, identifying and locating children who moved from their previous round school, identifying the regular and special education teachers of ECLS-K children and linking them to those children, distributing teacher questionnaires, distributing school administrator questionnaires, completing the facility checklist, and the conduct of quality control observations. A total of 76 field supervisors completed the home study training.

Assessors and field supervisors were trained on both the parent interview and child assessments. Child assessment home study materials included written exercises, self-study role plays, scripted role plays with partners, a training video focusing on improving assessment skills, and a final practice direct child assessment observed by the field supervisor. Parent interview home study materials included written exercises and trainees practicing the parent interviews in pairs on the telephone and completing a final role play on the parent interview with their field supervisor. Assessors again entered information designed specifically for training purposes into the CAPI system on laptop computers using training scripts to complete the home study exercises. The assessor home study program was 16 hours long. Field supervisors ensured that the assessors on their teams were proficient in conducting the direct child assessment by observing a practice assessment with a child of appropriate age and the parent interview by conducting a final role play over the telephone. A total of 175 assessors completed the home study training (see section 5.6 on data collection quality control).

Training Session for New Field Staff. The in-person training for new field staff (hired because of attrition in prior field staff) was held in March 2000 for nine days. This training incorporated any changes from the fall-first grade training, which were modified for an in-person training. The new spring-first grade parent and child home study materials were also incorporated into the training sessions. A total of 27 field supervisors and 142 assessors completed the in-person training.

5.3 Fall-First Grade Data Collection

The fall-first grade data collection was limited to 26.7 percent of the base year children in 30 percent of the ECLS-K originally sampled schools and was a design enhancement to measure the extent of and factors that contribute to summer learning/loss. See section 4.2 for details on the selection of the subsample. Only the direct child assessments and the parent interviews were included in this data collection.

5.3.1 Advance Mailings

Beginning in late summer 1999, letters were mailed to school administrators to confirm scheduled visits for the schools. A packet of material was also mailed to the school coordinators, who were identified by the school to act as a liaison with ECLS-K field staff, asking them to prepare for the preassessment visit to the school. The school coordinators were asked to distribute materials such as the study brochure, summary sheets describing the role of teachers in the study, and frequently asked questions to the first grade teachers.

5.3.2 Preassessment Contact

Most preassessment contact was made by telephone during September 1999 and at least one week before any assessment visit. During the preassessment contact, the field supervisor spoke with the school coordinator to confirm the dates of the assessment visits, answer any questions, review the list of ECLS-K children who were among the subsample selected for fall-first grade data collection, identify any of these children who were no longer enrolled at the school, collect locating information for those children, identify each enrolled child's regular teacher, review parental consent status and obtain accommodations information about the enrolled sampled children.

Reviewing Parent Consent

Although parental consent was obtained in the base year, field supervisors reviewed the parental consent with the school coordinator to determine if the base year consent was acceptable for fall-first grade. If the schools required consent to be re-obtained or they changed the type of consent that was required (e.g., from implicit to explicit), the supervisors re-obtained consent using the same procedures from the base year.

5.3.3 Conducting the Direct Child Assessment

The direct child assessment was administered during a 12-week field period that began in September and ended in late November 1999. In year-round schools, assessment teams made multiple visits to the school to conduct direct child assessments as not all children attended school at the same time. The assessment team visited the school when each track was in session to assess the sampled children. The direct child assessment was normally conducted in a school classroom or library. Before conducting the assessment, field supervisors and assessors set up the room for the assessment. They followed procedures for meeting children that were agreed upon during the preassessment contact with the school. Each child was signed out of his or her classroom prior to the assessment and signed back into the classroom upon the conclusion of the assessment

The direct child assessment took approximately 50 to 70 minutes per child. As in the spring-kindergarten data collection, for children with a language other than English in the home, the child's score on the oral language development scale (OLDS) administered in the prior round determined what path the child would follow in fall-first grade. Refer to section 5.4.2 of the ECLS-K Restricted-Use Base Year User's Manual (NCES 2000-097), August 2000, or section 5.4.2 of the ECLS-K Base Year Public-Use User's Manual (NCES 2001-029), February 2001, for more information on how home language was identified and how the OLDS was administered. Table 5-1 summarizes the OLDS routing patterns in fall-first grade.

Table 5-1.—Oral language development scale (OLDS) routing patterns in fall-first grade for previous round respondents

Home language	OLDS score in prior round	OLDS required in fall-first grade?	Fall-first grade OLDS score	Fall-first grade assessment path
English	Not applicable	No		English
Spanish	At or above cut score	No		English
	Below cut score	Yes		
			At or above cut score	English
			Below cut score	Spanish
Other language	At or above cut score	No		English
	Below cut score	Yes		
			At or above cut score	English
			Below cut score	Height/weight only

Children passing the cut score for the OLDS were administered the English direct child assessment and had their height and weight measured. Children who fell below the cut score for the OLDS and whose language was Spanish were administered the Spanish language version of the OLDS and parts of the direct child assessment that were translated into Spanish (the warm-up booklet and math). These children also had their height and weight measured. Children who fell below the cut score for the OLDS and whose language was other than Spanish had only their height and weight measured. Table 5-2 presents the percentage of children who were routed into the various assessment alternatives in fall-first grade. Overall five percent (281 children) of the sampled children were screened using the OLDS in the fall-first grade. Of the children whose home language was Spanish, 23 percent were at or above the cut score, and of the children whose home language was a language other than English or Spanish, 40 percent were at or above the cut score.

Table 5-2.—Oral language development scale (OLDS) routing results in fall-first grade

Category	Total screened (percent)	At or above cut score on OLDS (percent of those screened)	Below cut score on OLDS (percent of those screened)
Total sample	5	26	74
Spanish language children	80	23	77
Other language children	20	40	60

Slightly more than one percent of participating children were excluded from the direct child assessment due to a disability or required an accommodation offered in the assessment. Exclusion from the assessment and the identification and use of accommodations followed the procedures from the base year. Refer to section 5.4.2 of the ECLS-K Restricted-Use Base Year User's Manual (NCES 2000-097), August 2000 or section 5.4.2 of the ECLS-K Base Year Public-Use User's Manual (NCES 2001-029), February 2001 for more information on exclusions and identification of accommodations. Table 5-3 presents the number of children excluded from or requiring an accommodation to the direct child assessment.

5.3.4 Conducting the Parent Interview

For the fall-first grade round of data collection, the parent interview was administered between early September and mid-November 1999. The parent interview averaged 35 minutes. To administer the parent interview, assessors began by contacting parents using the contact information obtained during the previous parent interview.

Table 5-3.—Number of children excluded from or accommodated in the fall-first grade assessment

Category	Number of children
Excluded for disability	28
Setting accommodation (e.g. special lighting, adaptive chair)	10
Scheduling/timing accommodation	24
Health care aide present	5
Assistive device used/available	1

The ECLS-K fall-first grade parent interview was conducted primarily by telephone by field staff using CAI. The parent interview was conducted in person if the respondent did not have a telephone. Less than one percent of all completed parent interviews in fall-first grade were conducted in person. The parent interview was conducted primarily in English, but modifications were made to interview parents who spoke only Spanish. The questionnaire was translated into Spanish and printed on a hard-copy parent interview form. Bilingual interviewers were trained to conduct the parent interview in either English or Spanish. When the person answering the telephone was not able to speak English, and the field staff member was not bilingual and was unable to identify an English-speaking household member, the case was coded as a "language problem." The field supervisor reviewed the case and assigned it to a bilingual field staff person if the language was Spanish. Approximately six percent of the parent interviews were conducted in Spanish. Less than one percent of parent interviews could not be conducted because of language problems, meaning that the respondent spoke a language other than English or Spanish.

5.3.5 Conducting Data Collection on Children Who Withdrew from Their Previous Round School

During the preassessment contacts, field supervisors asked school coordinators to identify children who had withdrawn from the school since the spring of kindergarten. Of the base year participants that had transferred from their base year school, a random sample was identified to be included for data collection in the fall of first grade; see section 4.2 for more details. School staff was asked whether they knew the name and address of the school the child transferred into, as well as any new information about the child's household address. For the children who had moved from their spring-kindergarten 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-kindergarten 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.

A total of 1,178 children (18.2 percent of total sample in fall-first grade) were identified as having transferred schools between spring-kindergarten and fall-first grade. In addition, another 273 children were identified as having changed schools between fall-kindergarten and spring-kindergarten. Combining these two types of movers, there was a total of 1,451 children identified as "movers" in fall-first grade. Of the 1,451 mover children in fall-first grade, 671 were followed (46 percent of total movers). The remaining 780 mover children were part of the sample that would not be followed and were not included in the fall-first grade data collection. No child assessment or parent interview was conducted for these children.

Various data collection strategies were used for children who moved, depending on how they were classified. The following data collection approaches were attempted for children who moved and were flagged as "follow" in fall-first grade:

- Parent interviews were attempted for all children regardless of children's mover status.
- Data collected for children moving into cooperating base year sampled schools included the child assessment in the school.
- Data collected for children moving into nonsampled schools in base year cooperating districts included the child assessment conducted in the home.
- For children moving into sample schools that refused, schools in sampled districts that refused, or ineligible sampled schools, only the child assessment was conducted in the home.
- For children moving into schools in nonsampled districts or dioceses:
  - If the school was within the PSU, data collected included the child assessment in the home.
  - If the school was outside the PSU, no child data were collected.

- For children who were not enrolled in school in the spring (including children who were home schooled), data collected included the child assessment in the home, if the child was in the sample PSU. If the child was outside the sample PSU, no data other than the parent interview were collected.

As discussed in section 4.2, a random 50 percent of children who were included in the fall-first grade subsample were flagged to be followed for fall-first grade data collection in the event that they had transferred. Slightly more than one-half (54 percent, 780 children) of the children who moved were not followed and no data were collected for them or their parents, while 46 percent of the children who moved were followed. Of those flagged as "follow," 11 percent moved into a school outside the PSU, one percent were identified as out of the United States, and eight percent of the movers were unlocatable. One case flagged as "follow" was not fielded because, although the student was located, it was too late in the field period to field the transfer case to conduct a child assessment. This case is labeled *End of field period* in table 5-4. Of the 671 movers who were identified to be included in the fall-first grade data collection, 80 percent were fielded for data collection (i.e., were found and were eligible). Of the 533 cases that were finally located and eligible for data collection, 81 percent had a completed assessment. Table 5-4 presents the status of the 1,451 children who were movers in fall-first grade.

Table 5-4.—Number of children who moved in fall-first grade by completion category

Child in	Number of children	Percent
Total movers	1,451	100
Did not follow[1]	780	54
Followed[1]	671	46
Followed	671	100
Not fielded for assessment[2]	138	20
Unlocatable[2]	55	8
End of field period[2]	1	0
Nonsampled PSU[2]	72	11
Moved to outside the U.S.[2]	9	1
Deceased[2]	1	0
Fielded for assessment[2]	533	80
Fielded for assessment	533	100
Completed assessment[3]	430	81

[1] Percent based on total movers.

[2] Percent based on number of movers followed.

[3] Percent based on number of movers fielded.

5.4 Fall-First Grade Completion Rates

Table 5-5 presents weighted and unweighted child-level completion rates for the fall-first grade data collection, by school characteristics. On the ECLS-K, a completion rate is a response rate conditioned on the results of an earlier stage of data collection. For the first grade year of the ECLS-K, all completion rates are conditioned on the case having been a base year respondent, since data collection was attempted only for such cases. In fall-first grade, data collection was limited to the child assessment and parent interview. A weighted 90.3 percent of base year respondents completed the child assessment in fall-first grade and 88.6 percent of the children had a completed parent interview. By and large the completion rates are quite uniform across school characteristics. Students enrolled in Catholic schools, in rural schools outside of metropolitan statistical areas (MSA), and in schools with 750 or more enrolled completed the child assessment at higher than average rates (96.1, 96.9, and 95.9 percent respectively). The completion rate for the parent interview was lowest for students in large cities and those in schools with 90 percent to 100 percent minority enrollment (85.0 percent and 83.7 percent). It was highest for students enrolled in Catholic schools, in schools in large towns and in rural schools outside of MSAs (93.4, 94.7, and 94.7 percent respectively).

However, the category labeled “Unknown” in each of the different school characteristics has a substantially lower completion rate (table 5.5). Most of the children in this category are movers. The category includes children who were unlocatable as their whereabouts were unknown, whose cases could not be processed before the end of field period and those children who had moved into a nonsampled county. Under any of these circumstances if no information concerning the child’s school was available, they were included in the “unknown” category for each of the different school characteristics. Table 5-6 gives completion rates by mover status. A full 97 percent of nonmovers completed the child assessment. Movers who were located were assessed at a rate of 76 percent and, the others, including those not located, were not assessed. Even though these children were not administered a child assessment, wherever possible, a parent interview was conducted.

Table 5-7 contains child-level fall-first grade weighted and unweighted completion rates, this time broken out by child characteristics. Again, generally speaking completion rates do not differ greatly. The lowest completion rates for the child assessment are for blacks and for American Indian/Alaska Natives (87.6 and 87.9 percent respectively), the highest for Pacific Islanders (97.6 percent). For the parent interview the lowest completion rates are for Blacks and for Asians (83.4 and 85.1 percent respectively) and the highest for Whites and Pacific Islanders (92.0 and 92.5 percent).

Table 5-5.—ECLS-K fall-first grade: number of completed child-level cases and child-level completion rates, by school characteristic

School characteristics[1]	Child assessment: Completes[2]	Child assessment: Completion rates: Weighted	Child assessment: Completion rates: Unweighted	Parent interview: Completes[3]	Parent interview: Completion rates: Weighted	Parent interview: Completion rates: Unweighted
All children	5,291	90.3	93.6	5,071	88.6	89.8
School type						
Public	4,191	93.9	95.5	3,926	89.0	89.4
Private	1,060	94.8	95.8	1,028	93.0	92.9
Catholic	542	96.1	97.3	522	93.4	93.7
Other private	518	93.6	94.4	506	92.5	92.2
Unknown	40	28.7	26.0	117	73.3	76.0
Type of locale						
Large city	1,033	93.6	95.5	941	85.0	87.0
Mid-size city	882	93.9	96.1	834	89.3	90.8
Urban fringe of large city	1,379	93.2	94.4	1,286	88.4	88.0
Urban fringe of midsized city	296	93.2	94.3	286	90.4	91.1
Large town	255	94.5	97.0	252	94.7	95.8
Small town	513	94.0	97.3	491	90.6	93.2
Rural – Outside MSA	536	96.9	96.8	519	94.7	93.7
Rural – Inside MSA	271	94.3	95.4	266	93.4	93.7
Unknown	126	40.6	51.0	196	75.6	79.4
School size (total enrollment)						
1 to 299	1,246	94.7	95.6	1,178	89.7	90.3
300 to 499	1,420	93.1	95.6	1,355	90.0	91.2
500 to 749	1,327	92.8	94.4	1,262	89.1	89.8
750 or more	1,183	95.9	96.7	1,088	89.3	89.0
Unknown	115	40.9	49.4	188	76.3	80.7

[1] Based on ECLS-K survey data and not data from the sampling frame.

[2] Reading, math, or general knowledge assessment was scorable or child was language minority (not Spanish) or disabled.

[3] Family structure portion of parent interview was completed.

Table 5-5.—ECLS-K fall-first grade: number of completed child-level cases and child-level completion rates, by school characteristic (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%	1,604	95.1	95.9	1,568	93.3	93.8
11 – 49%	1,635	94.0	95.9	1,572	90.7	92.2
50 – 89%	1,005	92.2	94.5	926	87.4	87.0
90 – 100%	888	94.6	95.8	776	83.7	83.7
Unknown	159	46.8	56.4	229	77.4	81.2
Region						
Northeast	931	93.4	94.8	869	87.9	88.5
Midwest	1,299	92.0	94.7	1,232	87.8	89.8
South	1,758	95.2	96.4	1,676	91.1	91.9
West	1,263	94.7	95.8	1,177	89.8	89.2
Unknown	40	28.8	26.1	117	73.8	76.5

[1] Based on ECLS-K survey data and not data from the sampling frame.

[2] Reading, math, or general knowledge assessment was scorable or child was language minority (not Spanish) or disabled.

[3] Family structure portion of parent interview was completed.

Table 5-6.—ECLS-K fall-first grade: number of completed child-level cases and child-level completion rates, by child's mover status

	Child assessment			Parent interview		
		Completion rates			Completion rates	
Mover status[1]	Completes[2]	Weighted	Unweighted	Completes[3]	Weighted	Unweighted
All children	5,291	90.3	93.6	5,071	88.6	89.8
Mover status						
Mover	404	64.6	66.4	514	82.5	84.5
Located, followed	404	76.0	78.0	445	83.8	85.9
Other[4]	0	0.0	0.0	69	74.7	76.7
Nonmover	4,887	97.0	96.9	4,557	90.2	90.4

[1] This is the mover status used in weighting which does not consider children who moved into identified destination schools as movers.

[2] Reading, math, or general knowledge assessment was scorable or child was language minority (not Spanish) or disabled.

[3] Family structure portion of parent interview was completed.

[4] This category includes movers who could not be located, movers whose cases could not be processed before the end of the field period, and movers who moved into nomsampled PSUs.

Table 5-7.—ECLS-K fall-first grade: number of completed child-level cases and child-level completion rates, by child characteristic

Child characteristics[1]	Child assessment			Parent interview		
		Completion rates			Completion rates	
	Completes[2]	Weighted	Unweighted	Completes[3]	Weighted	Unweighted
All children	5,291	90.3	93.6	5,071	88.6	89.8
Gender						
Male	2,674	89.5	93.3	2,578	89.1	89.9
Female	2,606	91.0	94.0	2,493	88.6	90.0
Unknown gender	11	98.4	91.7	0	0.0	0.0
Race/ethnicity						
White (not Hispanic)	2,949	91.7	94.2	2,912	92.0	93.0
Black (not Hispanic)	783	87.6	92.2	718	83.4	84.6
Hispanic	878	89.1	93.3	809	85.6	86.0
Asian	285	89.8	93.4	260	85.1	85.2
Pacific Islander	96	97.6	97.0	89	92.5	89.9
American Indian or Alaskan Native	127	87.9	92.7	126	89.6	92.0
Other	153	88.8	93.9	150	87.7	92.0
Unknown race/ethnicity	20	81.6	80.0	7	20.2	28.0
Year of birth						
1992	1,611	89.1	93.6	1,577	89.6	91.6
1993	3,642	90.7	93.7	3,472	88.6	89.3
Other/unknown	38	95.3	92.7	22	47.9	53.7

[1] Based on ECLS-K survey data and not data from the sampling frame.

[2] Reading, math, or general knowledge assessment was scorable or child was language minority (not Spanish) or disabled.

[3] Family structure portion of parent interview was completed.

5.5 Spring-First Grade Data Collection

All children assessed during the base year, with a completed parent interview, or with a disability 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 moved ahead to second grade. Of those children who had transferred from their base year school, only a subsample was followed. In addition, children who were not in kindergarten in the United States during the 1998-99 school year and, therefore, did not have a chance to be selected to participate in the ECLS-K were added to the spring-first grade sample through a process referred to as "freshening." Freshening ensured that the data allowed for conclusions based on all first grade children, regardless of whether they were enrolled in kindergarten in the United States in fall 1998 or not. To accomplish the goals of following a subset of movers and freshening the sample, a random 50 percent subsample of base year schools were flagged. See section 4.3.2 for further details on how the ECLS-K sample was freshened to be representative of first graders in the United States and how movers were subsampled to be included in the survey.

5.5.1 Advance Contact with Respondents

In February 2000, letters were mailed to school coordinators confirming the scheduled visits to the school that had been arranged in the fall and winter. For those schools flagged for sample freshening, the school coordinators were also sent instructions for preparing the enrollment list of first grade students. Letters were also mailed to parents reminding them of the spring data collection activities.

5.5.2 Preassessment Contact

For schools that were not identified as part of the freshening and mover subsampling process, field supervisors conducted most preassessment activities by telephone starting in March 2000. The preassessment activities for these schools were similar to those conducted in previous rounds of data collection. For schools identified as part of the freshening and mover subsampling process, field supervisors visited each school in order to conduct freshening activities, as well as conduct other preassessment activities, such as identifying children who moved from their previous round school, identifying regular and special education teachers, linking them to children and distributing school administrator and teacher questionnaires. Section 4.3.2 describes the procedures used for identifying the subsample of schools that were part of the freshening and mover subsampling process.

Freshening Procedures

Each school that was sampled as a school from which children for the freshened sample would be selected, was asked to provide a list of all first grade students. When the field supervisor visited the school for the preassessment visit, he/she verified that the list was current (withdrawn students deleted, newly enrolled students included) and that the list contained no duplicate entries of student names. Once the list was ready, the supervisor identified the children previously sampled for the ECLS-K by highlighting their names on the list. On the freshening roster, the supervisors recorded the name of the first ECLS-K student on the enrollment list and the name of the student listed next if it was not that of another ECLS-K student. Then the supervisor asked the school coordinator to tell him/her if this non-ECLS-K student had been enrolled in kindergarten in the United States in fall 1998 and, based on the school coordinator's response, recorded either "Yes," "No," "Don't Know," or "Refused" for that child on the freshening roster. If the response was "Yes," that child was not added to the sample and the supervisor moved on to the next ECLS-K child on the enrollment list. If the response was "No," "Don't Know," or "Refused," the supervisor added the child from the enrollment list to the freshening roster. The supervisor continued with this process until he/she received a "Yes" response to the question "Was this child enrolled in kindergarten in the United States in fall 1998?" These procedures were repeated for all the ECLS-K sampled children on the enrollment list.

The definition of "enrolled in kindergarten in the United States in fall 1998," which was provided to the school coordinators, was very specific and included the following explanation of terms:

In Kindergarten in U.S. in Fall, 1998? The "In kindergarten in U.S. in fall, 1998?" question was to be taken literally. It was not meant to determine when or where a child was in kindergarten, but whether each currently enrolled first grade child was eligible for selection for the ECLS-K sample. To have been eligible for selection in the ECLS-K sample, the first grade child must have met three criteria. He/she must have been:

1. **In kindergarten:** In the ECLS-K, kindergarten is defined as the year of school primarily for five-year-olds prior to the first grade and includes all types of programs: public, private, full-day, part-day, regular, and transition programs. In addition to

"regular kindergartners" (i.e., the traditional year of school primarily for five-year-olds prior to first grade) found in "regular" public and private schools, the following kindergarten programs should be included:

- Kindergarten programs located in early childhood programs (e.g., nursery schools, early childhood learning centers, daycare centers);
- Transitional programs:
    1. **Transitional (or readiness) Kindergarten—**extra year of school for kindergarten-age-eligible children who are judged not ready for kindergarten.
    2. **Transitional First (or Pre-First) Grade—**extra year of school for children who have attended kindergarten and have been judged not ready for first grade.
- Ungraded/multigraded programs with kindergarten-age children
- Special education programs/classes with kindergarten-age children

If a child was enrolled in any of these types of programs the answer to "In Kindergarten" is "yes." If the child was in any other type of program or was in a different grade level (e.g., first grade), the answer to "in kindergarten" is "no."

2. **In United States:** The 50 states of the United States, that is, the 48 contiguous states, Hawaii, Alaska, and the District of Columbia.

3. **In fall 1998:** September through December 1998.

All nonsampled ECLS-K children recorded on the freshening roster for whom the school coordinator responded "No," "Don't Know," or "Refused" to the question "Was this child enrolled in kindergarten in the United States in fall 1998?" were added to the FMS and transmitted to the Westat home office. The supervisor collected contact information on each added child's parents from the school coordinator, prepared parent information packets, including consent forms, and sent them to the parents of the added children. Meanwhile, the Westat home office created the child assessment and parent interview case and fielded the case for data collection.

Prior to assessing children who were added to the sample through freshening, the supervisor contacted the parent(s) of each freshened child for two reasons: (1) to confirm that the freshened child was not enrolled in kindergarten in the Untied States in fall 1998, and (2) to obtain outstanding consent forms. If the parent(s) of the freshened child indicated that the child *was* enrolled in kindergarten in fall 1998, no data were collected from the child or the parent. However, if the parent(s) of the freshened child confirmed that the child *was not* enrolled in kindergarten in fall 1998, the child was included in the ECLS-K first grade sample, and both the child assessment and parent interview were conducted. Data for these freshened children were also collected from the teachers and school records. Of the 210 children who were added to the sample through freshening, 162 (77 percent) were confirmed by the parents as not enrolled in kindergarten in the United States in fall 1998.

5.5.3 Conducting Direct Child Assessments

The direct child assessments were conducted between March and June 2000. For children with a language other than English in the home, the child's score on the previous OLDS determined what path was followed in the spring. Table 5-8 summarizes the OLDS routing patterns in spring-first grade.

Table 5-8.—Oral language development scale (OLDS) routing patterns in spring-first grade for previous round respondents

Home language	OLDS score in prior round	OLDS required in spring-first grade?	Spring-first grade OLDS score	Spring-first grade assessment path
English	Not applicable	No		English
Spanish	At or above cut score	No		English
	Below cut score	Yes		
			At or above cut score	English
			Below cut score	Spanish
Other language	At or above cut score	No		English
	Below cut score	Yes		
			At or above cut score	English
			Below cut score	Height/weight only

Children who scored at or above the cut point on the OLDS in the previous assessment were automatically routed by CAPI to take the assessment in English; the OLDS was not administered again. Children who scored below the cut point in the OLDS in the previous assessment (spring-kindergarten for most, fall-first grade for most of the remaining, and fall-kindergarten for a few) were administered the OLDS again in spring-first grade and routed according to the new spring-first grade OLDS score. Children taking the direct assessment for the first time in the spring, for example students who were included in sample through the freshening process, with a language other than English in the home, were routed according to their home language as determined from school records or the child's teacher and were administered the OLDS. Table 5-9 presents the percentage of children who were routed into the various assessment alternatives in spring-first grade. Five percent of the sampled children (798 children) were screened using the OLDS in the spring-first grade. Overall 56 percent of the screened children were at or above the cut score on the OLDS and were administered the English direct child assessment. Of the children whose home language was Spanish, 50 percent were at or above the cut score, and of the children whose home language was a language other than English or Spanish, 77 percent were at or above the cut score.

Table 5-9.—Oral language development scale (OLDS) routing results in spring-first grade

Category	Total screened (percent)	At or above cut score on OLDS (percent of those screened)	Below cut score on OLDS (percent of those screened)
Total sample	5	56	44
Spanish language children	77	50	50
Other language children	23	77	23

Approximately 1.5 percent of participating children were excluded from the direct child assessment or required an accommodation offered in the assessment. The patterns for accommodations were the same as in previous rounds of data collection. Table 5-10 presents the number of children excluded from or requiring an accommodation to the direct child assessment in the spring of kindergarten.

Table 5-10.—Number of children excluded from or accommodated in the spring-first grade assessment

Category	Number of children
Excluded for disability	47
Setting accommodation (e.g. special lighting, adaptive chair)	55
Scheduling/timing accommodation	119
Health care aide present	21
Assistive device	12

5.5.4 Conducting the Parent Interview

Parent interview procedures mirrored those of the fall-first grade. The parent interview was administered, primarily by telephone interview using CAI, between March and early July 2000. The parent interview averaged 45 minutes. As in previous rounds of data collection, the parent interview was conducted in person if the respondent did not have a telephone. Three percent of all completed parent interviews were conducted in person. Six percent of the completed parent interviews were conducted in a language other than English with 94 percent of completed non-English interviews conducted in Spanish. The refusal conversion workshop developed for fall-kindergarten was implemented at the end of the spring field period to improve response rates. The special effort to build parent response rates was conducted between June 5 and July 8, 2000, and yielded an additional ten percentage points in the response rate. Four percent of the parent interviews were not completed because of locating problems.

5.5.5 Teacher and School Data Collection

Data were collected from school administrators, regular classroom teachers, and special education teachers between March and June 2000. Teachers were asked to complete individual ratings for the sampled children in their classrooms, and they were reimbursed seven dollars for each child rating (teacher questionnaire C) they completed. In addition, school staff was asked to complete a student record abstract after the school year closed. The school staff was reimbursed seven dollars for every student record abstract they completed. Field supervisors also completed a facilities checklist for each sampled school.

Field supervisors distributed the school and teacher questionnaires in a variety of ways, depending on the preference of the school staff. Questionnaires were distributed during the preassessment visit (if one was held), by mail, and during the assessment visits. During the field period, field supervisors followed up with school administrators and teachers by telephone and visits to the schools to ensure that completed questionnaires were mailed to Westat. To improve response rates, in early September 2000, a package was mailed to all schools with outstanding school administrator questionnaires, teacher questionnaires, or student record abstracts with a request to complete and return questionnaires. Field staff prompted by telephone for the return of school administrator questionnaires, teacher questionnaires, and student record abstracts through October 2000.

5.5.6 Conducting Data Collection on Children Who Withdrew from Their Previous Round School

During the preassessment contacts, field supervisors asked school coordinators to identify children who had withdrawn from the school since the spring of kindergarten. School staff was asked whether they knew the name and address of the school the child transferred into, as well as any new information about the child's household address. For the children who had moved from their spring-kindergarten 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-kindergarten 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.

A total of 3,454 children (16.1 percent of total sample in spring-first grade) were identified as having transferred from the school in which they were enrolled in the previous round of data collection (their spring-kindergarten school if they were not included in the fall-first grade subsample or their fall-first grade school if they were included in the fall subsample). In addition, another 2,125 children were identified as having changed schools between fall kindergarten and spring kindergarten. Combining these two types of movers, there was a total of 5,759 children identified as 'movers' in spring-first grade. Of the 5,759 mover children in spring-first grade, 2,911 were followed (51 percent of total movers). The remaining 2,848 mover children were part of the subsample that would not be followed and were not included in the spring-first grade data collection. No child assessment or parent interview was conducted for these children.

Different data collection strategies were followed for children who moved, depending on how they were classified. Data collection was attempted for children who moved and were flagged as "follow" in spring-first grade:

- Parent interviews were attempted for all children regardless of children's mover status.
- Data collected for children moving into cooperating base year sampled schools included the child assessment in the school, school administrator questionnaire, regular or special education teacher questionnaires, facilities checklist, and student record abstract forms.

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

- For children moving into sample 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, only the direct child assessment was 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 PSU, data collected included the child assessment in the school, school administrator questionnaire, regular or special education teacher questionnaires, facilities checklist, and student record abstract forms if school permission was obtained. If school permission was not obtained, the assessment was conducted in the home and no school or teacher data were collected.

  - 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 home schooled), data collected included the child assessment in the home if the child was in the sample PSU. If the child was outside the sample PSU, no data were collected.

As discussed in section 4.3.1, all children in a random 50 percent subsample of base year schools were flagged to be followed for spring-first grade data collection in the event that they had transferred. Slightly less than half (49 percent, 2,848 children) of the children who moved were not followed and no data were collected for them or their parents, while 51 percent of the children who moved were followed. Of those flagged as "follow," 16 percent moved into a school outside the PSU, two percent were identified as out of the United States, and nine percent of the movers were unlocatable. Another 2 percent were not fielded because, although the students were located, it was too late in the field period to field the transfer cases to conduct a child assessment. These cases are labeled *End of field period* in table 5-11. Of the 2,911 movers who were identified to be included in the spring-first grade data collection, 71 percent were fielded for data collection (i.e., were found and were eligible). Of the 2,070 cases that were finally located and eligible for data collection, 87 percent had a completed assessment. Table 5-11 presents the status of the 5,759 children who were movers in spring-first grade.

5.6 Data Collection Quality Control

A continuous quality assurance process was applied to all data collection activities, but with a particular focus on the assessments. The process was incorporated in all stages (i.e., during development, in the staff training program, through certification, and as part of the ongoing staff observations and evaluation activities).

Table 5-11.—Number of children who moved in spring-first grade by completion category

Child in	Number of children	Percent
Total movers	5,759	100
Did not follow[1]	2,848	49
Followed[1]	2,911	51
Followed	2,911	100
Not fielded for assessment[2]	840	29
Unlocatable[2]	271	9
End of field period[2]	70	2
Nonsampled PSU[2]	454	16
Moved to outside the U.S. [2]	44	2
Deceased[2]	1	0
Fielded for assessment[2]	2,070	71
Fielded for assessment	2,070	100
Completed assessment[3]	1,792	87

[1] Percent based on total movers.

[2] Percent based on number of movers followed.

[3] Percent based on number of movers fielded.

Data collection quality control efforts began with the additional development and testing of redesigned sections of the CATI/CAPI applications and 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 home study training, field managers certified supervisors and supervisors certified assessors on the parent interview and the child assessment. During the in-person assessor training, field staff practiced conducting the parent interview in pairs and practiced the direct child assessment with kindergarten, first, and second grade children brought to the training site for this purpose. 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. A sample of the assessor-completed OLDS score sheets was rescored in the home office for quality control purposes.

5.6.1 Child Assessment Observations

Field supervisors conducted on-site observations of the child assessments. In fall and spring-first grade, two observations were completed for each assessor. The first observation was within two weeks after the assessments began, and the second observation was completed within three weeks of the first observation.

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

- Rapport building and working with the child—use of neutral praise and the assessor's response to various child behaviors.
- Cognitive assessment activities—reading questions verbatim, the use of acceptable probes, the use of appropriate hand motions, and the absence of coaching.
- Specific assessment activities—correctly coding answers to open-ended questions in the assessment, weighing and measuring the child correctly, 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.

5.6.2 Parent Validations

Approximately ten percent of the completed parent interviews were validated. The first parent interview completed by an assessor was always validated. 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. This ensured that ten percent of each assessor's cases were selected for validation. The parent validation was approximately five minutes long, conducted by telephone.

Field supervisors used a standardized parent validation script when calling the parents. The script covered the following topics:

- Verification of the child's name, date of birth, and gender; and
- Between eight and ten questions from the current round interview were re-asked of the parent.

During the validation process, no evidence was found of parent interviews being falsified.

5.6.3 School Validations

To ensure that assessments proceeded smoothly, a validation call was completed with the school principal in at least two of each supervisor's assigned originally sampled schools in both the fall- and spring-first grade collections.

Field managers conducted the school validations by telephone. The first school that each team completed 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 provided, including in-person observation of the supervisor's next school if necessary. In fall-first grade, a total of 72 or 23 percent of the fielded 310 originally sampled schools were validated. In spring-first grade, a total of 198 or 20 percent of the fielded 970 originally sampled schools were validated.

Field managers used a standardized script when calling the school principals. The script covered the following topics:

- How well the ECLS-K supervisor organized and executed the sampling tasks;
- An overall rating of how the assessments went;
- Feedback about the study from the children and kindergarten teachers;
- Suggestions for improving procedures and making it easier for a school to participate; and
- General comments and suggestions.

No problems were encountered during the school validation process.

5.6.4 Quality Control of the Oral Language Development Scale Scoring

The OLDS used to screen children for English language proficiency included the "Let's Tell Stories" subtest. This subtest involved reading the child a short story and having the child repeat it back to the assessor. The child's responses were recorded verbatim and scored by the assessor. Responses to this subtest are unique to each child, and it was important for interviewers' and coders' scoring of the child's responses to match the preLAS®2000 standards.

ECLS-K assessors were trained to conduct the OLDS using audiotapes of the stories and children's responses to the stories. Assessors listened to the audiotaped stories and to the child's responses and recorded the child's responses verbatim. Then the assessor scored the story using the preLAS®2000 rules. Reasons for scoring each story a particular way were discussed in detail. Differences between the assessor's scores and the correct scores were discussed during training, so assessors could understand the difference between the scores. Several stories in each scoring category were provided for practice to fine-tune the assessor's scoring. Then the scoring ability of each assessor was tested. Only assessors who scored a 90 percent accuracy in scoring the training stories as matched against the preLAS®2000 samples were allowed to conduct the OLDS.

A ten percent sample of each assessor's OLDS stories was recoded in the home office by coders. The coders received the same training as the assessors. Coders then scored the stories independently. If the home office coders' scores differed from the assessor, the two scores were verified by the coding supervisor. All cases were adjudicated by lead trainers for the OLDS. As in the base year, approximately 66 percent of the stories had complete score agreement between the assessor, coder, and lead trainer. The additional 33 percent of the stories had score agreement by two of the three scorers.

5.6.5 Assessor Effects

Individual Test Administrator Effects and Design Effects

In the base year, a multilevel analysis[1] was carried out to estimate components of variance in fall- and spring-kindergarten direct cognitive scores associated with the (1) students, (2) schools, (3) team leaders, and (4) individual test administrators. A similar analysis was conducted using the spring-first grade direct cognitive scores. This secondary analysis was motivated by Westat's earlier finding of larger-than-expected design effects. In addition, the impact on the above sources of variance of the SES indicator (parent's education) was also estimated. It was expected that much of the clustering of students within neighborhood schools (hence higher design effects) could be explained by SES. To examine whether this held true for first grade, a similar analysis was conducted.

In addition to the potential clustering effects related to shared parent SES within schools, there was a concern that the individual mode of administration might inject additional and unwanted variance to both the individual and the between school components of variance in the cognitive scores. Since it is more difficult to standardize test administrations when tests are individually administered, this source of variance could contribute to the high design effects if the individual assessors differed systematically in their modes of administration.

The component of variance associated with the individual test administration effect was negligible in all three cognitive areas and thus had little or no impact on the design effects. Much of the design effects with respect to cognitive scores could be explained by parents' SES.

---

[1] Bryk, A. & Raudenbush, S.W. (1992). *Hierarchical Linear Models: Applications and data analysis methods.* New York: Sage Publications. Snijders, T. & Bosker, R. (1999). *Multilevel Analysis – An introduction to basic and advanced multilevel modeling.* London: Sage Publications.

The following table presents information on the intra-class correlations and design effects for the Item Response Theory (IRT) scaled measures and also for the unscaled routing test in reading and mathematics. The intra-class correlations and design effects are presented for the unscaled routing tests in order to evaluate how much the IRT scaling may be contributing to the design effects. Inspection of table 5-12 indicates that the intra-class correlations are consistently albeit trivially lower for the unscaled routing tests as compared to the IRT scaled counterpart tests. The slight difference observed are probably due to the shrinking of the tails that occurs in IRT scoring since they are estimates of "true scores."

Table 5-12.—Intra-class correlations and estimated design effects for the spring-first grade direct cognitive tests (unweighted)

Cognitive test	Level one variances (student)	Level two variances (school)	Intra-classroom correlations	Estimated design effect
Reading IRT	.318	.101	.241	2.75
Reading Routing	10.022	2.839	.221	2.61
Mathematics IRT	.503	.150	.230	2.67
Mathematics Routing	6.848	1.826	.210	2.53
General Knowledge IRT	.217	.112	.341	3.48
Highest Parent Education	.852	.443	.342	3.40

The intra-class correlation is also presented for the background variable "highest parental education" as an indicator of the potential for clustering with respect to cognitive-related variables that is likely to occur in neighborhood schools.

The estimated intra-class correlations and design effects for the cognitive tests in table 5-12 were computed on the cross-sectional spring-first grade sample using a two-level model without the sample weights. A t-test was done to examine the effects on the intra-class correlation of not using the weights. The test showed that weights and unweighted analysis gave similar results.

Table 5-13 presents the results of a three-level "null" model with student at level one, interviewer at level two, and "work area" at level three. It was suggested that the training and supervision of interviewers by a work area leader (the field supervisor) might lead to a clustering effect by work area. This model was used to examine the potential impact of the interviewer and field supervisor on the data collection process and subsequently on the intra-class correlations. Inspection of table 5-13 suggests that the interviewer effects in spring-first grade are relatively trivial and should not have any systematic impact on analysis. Similar conclusions apply to the work area effect also.

Table 5-13.—Components of variance for three-level model, including the interviewer effect for spring-first grade cognitive scores

Cognitive tests	Level one (student)	Level two (interviewer)	Level three (work area)
Reading (IRT)	.381 (94%)	.007 (2%)	.019 (4%)
Mathematics (IRT)	.587 (92%)	.011 (2%)	.039 (6%)
General Knowledge (IRT)	.276 (85%)	.012 (4%)	.036 (11%)

5.7 Spring-First Grade Completion Rates

In the sections that follow, spring-first grade completion rates are presented for three groups of students: (1) students sampled in kindergarten, (2) students sampled in first grade through the freshening procedure, and (3) both groups combined.

5.7.1 Students Sampled in Kindergarten

Table 5-14 presents weighted and unweighted child-level completion rates for spring-first grade data collection, broken out by school characteristics. These rates pertain to children who were sampled as part of the kindergarten cohort in the base year. (Rates for students sampled in first grade through the student sample freshening procedure can be found in table 5-18.) Relative to fall-first grade the overall completion rates for the child assessment (88.0 percent) and the parent interview (84.5 percent) dropped two and four points respectively in spring-first grade. The drop in overall child assessment completion rates is tied to lower completion rates for movers, particularly those moving outside of the sampled PSUs, which accounts for approximately 30 percent (unweighted) of the movers' assessment nonresponse. Nearly half (45 percent) of the nonresponse for the child assessment is associated with those children for whom the school type was "Unknown" in the base year (see section 5.4). For the vast majority of students the child assessment completion rate increased slightly between fall- and spring-first grade, and this is true for nearly all school characteristics. The decrease in

Table 5-14.—ECLS-K spring-first grade: number of completed child-level cases and child-level completion rates, by school characteristic—children sampled in the base year

School characteristics[1]	Child assessment			Parent interview			School administrator questionnaire		
		Completion rates			Completion rates			Completion rates	
	Completes[2]	Weighted	Unweighted	Completes[3]	Weighted	Unweighted	Completes[4]	Weighted	Unweighted
All school types	16,593	88.0	91.8	15,522	84.5	85.8	14,619	76.3	81.4
School type									
Public	13,259	94.7	95.3	11,957	85.9	86.0	11,652	81.6	83.8
Private	3,273	94.7	95.1	3,125	91.3	90.8	2,967	82.2	86.2
Catholic	1,960	97.4	97.3	1,847	92.6	91.7	1,828	88.1	90.7
Other private	1,313	91.3	92.0	1,278	89.5	89.6	1,139	74.6	79.8
Unknown school type	61	7.5	8.3	440	59.4	60.1	0	0.0	0.0
Type of locale									
Large city	2,965	92.2	93.6	2,524	78.6	79.7	2,422	72.3	76.5
Mid-size city	2,925	94.9	96.0	2,687	87.2	88.2	2,585	80.4	84.8
Urban fringe of large city	4,595	94.4	94.5	4,209	86.7	86.6	4,110	81.6	84.5
Urban fringe of mid-size city	1,052	94.0	95.0	956	84.9	86.4	906	77.6	81.8
Large town	442	96.3	96.9	422	91.0	92.5	408	88.8	89.5
Small town	1,358	97.0	96.9	1,278	91.1	91.2	1,254	88.8	89.5
Rural – Outside MSA	1,853	97.0	97.0	1,757	92.5	92.0	1,792	92.9	93.8
Rural – Inside MSA	1,069	94.0	95.2	1,006	88.1	89.6	998	89.4	88.9
Unknown	334	23.1	33.1	683	64.1	67.6	144	9.2	16.2
School size (total enrollment)									
1 to 299	3,777	93.9	95.0	3,536	87.6	88.9	3,457	83.7	86.9
300 to 499	4,831	95.4	95.9	4,481	88.8	89.0	4,294	82.7	85.2
500 to 749	4,261	93.7	94.4	3,862	85.5	85.6	3,890	83.1	86.2
750 or more	3,491	95.5	95.6	3,049	84.1	83.5	2,978	80.9	81.6
Unknown	233	18.2	25.7	594	62.8	65.5	0	0.0	0.0

[1] Based on ECLS-K survey data and not data from the sampling frame.

[2] Reading, math, or general knowledge assessment was scorable or child was language minority (not Spanish) or disabled.

[3] Family structure portion of parent interview was completed.

[4] A completed questionnaire was defined as one that was not completely left blank.

Table 5-14.—ECLS-K spring-first grade: number of completed child-level cases and child-level completion rates, by school characteristic—children sampled in the base year (continued)

School characteristics[1]	Child assessment			Parent interview			School administrator questionnaire		
		Completion rates			Completion rates			Completion rates	
	Completes[2]	Weighted	Unweighted	Completes[3]	Weighted	Unweighted	Completes[4]	Weighted	Unweighted
Percent non-White enrolled									
0 – 10%	5,320	95.9	96.0	5,102	91.9	92.1	4,879	87.0	88.1
11 – 49%	5,264	94.9	95.8	4,858	87.5	88.4	4,729	83.8	86.0
50 – 89%	2,901	92.9	93.1	2,561	83.1	82.2	2,508	78.8	80.5
90 – 100%	2,779	93.9	95.0	2,315	78.8	79.2	2,425	77.4	82.9
Unknown	329	25.3	32.6	686	65.5	68.0	78	7.8	8.8
Region									
Northeast	3,057	94.3	95.3	2,782	85.9	86.7	2,634	79.3	82.1
Midwest	4,239	95.3	96.3	3,921	88.3	89.1	3,682	80.2	83.7
South	5,535	95.5	95.5	5,068	86.9	87.5	5,072	86.0	87.5
West	3,700	92.5	93.5	3,310	84.0	83.6	3,231	77.3	81.6
Unknown	62	7.7	8.6	441	60.3	61.2	0	0.0	0.0

[1] Based on ECLS-K survey data and not data from the sampling frame.

[2] Reading, math, or general knowledge assessment was scorable or child was language minority (not Spanish) or disabled.

[3] Family structure portion of parent interview was completed.

[4] A completed questionnaire was defined as one that was not completely left blank.

Table 5-14.—ECLS-K spring-first grade: number of completed child-level cases and child-level completion rates, by school characteristic—children sampled in the base year (continued)

School characteristics[1]	Teacher questionnaire A			Teacher questionnaire B			Teacher questionnaire C		
		Completion rates			Completion rates			Completion rates	
	Completes[4]	Weighted	Unweighted	Completes[4]	Weighted	Unweighted	Completes[4]	Weighted	Unweighted
All school types	15,021	78.1	83.6	14,878	77.5	82.8	14,991	78.0	83.5
School type									
Public	11,883	83.1	85.4	11,795	82.6	84.8	11,875	83.1	85.4
Private	3,138	87.2	91.2	3,083	86.0	89.6	3,116	86.6	90.5
Catholic	1,912	92.9	94.9	1,910	92.7	94.8	1,894	92.4	94.0
Other private	1,226	80.0	85.9	1,173	77.5	82.2	1,222	79.2	85.6
Unknown school type	0	0.0	0.0	0	0.0	0.0	0	0.0	0.0
Type of locale									
Large city	2,502	73.3	79.0	2,448	71.1	77.3	2,519	73.9	79.5
Mid-size city	2,705	84.7	88.7	2,658	83.9	87.2	2,700	84.8	88.6
Urban fringe of large city	4,195	82.9	86.3	4,140	81.9	85.2	4,150	82.1	85.4
Urban fringe of mid-size city	873	77.0	78.9	892	78.5	80.7	881	77.3	79.7
Large town	440	95.7	96.5	441	95.7	96.7	437	94.4	95.8
Small town	1,300	91.2	92.8	1,297	91.5	92.6	1,294	90.9	92.4
Rural – Outside MSA	1,770	92.8	92.8	1,774	93.2	93.0	1,784	93.7	93.5
Rural – Inside MSA	1,009	88.5	89.8	1,000	87.9	89.0	1,000	87.2	89.0
Unknown	227	14.3	25.5	228	14.4	25.6	226	14.4	25.4
School size (total enrollment)									
1 to 299	3,417	82.0	85.9	3,382	81.7	85.1	3,429	82.8	86.2
300 to 499	4,497	85.8	89.3	4,434	84.5	88.1	4,483	85.6	89.0
500 to 749	3,805	81.9	84.3	3,757	81.3	83.2	3,772	81.0	83.6
750 or more	3,173	85.6	86.9	3,174	85.4	87.0	3,178	85.5	87.1
Unknown	129	7.7	16.4	131	7.9	16.6	129	7.8	16.4

[1] Based on ECLS-K survey data and not data from the sampling frame.

[2] Reading, math, or general knowledge assessment was scorable or child was language minority (not Spanish) or disabled.

[3] Family structure portion of parent interview was completed.

[4] A completed questionnaire was defined as one that was not completely left blank.

Table 5-14.—ECLS-K spring-first grade: number of completed child-level cases and child-level completion rates, by school characteristic—children sampled in the base year (continued)

School characteristics[1]	Teacher questionnaire A			Teacher questionnaire B			Teacher questionnaire C		
		Completion rates			Completion rates			Completion rates	
	Completes[4]	Weighted	Unweighted	Completes[4]	Weighted	Unweighted	Completes[4]	Weighted	Unweighted
Percent non-White enrolled									
0 – 10%	4,898	86.5	88.4	4,899	86.7	88.4	4,905	86.7	88.5
11 – 49%	4,931	86.8	89.8	4,854	85.9	88.4	4,869	85.4	88.6
50 – 89%	2,586	80.7	83.0	2,523	79.1	81.0	2,563	80.3	82.3
90 – 100%	2,389	77.2	81.7	2,383	76.6	81.5	2,437	79.1	83.3
Unknown	217	16.1	24.4	219	16.3	24.6	217	16.3	24.4
Region									
Northeast	2,695	80.6	84.0	2,711	81.3	84.5	2,723	81.3	84.9
Midwest	3,792	82.4	86.2	3,776	82.1	85.9	3,791	82.6	86.2
South	5,204	87.7	89.8	5,130	86.7	88.5	5,171	86.9	89.2
West	3,330	79.9	84.1	3,261	78.4	82.4	3,306	79.8	83.5
Unknown	0	0.0	0.0	0	0.0	0.0	0	0.0	0.0

[1] Based on ECLS-K survey data and not data from the sampling frame.

[2] Reading, math, or general knowledge assessment was scorable or child was language minority (not Spanish) or disabled.

[3] Family structure portion of parent interview was completed.

[4] A completed questionnaire was defined as one that was not completely left blank.

parent interview rates, on the other hand, is widespread and cannot be attributed to one or two anomalous reporting categories. It is most pronounced for large cities where the 78.6 percent completion rate is 6.4 points down from the fall.

The overall weighted completion rate for the school administrator questionnaire is 76.3 percent. This is lower than the overall rate for the SAQ in the base year by close to 10 percentage points. The extremes range from 72.3 percent for children in large cities to 88.8 percent or higher for those in large towns, small towns, or rural areas. The lower rates for the SAQ in the first-grade year are due to movers. The transfer schools that received movers were given an SAQ designed for schools that were not already in the ECLS-K sample. The completion rate for these *new* SAQs is only 34.7 percent compared with 88.6 percent for *returning* SAQs given to schools that are in the original school sample. If not for the movers, the completion rate for the SAQ is in fact slightly higher than the rate for the base year (88.6 percent versus 85.9 percent). It is worth noting that the completion rates for the SAQ continue to be lower for schools with higher percentage minorities. In fact, as the percent minority enrolled in the schools increases, the completion rate for the SAQ decreases. However, this disparity decreased considerably in the first-grade year compared to the base year, reflecting increased data collection efforts targeted toward their schools.

All three of the teacher questionnaires were completed at an overall rate of approximately 78 percent. Often the rates are substantially higher, over 90 percent for Catholic schools and for the less urban areas. At the other end of the spectrum are schools in large cities which had completion rates in the low 70s, and schools with 90 percent or more minority enrollment which had rates in the mid to upper 70s.

The rate at which these survey instruments were completed varies markedly by mover status and within movers, by whether or not the child was located and followed. As presented in table 5-15 the completion rate for the child assessment was 95.9 percent for children still enrolled in their base year school. For movers it dropped over ten points to 85.5 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 88.0 percent for nonmovers to 78.2 percent for movers who were located and followed for the purposes of the child assessment, to 60.4 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 assessments, wherever possible, a

Table 5-15.—ECLS-K spring-first grade: number of completed child-level cases and child-level completion rates, by child's mover status—children sampled in the base year

	Child assessment			Parent interview			School administrator questionnaire		
		Completion rates			Completion rates			Completion rates	
Mover status[1]	Completes[2]	Weighted	Unweighted	Completes[3]	Weighted	Unweighted	Completes[4]	Weighted	Unweighted
All school types	16,593	88.0	91.8	15,522	84.5	85.8	14,619	76.3	81.4
Mover status									
Mover	1,519	63.1	64.1	1,756	73.5	74.1	778	35.3	34.6
Located, followed	1,519	85.5	85.1	1,397	78.2	78.2	778	48.7	46.7
Other[5]	0	0.0	0.0	359	60.4	61.5	0	0.0	0.0
Nonmover	15,074	95.9	95.9	13,766	88.0	87.6	13,841	88.7	88.1

[1] This is the mover status used in weighting, which does not consider children who moved into identified destination schools as movers.

[2] Reading, math or general knowledge assessment was scorable or child was language minority (not Spanish) or disabled.

[3] Family structure portion of Parent Interview was completed.

[4] A completed questionnaire was defined as one that was not completely left blank.

[5] This category includes movers who could not be located, movers whose cases could not be processed before the end of the field period, and movers who moved into nonsampled PSUs.

Table 5-15.—ECLS-K, spring-first grade: number of completed child-level cases and child-level completion rates, by child's mover status—children sampled in the base year (continued)

	Teacher questionnaire A			Teacher questionnaire B			Teacher questionnaire C		
		Completion rates			Completion rates			Completion rates	
Mover status[1]	Completes[4]	Weighted	Unweighted	Completes[4]	Weighted	Unweighted	Completes[4]	Weighted	Unweighted
All school types	15,021	78.1	83.6	14,878	77.5	82.8	14,991	78.0	83.5
Mover status									
Mover	856	38.5	38.0	844	38.1	37.5	858	38.5	38.1
Located, followed	856	53.2	51.4	844	52.6	50.7	858	53.2	51.5
Other[5]	0	0.0	0.0	0	0.0	0.0	0	0.0	0.0
Nonmover	14,165	90.1	90.2	14,034	89.5	89.3	14,133	90.0	90.0

[1] This is the mover status used in weighting, which does not consider children who moved into identified destination schools as movers.

[2] Reading, math, or general knowledge assessment was scorable or child was language minority (not Spanish) or disabled.

[3] Family structure portion of parent interview was completed.

[4] A completed questionnaire was defined as one that was not completely left blank.

[5] This category includes movers who could not be located, movers whose cases could not be processed before the end of the field period, and movers into nonsampled PSUs.

parent interview was conducted by telephone leading to the 60.4 percent response rate for this category. The SAQ showed a 40-point drop in completion rate for movers, even when they were located and followed. There are several reasons for this drop: 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 SAQ; 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 88.7 percent. For located and followed movers it was 48.7 percent. For all three teacher questionnaires the completion rates were approximately 90 percent if the child had not moved, about 53 percent if the child moved, was located, and followed, and 0 if not located or followed. The reasons for lower completion rates from teachers if the child had moved are similar to the reasons that affected the SAQ completion rates for movers.

Table 5-16 presents child-level completion rates for the spring-first grade data collection, this time broken out by child characteristics for children who were sampled as part of the kindergarten cohort in the base year. The differences in completion rates by gender and by year of birth are inconsequential but for race and ethnicity they are more substantial. For the child assessment the completion rate was highest for Asians and Pacific Islanders (90.7 percent and 89.6 percent respectively) and lowest for Native Americans (84.0 percent). For the parent interview it was highest for whites (88.6 percent), lowest for Asians (73.2 percent), and second lowest for black students (77.5 percent). The low rate for Asians is partly due to language problems; 7.3 percent (unweighted) of parent interviews for Asian children were nonresponse because of language. The ECLS-K sample of Pacific Islanders is very clustered and has unusually high completion rates for the instruments filled out by school personnel, 87.1 percent for the school administrator questionnaire and over 86 percent for each of the teacher questionnaires. The lowest completion rate for the school administrator questionnaire is for Native Americans (68.6 percent). For the teacher questionnaires the lowest rates are in the 73 to 75 percent range and are associated with blacks and Hispanics. Since almost 70 percent of the black and Hispanic students are enrolled in high minority schools (50 percent or higher), this may be associated with lower levels of response for the SAQ from high minority schools. Of the 21.6 percent (unweighted) of black and Hispanic students with no SAQ data, 20 percent are enrolled in high minority schools.

In addition to the child assessment, parent interview, teacher questionnaires, and school administrator questionnaires whose completion rates have been summarized in the preceding tables, various other types of data were collected during spring-first grade as well. Table 5-17 presents counts of completes and weighted and unweighted completion rates at the overall

Table 5-16.—ECLS-K spring-first grade: number of completed child-level cases and child-level completion rates, by child characteristic—children sampled in the base year

	Child assessment			Parent interview			School administrator questionnaire		
		Completion rates			Completion rates			Completion rates	
Child characteristics[1]	Completes[2]	Weighted	Unweighted	Completes[3]	Weighted	Unweighted	Completes[4]	Weighted	Unweighted
All school types	16,593	88.0	91.8	15,522	84.5	85.8	14,619	76.3	81.4
Gender									
Male	8,476	87.7	91.7	7,937	84.2	85.9	7,464	76.2	81.3
Female	8,106	88.3	91.8	7,585	84.9	85.9	7,144	76.3	81.4
Unknown gender	11	98.4	91.7	0	0.0	0.0	11	98.4	91.7
Race/ethnicity									
White (not Hispanic)	9,401	88.6	92.1	9,214	88.6	90.3	8,477	79.2	83.5
Black (not Hispanic)	2,349	87.3	90.5	2,051	77.5	79.0	2,023	72.5	78.7
Hispanic	2,838	87.2	91.3	2,559	81.0	82.3	2,391	72.0	77.7
Asian	1,071	90.7	93.5	838	73.2	73.1	919	76.3	80.6
Pacific Islander	188	89.6	93.1	165	81.9	81.7	184	87.1	91.5
American Indian or Alaska Native	296	84.0	89.2	288	82.9	86.7	247	68.6	75.5
Other	397	85.2	91.5	385	86.7	88.7	333	71.4	77.4
Unknown race/ethnicity	53	77.9	84.1	22	32.8	34.9	45	65.6	72.6
Year of birth									
1992	4,863	88.2	92.0	4,571	84.4	86.5	4,320	77.5	82.3
1993	11,616	87.9	91.7	10,856	84.7	85.7	10,197	75.8	81.0
Other/unknown	114	83.0	86.4	95	65.7	72.0	102	74.6	79.1

[1] Based on ECLS-K survey data and not data from the sampling frame.

[2] Reading, math or general knowledge assessment was scorable or child was language minority (not Spanish) or disabled.

[3] Family structure portion of parent interview was completed.

[4] A completed questionnaire was defined as one that was not completely left blank.

Table 5-16.—ECLS-K spring-first grade: number of completed child-level cases and child-level completion rates, by child characteristic—children sampled in the base year (continued)

	Teacher questionnaire A			Teacher questionnaire B			Teacher questionnaire C		
		Completion rates			Completion rates			Completion rates	
Child characteristics[1]	Completes[4]	Weighted	Unweighted	Completes[4]	Weighted	Unweighted	Completes[4]	Weighted	Unweighted
All school types	15,021	78.1	83.6	14,878	77.5	82.8	14,991	78.0	83.5
Gender									
Male	7,648	77.6	83.3	7,569	77.0	82.5	7,631	77.4	83.2
Female	7,373	78.7	84.0	7,309	78.1	83.3	7,360	78.7	83.9
Unknown gender	0	0.0	0.0	0	0.0	0.0	0	0.0	0.0
Race/ethnicity									
White (not Hispanic)	8,699	80.7	85.7	8,638	80.4	85.1	8,657	80.2	85.3
Black (not Hispanic)	2,037	73.6	79.3	2,027	72.7	78.9	2,075	74.7	80.8
Hispanic	2,480	75.3	80.6	2,451	74.6	79.6	2,465	75.0	80.1
Asian	963	79.5	84.5	927	76.1	81.3	951	79.1	83.4
Pacific Islander	182	86.1	90.5	182	86.1	90.5	180	86.4	89.6
American Indian or Alaska Native	283	78.4	86.5	283	78.4	86.5	283	79.2	86.5
Other	345	71.9	80.2	338	71.2	78.6	348	73.1	80.9
Unknown race/ethnicity	32	36.7	51.6	32	36.7	51.6	32	36.7	51.6
Year of birth									
1992	4,462	79.4	85.0	4,415	78.8	84.1	4,462	79.7	85.0
1993	10,466	77.7	83.2	10,370	77.1	82.4	10,441	77.4	83.0
Other/unknown	93	61.6	72.1	93	61.6	72.1	88	58.6	68.2

1 Based on ECLS-K survey data and not on data from the sampling frame.

2 Reading, math, or general knowledge assessment was scorable or child was language minority (not Spanish) or disabled.

3 Family structure portion of parent interview was completed.

4 A completed questionnaire was defined as one that was not completely left blank.

student level for these other data collection efforts. The facilities checklist has a 79.1 percent completion rate, which is about three points higher than that for the school administrator questionnaire, the only other school-level survey instrument. The student record abstract, which was to have been completed for all students except for those who moved and could not be found in a school, had a 72.2 percent completion rate. There are fewer special education teacher questionnaires overall but their completion rates are higher, 88.7 percent for part A, which captures teacher information, and 82.9 percent for part B, which relates to individual students who receive special education services. Finally there is the adaptive behavior scale (68.5 percent), which was collected only for students who were not assessed in spring-first grade because of physical or mental disability.

Table 5-17.—ECLS-K spring-first grade: number of completed instruments and child-level completion rates for additional data collected—children sampled in the base year

Category	Completes	Completion rates	
		Weighted	Unweighted
Facility checklist*	15,319	79.1	85.3
Student record abstract*	13,928	72.2	77.5
Special education—Part A*	693	88.7	88.5
Special education—Part B*	650	82.9	83.0
Adaptive behavior scale*	22	68.5	68.8

* A completed instrument was defined as one that was not completely left blank.

5.7.2 Students 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 did not attend kindergarten in the United States or were in first grade in the fall of 1998. (For a detailed description of the freshening procedure see section 4.3.2.) Nonresponse attrition 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 (first component) and then during data collection after the freshened students had been identified (second component). The first component alone can further be decomposed into two sources: attrition due to entire schools refusing to implement the freshening procedure (*the school term*), and attrition due to ECLS-K sample children moving to other schools (*the child term*). To contain costs, students who transferred from schools targeted for freshening were not used as links to identify freshened students, even when they were otherwise followed for data collection. These movers were considered freshening nonrespondents in the *child term*.

Table 5-18 presents weighted completion rates for freshened students. The two components of the completion rates are presented separately in table 5-18. The actual 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 65.4 percent. The reasons that schools did not participate in the freshening process included refusing or being unable to provide the requested information in order to complete the procedures. Within the schools that agreed to freshen, the freshening completion rate is 98.3 percent, the slight loss due to students who transferred to other schools (*the child term*). Multiplying these two terms together gives a first component completion rate of 64.3 percent. The second component varies by survey instrument. The rates for the paper-and-pencil instruments range from 78.2 percent for the student record abstract to 100 percent for the special education teacher—part A questionnaire and are uniformly higher than for the kindergarten sample. The child assessment at 84.0 percent is four points lower than for the kindergarten sample and the parent interview, at 64.9 percent, is nearly 20 points down. These figures demonstrate that, except for the parent interview, data collection went very well once the freshening procedure had been implemented. The final completion rate for each instrument is the product of the two components. Because of the poor showing at the sampling stage these range from a high of 64.3 percent for the special education teacher—part A questionnaire to 41.7 percent for the parent interview.

5.7.3 Spring-First Grade Completion Rates – All Children

Table 5-19 presents final spring-first 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 brings down the combined rate only one- or two-tenths of a percent relative to the children sampled in kindergarten rates, even though the children sampled in first grade rates are quite low.

Table 5-18.—ECLS-K spring-first grade: number of completed child-level cases and child-level completion rates for children sampled in first grade

Category	Completes	Completion rates[1]	
		Weighted	Unweighted
First component	7,135	64.3	77.9
School term[2]	7,192	65.4	78.5
Child term[3]	7,135	98.3	99.2
Second component			
Child assessment[4]	134	84.0	81.2
Parent interview[5]	104	64.9	63.0
Teacher questionnaire—Part A[6]	145	89.7	87.9
Teacher questionnaire—Part C[6]	144	88.9	87.3
Teacher questionnaire—Part B[6]	132	81.9	80.0
Special education—Part A[6]	15	100.0	100.0
Special education—Part B[6]	14	96.5	93.3
School administrator questionnaire[6]	145	91.8	87.9
Facility check list[6]	159	96.9	96.4
Student records abstract[6]	120	78.2	72.7
Completion rates			
Child assessment[4]	134	54.0	63.3
Parent interview[5]	104	41.7	49.1
Teacher questionnaire—Part A[6]	145	57.7	68.4
Teacher questionnaire—Part B[6]	144	57.2	68.0
Teacher questionnaire—Part C[6]	132	52.7	62.3
Special education—Part A[6]	15	64.3	77.9
Special education—Part B[6]	14	62.1	72.7
School administrator questionnaire[6]	145	59.0	68.4
Facility check list[6]	159	62.3	75.1
Student records abstract[6]	120	50.3	56.6

[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, math, or general knowledge assessment was scorable or child was language minority (not Spanish) or disabled.

[5] Family structure portion of parent interview was completed.

[6] A completed questionnaire was defined as one that was not completely left blank.

Table 5-19.—ECLS-K spring-first grade: number of completed child-level cases and child-level completion rates, for children sampled in kindergarten and first grade, by survey instruments

	Children sampled in kindergarten			Children sampled in first grade			All children		
		Completion rates			Completion rates			Completion rates	
Survey instrument	Completes	Weighted	Unweighted	Completes	Weighted	Unweighted	Completes	Weighted	Unweighted
Child assessment[1]	16,593	88.0	91.8	134	54.0	63.3	16,727	87.2	91.6
Parent interview[2]	15,522	84.5	85.8	104	41.7	49.1	15,626	83.5	85.6
Teacher questionnaire—Part A[3]	15,021	78.1	83.6	145	57.7	68.4	15,166	77.6	83.5
Teacher questionnaire—Part B[3]	14,878	77.5	82.8	144	57.2	68.0	15,022	77.0	82.7
Teacher questionnaire—Part C[3]	14,991	78.0	83.5	132	52.7	62.3	15,123	77.4	83.3
Special education—Part A[3]	693	88.7	88.5	15	64.3	77.9	708	88.1	88.4
Special education—Part B[3]	650	82.9	83.0	14	62.1	72.7	664	82.4	82.9
School administrator questionnaire[3]	14,619	76.3	81.4	145	59.0	68.4	14,764	75.9	81.3
Facility check list[3]	15,319	79.1	85.3	159	62.3	75.1	15,478	78.7	85.2
Student records abstract[3]	13,928	72.2	77.5	120	50.3	56.6	14,048	71.7	77.3
Adaptive behavior scale	22	68.5	68.8	-	-	-	22	68.5	68.8

[1] Reading, math, or general knowledge assessment was scorable or child was language minority (not Spanish) or disabled.

[2] Family structure portion of parent interview was completed.

[3] A completed questionnaire was defined as one that was not completely left blank.

6. DATA PREPARATION

As described in chapter 5, two types of data collection instruments were used for the Early Childhood Longitudinal Study-Kindergarten Class of 1998-99 (ECLS-K) data collection in the fall- and spring-first grade: computer-assisted and self-administered paper forms (hard copy). The data preparation approach differs with the mode of data collection. The direct child assessment and parent interview were conducted using computer-assisted interviewing (CAI) techniques. Editing specifications were built into the computer programs used by assessors to collect these data. The teacher and school administrator forms were self-administered. When the field supervisors returned these forms, coders recorded the receipt of these forms into a project-specific forms tracking system. Coders reviewed the questionnaires to ensure data readability for transfer into an electronic format. The visual review included changing (upcoding) any "Other, specify" responses that actually fit within the available response categories of the question. There were some items where upcoding was conducted after the data were keyed due to the large volume of other responses. Once they finished this review, the coders sent the instruments to data entry to be manually transferred to an electronic format and reviewed for range and logic consistency. The following sections describe the data preparation activities for both modes of data collection in more detail.

6.1 Coding and Editing Specifications for Computer-Assisted Interviweing

The very nature of designing a computer-assisted interview forces decisions about edit specifications to be made up front. Both acceptable ranges and logic consistency checks were pre-programmed into the electronic questionnaire. The next few sections describe the coding and editing of the data collected using CAI. Though the child assessments and the parent interviews were both collected using CAI, the child assessment did not contain some of the additional range and edit checks contained in the parent interview. The following sections describe the coding and editing that was conducted on the CAI parent interview.

6.1.1 Range Specifications

Within the CAI parent interview instruments, respondent answers were subjected to both "hard" and "soft" range edits during the interviewing process. (The child assessment did not have such hard and soft ranges.) A "soft range" is one that represents the reasonable expected range of values but does not include all possible values. Responses outside the soft range were confirmed with the respondent and entered a second time. For example, the number of hours each week a child attends a daycare center on a regular basis had a soft range of one to 50. A value outside this range could be entered and confirmed as correct by the assessor as long as it was within the hard range of values (one to 70).

"Hard ranges" are those that have a finite set of parameters for the values that can be entered into the computer, for example, "0-20 times" for the number of times the child was late for school in the past four months. Out-of-range values for closed-ended questions were not accepted. If the respondent insisted that a response outside the hard range was correct, the assessor could enter the information in a comments data file. Data preparation and project staff reviewed these comments. Out-of-range values were accepted and entered into the data file if the comments supported the response.

6.1.2 Consistency Checks (Logical Edits)

Consistency checks, or logical edits, examine the relationship between responses to ensure that they do not conflict with one another or that the response to one item does not make the response to another item unlikely. For example, in the household roster, one could not be recorded as a mother and male. When a logical error such as this occurred during a session, the assessor saw a message requesting verification of the last response and a resolution of the discrepancy. In some instances, if the verified response still resulted in a logical error, the assessor recorded the problem either in a comment or on a problem report. Consistency checks were not applicable to the child assessments.

6.1.3 Coding

Additional coding was required for some of the items collected in the CAI instruments. These items included "Other, specify" text responses, occupation, race-ethnicity, and home language. Assessors keyed verbatim responses to these items. Once the data were keyed, staff were trained to code these data using coding manuals designed by Westat and the National Center for Education Statistics (NCES) to support the coding process. In this section, we describe the coding activities for the CAI instruments.

Review of "Other, specify" Items

The "Other, specify" open-ended parent interview responses were reviewed to determine if they should be coded into one of the existing response categories. During the data collection, when a respondent selected an "other" response in the parent interview, the assessor entered the text into a "specify" overlay that appeared on the screen. The data preparation staff reviewed these text "specify" responses and, where appropriate, coded them into one of the existing response categories. If a text "specify" response for which there was no existing response category occurred frequently enough, new codes (response categories) would have been added. No new codes were added for first grade parent data. There were no "Other, specify" items in the child assessments.

Parent Occupation Coding

As in the base year, occupations were coded using the *Manual for Coding Industries and Occupations*, March 1999 (National Household Education Survey, NHES: 99). This coding manual was created for NHES and used an aggregated version of industry and occupation codes. The industry and occupation codes used by NHES were originally developed for the National Postsecondary Student Aid Study (NPSAS, 1990) and contained one to four digits. Analysis of the NPSAS categories revealed that some categories had very small numbers of cases and some categories that are similar in industry or occupation had similar participation rates, suggesting that the separate codes could be collapsed without significant loss of information. The NHES industry and occupation code categories use a two-digit code, the highest level of aggregation, to have sufficient numbers of cases to support analysis without collapsing categories. There are 13 industry codes and 22 occupation codes in the NHES coding scheme. If an industry or occupation could not be coded using this manual, the *Index of Industries and Occupations—1980* and *Standard Occupational Classification Manual—1980* were used. Both of these manuals use an expanded coding system and at the same time are directly related to the much more condensed NHES coding scheme. The 1980 manuals were used for reference in cases where the NHES coding scheme did not adequately cover a particular situation. (See Chapter 7, section 7.4 for an expanded description of the industry and occupation codes.*)

Occupation coding began with an autocoding procedure using a computer string match program developed for the NHES. The program searched the responses for strings of text for each record/case and assigned an appropriate code. About 25 percent of the cases were autocoded and were not verified because there was an exact match between the respondent's answer and the occupation code.

Cases that could not be coded using the autocoding system were coded manually by coders using a customized coding utility program designed for coding occupations. The customized coding utility program brought up each case for coders to assign the most appropriate codes. In addition to the text strings, other information, such as main duties, highest level of education, and name of the employer, was available for the coders. The coders used this information to ensure that the occupation code assigned to each case was appropriate.

Verification of coding is an important tool to assure quality control and as an extension of coder training. As a verification step, two coders independently assigned codes (i.e., double-blind coding) to industry and occupation cases. A coding supervisor arbitrated disagreements between the two codes, the initial code and the verification code. In the early stages of each coder's work, 100 percent of each coder's work was reviewed. Once the coder's error rate had dropped to one percent or less, ten percent of the coder's work was reviewed.

Race-Ethnicity Coding

The same coding rules used in the base year were used to code all race-ethnicity variables for children, resident parents, and nonresident parents. See chapter 7, section 7.4.1 for details on how the race variables were coded and how the race-ethnicity composite was created.

---

* Office of Budget and Management, Executive Office of the President (1980). *Standard Industrial Classification Manual.* Springfield, VA, and Office of Federal Statistical Policy and Standards, U.S. Department of Commerce (1980). *Standard Occupational Classification Manual* (2nd ed.). Washington, DC: Superintendent of Documents, U.S. Government Printing Office.

Language Coding

The same coding rules used in the base year were used to code home languages reported by respondents.

Partially Complete Parent Interviews

A "completed" parent instrument was defined by whether the section on family structure (FSQ) was completed by the respondent. Only completed interviews were retained in the final data file. A small number of interviews in each round, approximately five (less than one percent) in fall-first grade and 113 (less than one percent) in spring-first grade, terminated the parent interview after the FSQ section but before the end of the instrument. These interviews were defined as "partially complete" cases and were included in the data file. All instrument items after the interview termination point were set to -9 for "not ascertained."

Household Roster in the Parent Interview

Several tests were run on the household roster to look for missing or inaccurate information. These tests are the same tests run on the kindergarten year data; the difference in the first grade files is the absence of detailed edit flags. Rather than carrying seven different flags describing each type of edit performed on a particular case, as was done in the kindergarten year, one flag was used per round to identify cases that were edited for any of the reasons described below. These flags are P3EDIT and P4EDIT; the flag is set to 1 if the case was edited in the given round. There were 212 cases requiring edits in the round 3 roster data, and 226 cases requiring edits in the round 4 data.

There were essentially three general types of roster tests performed to determine which cases required editing. First, the relationship of an individual to the focal child was compared to the individual's listed age and gender. Any problems found were corrected based on data from prior data collections wherever possible. Second, households with more than one mother or more than one father were scrutinized for errors. While it is possible to have more than one mother in a household—for example, a household could contain one biological and one foster mother of the focal child—such cases warranted closer inspection. There were 37 cases with more than one mother or father in the household. Corrections were made wherever clear errors and a clear resolution existed. Lastly, the relationship of an individual to both the focal child and the reference person was examined, as there are cases in which the relationship of an individual to the focal child conflicts with his status as the spouse/partner of the reference person. For example, in a household containing a child's grandparents but not his or her parents, we may designate the grandmother as the "mother" figure, and the grandfather thus becomes the "father" by virtue of his marriage to the grandmother. These cases were examined but left unchanged. Both the original—and correct (grandfather)—relationship data and the new "parent-figure" designation (father) that had been constructed were kept. In fall-first grade, four cases remained with errors after these tests were run; the cases can be identified by the flag "P3ERRFLG". In spring-first grade, 82 cases remained in error after these tests were run; the cases can be identified by the flag "P4ERRFLG".

6.2 Coding and Editing Specifications for Hard-Copy Questionnaires

6.2.1 Receipt Control

In order to monitor the more than 45,000 documents that were to be received in the first grade year, the project-specific receipt and document control system developed in the base year was used. The receipt and document control system was initially loaded with the identifying information, such as identification numbers for schools, teachers, and children; the links between teachers and children; and the questionnaires that were expected from each school and teacher, for each cooperating school in the sample. As data were collected in the field, field supervisors completed transmittal forms for each school to indicate which questionnaires were being mailed to the home office. Once data collection started, receipt control clerks compared the transmittal forms to the questionnaires sent in from the field for accuracy and completeness. The identification number on each form was matched against the identification numbers in the tracking system to verify that the appropriate number of forms for each school was returned. The forms were then logged into the receipt and document control system. Once forms were logged in, if they had any data (some forms had no data due to refusal by the respondent to complete them), they were then coded; the data were entered into electronic format, after which the data were edited. The following sections describe the coding, data entry, and editing processes for hard-copy questionnaires.

6.2.2 Coding

The hard-copy questionnaires required coding of race-ethnicity for teachers, review of "Other, specify" text responses, and a quick visual review of particular questions in each questionnaire. The quick visual review was to assure that the questionnaire values were accurate and complete and were consistent across variables and that the numbers were converted to the appropriate unit of measurement prior to converting data to an electronic format. The coding staff were trained on the coding procedures and had coding manuals to support the coding process. This staff also did the data editing after data entry was complete. Senior coders verified coding. The verification rate was set at 100 percent for each coder until accuracy of less than one percent error rate was established. After that point, work was reviewed at a rate of ten percent.

Review of "Other, specify" Items

The "Other, specify" text responses were reviewed by the data editing staff and, where appropriate, upcoded into one of the existing response categories. The small number of specify responses that remained after upcoding did not fit into any preexisting category and were of insufficient numbers to warrant an additional category. No new codes were added.

Coding Teacher Race-Ethnicity

"Other, specify" text responses for race-ethnicity in the teacher questionnaire part B were coded using the base year procedures.

Coding Teacher Language

"Other, specify" text responses for language in the teacher questionnaire part A were coded using the base year procedures.

6.2.3 Data Entry

Westat data entry staff keyed the forms in each batch. The data were rekeyed by more senior data entry operators at a rate of 100 percent to verify the data entry. The results of the two data entry passes were compared and differences identified. The hard-copy form was pulled and examined to determine what corrections had to be made to the keyed data. These corrections were rekeyed, resulting in an accuracy rate exceeding 99 percent. The verified batches were then transmitted electronically to Westat's computer system for data editing.

6.2.4 Data Editing

The data editing process consisted of running range edits for soft and hard ranges, running consistency edits, and reviewing frequencies of the results.

Range Specifications

Hard-copy range specifications set the parameters for high and low acceptable values for a question. Where values were printed on the forms, these were used as the range parameters. For open-ended questions, such as, "Counting this school year, how many years have you taught in your **current school** including part-time teaching?" high and low ranges were established as acceptable values. Data frequencies were run on the range of values to identify any errors. Values outside the range were identified as an error and were printed on hard copy for a data editor to review. Cases identified with range errors were identified, and the original response was updated. Data frequencies were then rerun and reviewed. This iterative process was repeated until no further range errors were found.

Consistency Checks (Logical Edits)

By programming logical edits between variables, consistency between variables not involved in a skip pattern was confirmed. For example, in the school administrator questionnaire, the number of children eligible for free breakfast could not exceed the total number of children enrolled in the school. These logical edits were run on the whole database after all data entry and range edits were complete. The logical edits were run separately for each form. All batches of data were combined into one large data file, and data frequencies were produced. The frequencies were reviewed to ensure the data remained logically consistent within the form. When an inconsistency was found, the case was identified and the inconsistency was printed on paper for an editor to review. The original value was corrected, and the case was then rerun through the consistency edits. Once the case passed the consistency edits, it was appended back into the main data set. The frequencies were then rerun and reviewed. This was an iterative process; it was repeated until no further inconsistencies were found.

Frequency and Cross-Tabulation Review

Frequencies and cross-tabulations were run to determine consistency and accuracy across the various forms and matched against the data in the field management system. If discrepancies could not be explained, no changes were made to the data. For example, in teacher questionnaire part A, an item asking about languages other than English spoken in the classroom includes a response option of "No language other than English." If a respondent circled that response but also answered (in subsequent items) that other languages besides English were spoken in the classroom, then the response was left as recorded by the respondent because the discrepancy could not be resolved. In the student record abstract, for item 13 a code was added, "Both option 1 and 2 are circled" as in 99 cases the respondent indicated both options.

7. DATA FILE CONTENT AND COMPOSITE VARIABLES

This chapter describes the content of the Early Childhood Longitudinal Study-Kindergarten Class of 1998-99 (ECLS-K) First Grade Public-Use Data File and focuses largely on the composite variables that have been created. The file can be used for longitudinal analysis in combination with the file from the base year (kindergarten year) (see chapter 9 for details about longitudinal analyses). The Base Year User's Manual is included in the first grade Electronic Code Book (ECB) with the First Grade User's Manual for reference.

Unlike the ECLS-K files for the base year for which there were child, teacher, and school "catalogs," there is only one child-level first grade data file or catalog, as explained in chapter 1. This file was constructed at the child level, and each child record contains data from the various respondents associated with the child (the child herself/himself, a parent, one or more teachers, and a school administrator, the facilities checklist, school records) as well as from the field management system (FMS).

The first grade child catalog contains one record for each of the 17,487 responding students in fall- and spring-first grade. (There were 5,424 responding students in fall-first grade, 17,324 in spring-first grade, and 17,487 responding students in the combined file for first grade.) First grade school and teacher-level data, including composites, are also stored on the child catalog. The file, named child1p.dat, is stored in the root directory of the CD-ROM as an ASCII file. However, it is strongly recommended that users access the data using the 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.

There is also a school catalog on the file that is from the base year of the study. This is included in order to allow analysis of data in conjunction with the base year file and to provide school data that were not collected again in the first grade (details about these data are provided in section 7.7).

This chapter is divided into nine sections. Sections 7.1 through 7.3 focus on the conventions used in the study and describe identification variables, missing values, and variable names. Section 7.4 provides details about the creation of composite variables on the first grade file. Composites created from the base year data that are made available on the first grade file are also described. Sections 7.5 through 7.7 focus on issues that users may want to know for particular analyses—the relationship among composites on the child catalog, variables to identify children who changed schools, and merging base year school level data with first grade child level data. Section 7.8 contains tables of the composite variables. Finally, Section 7.9 describes masked variables.

7.1 Identification Variables

The first grade data file contains a child identification (ID) variable (CHILDID) that uniquely identifies each record. Teachers on the child records are identified with the ID variable T4_ID; schools are identified by the ID variables S3_ID (fall-first grade) and S4_ID (spring-first grade). The ID variables S1_ID and S2_ID are also included on the file to identify children who changed schools. These school identifiers indicate which school the child attended during each round of data collection, e.g. S3_ID indicates the school the child attended during the fall-first grade data collection. Section 7.6 provides further details on school identifiers.

Each type of respondent (child, parent, regular teacher, special education teacher and school) has a unique ID number. The original school ID number (S_ID) is the base for all the subsequent ID numbers as children, parents, and teachers were sampled from schools during the base year. The school ID number is a four-digit number assigned sequentially to sampled schools. The number has a series of ranges: 0001-1299 for originally sampled schools; 2000 series for new schools added to the sample during the sample freshening process; 3000 series for substitute schools that replaced nonresponding original sample schools; and 4000 through 6000 series for transfer schools, which were assigned during processing at the home office. (See chapter 4 for a complete description of the ECLS-K sample.) There was also a 9000 series of S_ID numbers that referred either to children who did not attend regular school because they were schooled at home or to children for whom information about the current school was not in the ECLS-K database, but it was thought that the parent could be located. All children who fit one of these conditions were assigned 9000 series ID's according to work area numbers (e.g., the S_ID was 9 + the work area number. In work area 101, the S_ID was 9101). There are also several specific 9000 series codes for children who were not located or not followed at the end of a round. The school ID numbers start with 999 for these cases. These are described in section 7.6.

The child ID number (CHILDID) is a concatenation of the school ID where the child was sampled, a three-digit student number and the letter "C." For example, 0001001C is the ID number of the first child sampled in school 0001. The teacher ID number (T4_ID) is a concatenation of the school ID where the teacher was sampled, the letter "T," and a two-digit teacher number. For example, 0001T01 is the ID number for the first teacher sampled in school 0001. The parent ID number (PARENTID) is linked to the child ID number and is a concatenation of the four digit school ID, the three digit student number, and the letter "P." It is the same number as the child ID with a letter "P" instead of a letter "C" at the end. For example, 0001001P is the ID number of the parent of the first child sampled in school 0001. If twins are sampled, the ID of the first child sampled is used to generate the parent ID. For twins, there will be two child-level records with the same parent ID. Children with the same teacher can be identified by finding all children on the child file with the same teacher ID.

It should be noted that there is a difference in the variable names between the base year and first grade special education teacher IDs. In the base year of the study information from special education teachers was included in a separate file and was not part of the child or teacher catalogs. The ID number for special education teachers in the base year special education file was T_ID. In the first grade file, the special education teacher information is included with the rest of the first grade data, thus ID numbers were needed to distinguish special education teachers from regular education teachers. In the first grade file, T4_ID is used to identify regular education teachers and D4T_ID is used to identify special education teachers.

7.2 Missing Values

All variables on these files use a standard scheme for missing values. Codes are used to indicate item nonresponse, legitimate skips, and unit nonresponse.

-1	Not Applicable, including legitimate skips
-7	Refused (a type of item nonresponse)
-8	Don't Know (a type of item nonresponse)
-9	Not Ascertained (a type of item nonresponse)
(blank)	System Missing, including unit nonresponse

The "Not Applicable" code (-1) has two purposes. Its primary purpose is to indicate that a respondent did not answer the question due to skip instructions within the instrument or external reasons that led a respondent not to participate. In the parent interview, where the parent or guardian was a respondent, a “Not Applicable” is coded for questions that were not asked of the respondent because of a previous answer given. For example, a question about a sibling’s age is not asked when the respondent has indicated that the child has no siblings. A “Not Applicable” code is also used in the direct child assessment if a child did not participate in any section due to language or a disability. For the teacher and school files where the instruments are self-administered, a “Not Applicable” is coded for questions that the respondent left blank because the written directions instructed them to skip the question due to a certain response on a previous question.

Another use of the “Not Applicable” code is when it is not known whether a respondent would have answered a question series following a lead question. One example of this use of “Not Applicable” is in the Facility Checklist. There are several skip patterns in this questionnaire that result in the use of “Not Applicable” codes. To illustrate, question 1a (whether a particular facility is available) is a lead question that skips to question 1b (whether a particular facility was observed) if the facility is available. Question 1b is a lead question that skips to question 2 if a facility is observed. A “Not Applicable” for question 1b (whether a particular facility was observed) means that the facility was not available or that the observer failed to answer whether the facility was available. A “Not Applicable” for question 2 (observer rating of the facility) or question 3 (handicap accessibility) means that the specific facility was not available, that the observer did not answer whether it was available, the facility was available but not observed, or that the observer failed to answer whether it was observed. In these cases, the “Not Applicable” code does not simply mean that the facility was not available at the school.

Another example of the use of the “Not Applicable” code is illustrated in the skip pattern link between question 4 (number of portable classrooms) in the Facility Checklist and questions 1a and 1b about whether or not the observer reported available or observed portable classrooms. If an observer failed to answer in question 1a(m) whether portable classrooms were available, the number of portable classrooms in question 4 was coded “Not Applicable.” Also, if an observer answered “yes” that there were portable classrooms in question 1a(m) but then did not answer whether he/she observed them in question 1b(m,) question 4 about the number of portable classrooms was also coded as “Not Applicable.”

The “Refused” code (-7) indicates that the respondent specifically told the interviewer that he or she would not answer the question. This, along with the “Don’t Know” code and the “Not Ascertained” code, indicates item nonresponse. The “Refused” code rarely appears in the school and teacher data because it indicates the respondent specifically wrote something on the questionnaire indicating an unwillingness to answer the question.

The "Don't Know" code (-8) indicates that the respondent specifically told the interviewer that he or she does not know the answer to the question (or in rare cases on the self-administered questionnaires, "I don't know" was written in for the question). For questions where "Don't Know" is one of the options explicitly provided, a "-8" will not be coded for those that choose this option; instead the "Don't Know" response will be coded as indicated in the value label information for that question.

The "Not Ascertained" code (-9) indicates that the respondent left the question blank that he or she should have answered. For the school and teacher self-administered questionnaires, this is the primary code for item nonresponse. For data outside the self-administered questionnaires (e.g., direct assessment scores), a "-9" means that a value was not ascertained or could not be calculated due to nonresponse.

"System Missing" appears as a blank when viewing code book frequencies and in the ASCII data files. System missing codes (blanks) in the first grade files have one of two meanings—either that an entire instrument or assessment is missing due to unit nonresponse or that some questions were not asked in one form of a questionnaire for returning schools but were asked in another form of a questionnaire for new schools (this applies only to the school administrator questionnaire). An example of the first meaning of system missing is a child's parent not participating in the parent interview. In this case, all questions from the parent interview will be blank (system missing). These may be translated to another value when the data are extracted into specific processing packages. For instance, SAS will translate these blanks into periods (".") for numeric variables.

The second meaning of a system missing code occurs only in the school administrator questionnaire. School administrator questionnaires were distributed in spring-kindergarten and spring-first grade. In order to reduce burden on the school administrator, if it was the same school for both time periods (a "returning" school), questions that were not expected to change (e.g., the grade levels included in the school) were not asked again in spring-first grade. However, if the school had not participated in spring-kindergarten (a "new" school), all questions were asked. Those questions that were asked in new schools but not in returning schools appear as system missing in returning schools. Details about which questions were not re-asked in returning schools are provided in section 7.7.

Depending on the research question being addressed, cases with missing values may need to be recoded. Because the missing value codes were used somewhat differently in different instruments, it is advised that users cross-tabulate all lead questions (e.g., whether the child received child care from a relative) and follow-up questions (e.g., hours of child care from a relative) before proceeding with any recodes or use of the data.

Missing values for composite variables were coded using the same general coding rules as those used for other variables. If a particular composite was inappropriate for a given household—as the variable P4MOMID was for a household with no resident mother—that variable was given a value of "-1" (Not Applicable). In instances where a variable was appropriate, but complete information to construct the composite was not available, the composite was given a value of –9 (Not Ascertained). The "Refused" and "Don't Know" codes were not used for the composites, except in the calculations of the height, weight, and body mass index (BMI) composites for fall-first grade and spring-first grade.[1]

The ECLS-K First Grade Public-Use Data Files are provided on a CD-ROM and are accessible through an ECB that allows data users to view variable frequencies, tag variables for extraction, and create the SAS, SPSS for Windows, or STATA code needed to create an extract each file for analysis. The two data files on the ECB—school and child—are each referred to as a "catalog." Instructions for using the CD-ROM and ECB are provided in chapter 8.

7.3 Variable Naming Conventions

Variables were named according to the data source (e.g., parent interview, teacher questionnaire) and the data collection point (i.e., a number is used to indicate in which round of data collection the variable was obtained, as follows: 1 for fall-kindergarten, 2 for spring-kindergarten, 3 for fall-first grade, and 4 for spring-first grade). These variable names are used consistently throughout the catalog. The prefixes listed here are in two categories: (1) first grade variables, and (2) kindergarten year and cross-round variables. In general, variable names start with the following prefixes:

---

[1] Children's height and weight measurements were each taken twice to prevent error and provide an accurate reading. Children's BMI was calculated based on height and weight. The rules for using "Don't Know" and "Not Ascertained" codes for these values was as follows. If both the first and second measurement of height in the child assessment were coded as -8 (Don't Know), then the height composite was coded as -8 (Don't Know). If both the first and second measurements of weight were coded as -8 (Don't Know), the weight composite was coded as -8 (Don't Know). If either the height or weight composites were coded as not ascertained (-9), the BMI composite was coded as not ascertained (-9). If neither the height nor weight composites were coded as not ascertained, and either the height or weight composite was coded as -8 (Don't Know), then the BMI composite was coded as -8 (Don't Know).

First Grade Variables

A4	Data collected/derived from spring-first grade teacher questionnaire A
B4	Data collected/derived from spring-first grade teacher questionnaire B
C3	Data/scores collected/derived from fall-first grade direct child assessment and fall-first grade weight variables
C4	Data/scores collected/derived from spring first grade direct child assessment and spring-first grade weight variables
D4	Data collected from spring-first grade special education teacher questionnaire A
E4	Data collected from spring-first grade special education teacher questionnaire B
F3	Data from fall-first grade FMS
F4	Data from spring-first grade FMS
IF	Imputation flags
K4	Data collected/derived from spring-first grade school facility checklist
P3	Data/scores collected/derived from fall-first grade parent interview
P4	Data/scores collected/derived from spring-first grade parent interview
R3	Derived child demographic or child status variables for fall-first grade
R4	Derived child demographic or child status variables for spring-first grade
S4	Data collected/derived from spring-first grade school administrator questionnaire
T4	Data/scores collected/derived from spring-first grade teacher questionnaire C
U4	Data collected/derived from spring-first grade student record abstract
W1	First grade (cross-round) parent composite variables

Kindergarten and Cross-round Variables

C23C	Child-level panel weight variable from spring-kindergarten and fall-first grade
C123C	Child-level panel weight variable from fall-kindergarten, spring-kindergarten, and fall-first grade

C24C	Child-level panel weight variable from spring-kindergarten and spring-first grade
C34C	Child-level panel weight variable from fall-first grade and spring-first grade
C124C	Child-level panel weight variable from fall-kindergarten, spring-kindergarten, and spring-first grade
C1_4C	Child-level panel weight variable from fall-kindergarten, spring-kindergarten, fall-first grade, and spring-first grade
C23P	Child-level panel weights for parent data from spring-kindergarten and fall-first grade
C123P	Child-level panel weights for parent data from fall-kindergarten, spring-kindergarten, and fall-first grade (see chapter 9 for details)
C24P	Child-level panel weights for parent data from spring-kindergarten and spring-first grade
C34P	Child-level panel weights for parent data from fall-first grade and spring-first grade
C124P	Child-level panel weights for parent data from fall-kindergarten, spring-kindergarten, and spring-first grade (see chapter 9 for details)
C1_4P	Child-level panel weights for parent data from fall-kindergarten, spring-kindergarten, fall-first grade, and spring-first grade
F2	Data from spring-kindergarten FMS
HS	Head Start variables from fall-kindergarten
K2	Data from spring-kindergarten facilities checklist
P2	Data/scores collected/derived from spring-kindergarten parent interview
S2	Data collected/derived from spring-kindergarten school administrator questionnaire
U2	Data collected/derived from spring-kindergarten student record abstract
WK	Base year parent composite variables
Y2COM	Child-level panel weight that is nonzero for children whose C124W0 is nonzero and the base year combined weight BYCOMW0 is nonzero and whose parent completed the FSQ block in the parent questionnaire or whose teacher completed the TQB in the teacher questionnaire in the spring of first grade.

A few exceptions that do not follow the previously mentioned prefix convention are as follows:

- The identifiers CHILDID, PARENTID, T4_ID, S1_ID, S2_ID, S3_ID, and S4_ID.
- Some composite variables. In general, all composites derived from a given source maintain the same prefix as the source. Some composite variables, however, combined information across data collection points and/or several sources and are not associated with any prefixes. Derived child demographic variables, gender, race-ethnicity, and date of birth were created from the best source of data and are named GENDER, RACE, DOBMM, DOBDD, and DOBYY. Other such derived variables include R4R3SCHG, R4R2SCHG, and R4R2TCHG. These variables indicate change in school and teacher between the rounds indicated in the prefix (e.g., "R4R2" indicates a change between spring-first grade and spring-kindergarten). Sources and other details for these and all other composite variables can be found in tables 7-13 (fall-first grade composite variables) and 7-14 (spring-first grade composite variables).

7.4 Composite Variables

To facilitate analysis of the survey data, composite variables were created and added to the child data file. Most composite variables were created using two or more variables, each of which is named in the text that explains the composite variable. Other composite variables are recodes of single variables. Variables based on the child assessment include height, weight, and BMI. Variables based on the teacher data include class type (e.g., AM, PM, or all-day kindergarten class, or nonkindergarten class), teacher age, and student grade level. Variables constructed from the school data include the percentage of minority students, school type, and school instructional level. Variables constructed from the parent interview data include parent identifiers, parent demographics, household composition, household income and poverty, child care, and child demographics. Certain composites were created using data from the Field Management System (FMS).

Tables 7-13 and 7-14 list all the composite variables for fall-first grade and spring-first grade, respectively. Although there were far fewer composites in fall-first grade than in spring-first grade, the general order of both tables is the same (with some areas relevant only to spring-first grade.) All basic child demographic items are listed first. Child care variables follow the demographics and then household composition. For spring-first grade, imputed variables are listed next, then demographics for parents (resident father and mother characteristics are followed by characteristics of nonresident biological parents and nonresident adoptive parents). Teacher, classroom, and school variables are listed last. Once the user identifies the composites of interest, he or she can refer to table 8-8 for instructions on accessing the variables from the ECB.

7.4.1 Child Composite Variables

There are many child-level composite variables on the child catalog. Tables 7-13 and 7-14 describe all of the composites. Some of the composites are described in further detail here.

Child's Age at Assessment (R3_AGE and R4_AGE)

The child's age was calculated by determining the number of days between the date when the child completed the ECLS-K direct child assessment (e.g., the spring-first grade source variables were C4ASMTMM, C4ASMTDD, C4ASMTYY) and the child's date of birth (DOBMO, DOBDA, DOBYR). The total number of days was then divided by 30 to calculate the age in months. The child assessment date was tested for the appropriate range (September to November 1999 for round 3; March to July 2000 for spring-first grade). If the assessment date fell outside these ranges, the modal assessment date for the child's school was used.

Gender (GENDER)

The gender composite was derived using the gender indicated in the parent interview (CHILDGEN) and, if it was missing, the FMS. (The composite for GENDER is on the file but not the source variable, CHILDGEN.) Gender was only asked in the fall and spring of first grade if the household was new to the study. For most of the cases the data were collected in the base year.

Child's Date of Birth (DOBYY, DOBMM, and DOBDD)

The child date of birth composites were created using parent interview data (CHILDDOB) and, in cases in which the parent interview data did not exist or were outside of the criteria for inclusion, using the FMS data. If the date of birth given was before June 1, 1990, or after March 31, 1995, the data were excluded. This composite was updated from the base year, with data collected from households that were new to the study in the fall and spring of first grade.

Race-Ethnicity (W1AMERIN, W1ASIAN, W1PACISL, W1BLACK, W1WHITE, W1HISP, W1MT1RAC, W1RACETH, and RACE)

The composites for the child's race-ethnicity are presented in the ECLS-K files in three ways—(1) as dichotomous variables for each race-ethnicity category (W1AMERIN, W1ASIAN, W1PACISL W1BLACK, W1WHITE, W1HISP, W1MT1RAC); (2) as a single race-ethnicity composite taken from the parent interview data (W1RACETH); and (3) as a race-ethnicity composite taken from both the parent data and the FMS, if the parent data were missing (RACE). If the child's race-ethnicity information was available from a prior data collection, this value was used and copied forward. If the data were missing for a child in the kindergarten year, they were collected in first grade.

Respondents were allowed to indicate that they belonged to more than one of the five race categories (white, black or African American, American Indian or Alaska Native, Asian, Native Hawaiian or other Pacific Islander). From these responses, a series of five dichotomous race variables were created that indicated separately whether the respondent belonged to each of the five specified race groups. In addition, one more dichotomous variable was created for those who had simply indicated that they were multiracial without specifying a race (e.g., biracial). The retention of the dichotomous variables on the file allows users to create different composites as needed.

Data were collected on ethnicity as well. Respondents were asked if they were Hispanic or not. Using the six race dichotomous variables and the Hispanic ethnicity variable (P4HSP_1 to P4HSP_25, depending on household size), the race-ethnicity composite variables (W1RACETH and RACE) were created. The categories were: white, non-Hispanic; black or African American, non-Hispanic; Hispanic, race specified; Hispanic, no race specified; Asian; Native Hawaiian or other Pacific Islander; American Indian or Alaska Native, and more than one race specified, non-Hispanic. The child composites W1RACETH (race-ethnicity) and RACE (race-ethnicity) both share these categories; however, FMS data were used to fill in missing parent report data for the variable RACE and only parent report data were used for the variable W1RACETH. A child was classified as Hispanic if a respondent indicated the child's ethnicity was Hispanic regardless of whether a race was identified and what that race was.

It should be noted that there were additional race and ethnicity data collected for children in the spring of first grade. These questions ask whether the child is a member of a Spanish/Hispanic/Latino group, and if so, what group (FSQ.210 and FSQ.215); if the child is a member of an Asian group, and if so, what group (FSQ.220 and FSQ.225); and if the child is a member of a Pacific Islander group, and if so, what group (FSQ.230 and FSQ.235). These variables are available to users to make race and ethnicity distinctions, but they were not incorporated into the composite variables.

Child's Height (C3HEIGHT and C4HEIGHT)

Children's height was measured twice at each data collection point. For each height composite (C3HEIGHT and C4HEIGHT), if the two height values from the instrument (e.g., C4HGT1 and C4HGT2 for spring-first grade) were less than two inches apart, the average of the two height values was computed and used as the composite value. Otherwise, the value that was closest to 45.5 inches (the median height for six-year-olds as developed by the National Center for Health Statistics (NCHS) in collaboration with the National Center for Chronic Disease Prevention and Health Prevention [NCCDPHP]) was used as the composite value.

Child's Weight (C3WEIGHT and C4WEIGHT)

Children's weight was also measured twice at each data collection point. For each weight composite (C3WEIGHT and C4WEIGHT), if the two weight values from the instrument (e.g., C4WGT1 and C4WGT2 for spring-first grade) were less than five pounds apart, the average of the two values was computed and used as the composite value. Otherwise, the value that was closest to 45.5 pounds (the average weight for six-year-olds as developed by NCHS in collaboration with the NCCDPHP) was used as the composite value.

Child's Body Mass Index (BMI) (C3BMI and C4BMI)

Composite BMI (C3BMI and C4BMI) was calculated by multiplying the composite weight in pounds by 703.0696261393 and dividing by the square of the child's composite height in inches.

Child's Disability Status (P4DISABL)

A composite variable was created to indicate whether a child had a disability that was diagnosed by a professional. Questions in the parent interview about disabilities in spring-first grade asked about the child's ability to pay attention and learn, overall activity level, ability to communicate, difficulty hearing and understanding speech, and eyesight. (In fall-kindergarten, parents were also asked about their children's coordination in using their arms or legs and diagnosis of problems with use of limbs. These questions were not included in spring-first grade, thus the composite variable is slightly different than the kindergarten version.) For each disability, a question was asked about whether a diagnosis of a problem was obtained by a professional (CHQ050, CHQ110, CHQ170, CHQ210, CHQ300). A question was also asked about receipt of therapy services or participation in a program for children with disabilities (CHQ340).

The composite variable P4DISABL was coded 1 (Yes) if any of the source variables about diagnosis or therapy services were coded 1 (Yes) (CHQ050, CHQ110, CHQ170, CHQ210, CHQ300, CHQ340). This was done even if data for some of the source variables were missing. If data for all the source variables were missing, the composite was coded -9 (Not Ascertained). Otherwise, P4DISABL was coded 2 to indicate no reported disability.

Nonparental Care (P3CARNOW and P4CARNOW)

There are several composite variables on the file that can be used to describe children's child care arrangements based on information from the parent interview. Two of these (P3CARNOW and P4CARNOW) describe whether the child had any type of nonparental care during the reference period. The variables P3CARNOW and P4CARNOW were created similarly; however, P3CARNOW refers to any regular nonparental care during the summer before first grade and P4CARNOW refers to any nonparental care during spring-first grade. The creation of P4CARNOW was as follows (item references are only to spring-first grade for brevity, but the logic is the same for both variables). If the child was receiving care from a relative (CCQ010), a nonrelative (CCQ150), or a day care center or before or after school program at a school or in a center (CCQ260), P4CARNOW was equal to 1 (Yes). Otherwise, if any of the three variables was unknown, P4CARNOW was coded as -9 (Not Ascertained). If the respondent indicated that the child was not currently receiving any of the three types of care (CCQ010, CCQ150, and CCQ120 all equal 2 [No]), P4CARNOW was coded as 2 (No).

Hours per Week in Child Care (P3HRSNOW and P4HRSNOW)

Another set of child care composites indicates the number of hours per week the child spent in child care. The variable P3HRSNOW indicates the total number of hours per week spent in all primary, nonparental child care the summer before first grade, and P4HRSNOW indicates the total number of hours per week the focal child spent in care during spring-first grade. The variable combines hours in child care arrangements in which the child spent the most time with hours from additional regular child care arrangements. It was coded as follows (item references are to spring-first grade). If the relevant child care receipt variables for relative, nonrelative, and center-based care (CCQ010, CCQ150, or CCQ260) were equal to 2 (No Receipt), or if the indicator for regular receipt of that type of care (CCQ080, CCQ180, and CCQ340) was equal to 2 (No Regular Receipt), the number of hours for that type of care was coded to 0. If the receipt variables or regular receipt of care variables were refused or unknown, then the number of hours for that type of care was coded as -9 (Not Ascertained). Also, if the regular receipt variable was coded as 1 (Yes), but the hours given was refused or unknown, then the number of hours for that type of care was coded as -9 (Not Ascertained). Otherwise, if the indicator for regular receipt of care was equal to 1 (Yes), and the hours given were greater than or equal to 0, then the number of hours for that type of care was coded as the number of hours given.

The composites also include hours spent with additional regularly scheduled providers of care of the same type. This was done to include child care arrangements such as those in which two different relatives cared for the child on a regular basis or two different child care programs were attended. For each type of care, if the care receipt variables indicated no care of that type, or if the number of providers of that type of care (questions CCQ060, CCQ165, and CCQ325 indicated number of regular providers of each type), was equal to 1, then additional hours were coded to 0. Otherwise, if the number of providers or the number of additional hours (questions CCQ140, CCQ250, and CCQ403 indicated number of hours spent with additional providers) was refused or unknown, then the number of additional hours was coded as -9 (Not Ascertained). Also, if the child did not have a regular weekly schedule for the main provider of a given type—and thus hours for that provider were not elicited—but the respondent gave hours for the secondary provider of a given type, these hours were not included in the composite. Otherwise the number of additional hours was coded to equal the appropriate number of additional hours variables in the instrument (CCQ140, CCQ250, or CCQ403).

This process was followed three times, once each for relative care, nonrelative care, and center-based care. If any of the three primary caregiver hour variables was missing then the total number of hours was coded as -9 (Not Ascertained). Otherwise the total number of hours in regularly scheduled child care was coded as the sum of the six hour variables.

It should be noted that while P3HRSNOW and P4HRSNOW were created almost identically to the same composite variable in kindergarten (P1HRSNOW), there was one difference. In kindergarten, questions were asked about whether the child was ever in a particular type of care. If not, P1HRSNOW was set to 0. Because questions about the child having ever been in a particular type of care were not included in first grade, they were not part of the composite variable definition for the first grade variables.

Number of Child Care Arrangements (P3NUMNOW and P4NUMNOW)

Two other composite variables were used to indicate the total number of all types of care arrangements the focal child had on a regular basis during the summer before first grade (P3NUMNOW) and during spring-first grade (P4NUMNOW). The variables were created as follows (item references are to spring-first grade). If any of the child care receipt variables for relative, nonrelative, or center-based care (CCQ010, CCQ150, or CCQ260) was refused, unknown, or missing then P4NUMNOW was coded as -9 (Not Ascertained). If any of the care receipt variables was equal to 1 (Yes), but its corresponding number of arrangements variable (CCQ060, CCQ165, and CCQ325) was refused, unknown, or missing, then P4NUMNOW was again coded as -9 (Not Ascertained). Otherwise, the number of arrangements indicated in CCQ060, CCQ165, and CCQ325 were summed to obtain the total number of current child care arrangements.

Primary Nonparental Child Care Arrangement (P3PRIMNW and P4PRIMNW)

Composite variables were also created to indicate the primary, nonparental child care arrangement in which the child spent the most hours per week during the summer before first grade (P3PRIMNW) and during spring-first grade (P4PRIMNW). The values for this variable are as follows:

- 0=No nonparental care,
- 1=Relative care in child's home,
- 2=Relative care in another home
- 3=Nonrelative care in child's home,
- 4=Nonrelative care in another home
- 5=Center-based program
- 6=Two or more programs
- 7=Location of care varies

To obtain the composite (item references are to spring-first grade), hours were compared for relative care in child's home (CCQ090) or in other home (CCQ070); nonrelative care in child's home (CCQ190) or in other home (CCQ170); and center/program care (CCQ355). First, the composite P4HRSNW, described earlier, was used to code individuals missing current hours of care (P4HRSNW=-9) or with no hours of nonparental care (P4HRSNW=0). Those with missing hours of care were coded as -9 (Not Ascertained); those with no hours of care were coded as 0.

For the remaining cases, if the number of hours of either relative or nonrelative care (given in CCQ090 and CCQ190) were higher than all other hours of care, the variable indicating location of care for that type was examined using instrument items CCQ070 and CCQ170. If location of care was missing, then P4PRIMNW was coded as -9 (Not Ascertained); if P4PRIMNW was not missing, then P4PRIMNW was coded 1, 2, 3, or 4, depending on the type (relative/nonrelative) and location (child's home/other home) of care. Otherwise, if the number of hours of care in center-based programs (CCQ355) was higher than for relative or nonrelative care, then P4PRIMNW was coded as 5. If the number of hours of care was equal for two or more types of care, P4PRIMNW was coded as 6. P4PRIMNW was coded as 7 if the location of care varied between the two homes.

7.4.2 Family and Household Composite Variables

Many composites were created to capture information about the sampled children's family and household characteristics. Several of these are described below. All of the family and household composites are listed and described in tables 7-13 and 7-14.

Number of Siblings (P4NUBSIBS)

The composite P4NUMSIBS indicates the total number of siblings (full/step/adoptive or foster) with whom the child lived in the household (FSQ160 and FSQ170). Siblings were identified through the respondents' stated relation of the sibling to the focal child. In addition, any child that was reported to be a child of the focal child's parent/guardian was considered a sibling of the focal child.

Parent and Household Members' Age (P4LESS18, P4OVER18, P4HDAGE, and P4HMAGE)

There are several composite variables on the file that refer to the ages of adults and children in the household. These are P4LESS18 (total number of people in the household under age 18, including focal child, siblings, and other children), P4OVER18 (total number of people in the household age 18 or older, siblings, and other children), P4HDAGE (age of resident father), and P4HMAGE (age of resident mother). The ages of these persons in the household were collected during the fall of kindergarten in the household matrix. However, in the spring of kindergarten and the fall and spring of first grade, age was not updated for household members who were previously in the household. In the first grade, for the composites (P4LESS18, P4OVER18, P4HDAGE, and P4HMAGE), 1 year was added to the age of each person who had been in the household in fall-kindergarten or spring-kindergarten. There were no changes to the ages of those who entered the household in fall-first grade or spring-first grade.

Socioeconomic Status and Poverty

Socioeconomic status (SES) was computed at the household level using data for the set of parents who completed the parent interview in fall-first grade or spring-first grade. The SES variable reflects the socioeconomic status of the household at the time of data collection for spring-first grade (spring 2000). The components used to create the SES were as follows:

- Father/male guardian's education;
- Mother/female guardian's education;
- Father/male guardian's occupation;
- Mother/female guardian's occupation; and
- Household income.

Occupation was recoded to reflect the average of the 1989 General Social Survey (GSS) prestige score[2] of the occupation. It was computed by averaging the corresponding prestige score of the 1980 Census occupational category codes covered by the ECLS-K occupation. Table 7-14 provides details on the prestige score values (W1DADSCR, W1MOMSCR).

The variables were collected as follows:

1. **Income.** The information about income was collected in spring-first grade. As a result, income is missing for all households with parents who did not participate in the survey in spring-first grade. Broad-range and detailed-range income questions were asked of all participants. The broad range classifies household income as $25,000 and less per year, or as greater than $25,000. The detailed range classifies household income as shown in table 7-1:

[2] Nakao, K., and Treas, J. (1992). *The 1989 Socioeconomic Index of Occupations: Construction from the 1989 Occupational Prestige Scores:* General Social Security Methodological Report No. 74. Chicago: NORC.

Table 7-1.—Levels of the detailed income range

Detailed income range	Total household income
1	$5,000 or less
2	$5,001 to $10,000
3	$10,001 to $15,000
4	$15,001 to $20,000
5	$20,001 to $25,000
6	$25,001 to $30,000
7	$30,001 to $35,000
8	$35,001 to $40,000
9	$40,001 to $50,000
10	$50,001 to $75,000
11	$75,001 to $100,000
12	$100,000 to $200,000
13	$200,001 or more

Households that were determined to meet the size and income criteria related to poverty shown in table 7-2 were asked to report income to the nearest $1,000. (We call this exact income for simplicity.)

2. **Parent's education.** The information about parent's education was collected in spring-first grade. Any household not interviewed in the base year had this information collected in fall-first grade as well if it was part of the fall-first grade subsample.

3. **Parent's occupation.** The information about parent's occupation was collected in spring-first grade only.

Table 7-2.—Households asked to report income to the nearest $1,000

Household size	Total household income
Two to three	$15,000 or less
Four to five	$20,000 or less
Six to seven	$25,000 or less
Eight	$30,000 or less
Nine	$35,000 or less

Because not all parents responded or responded to all of the questions, there were missing values for some of the components of the SES composite variable, but only a small percentage of values for the education, occupation, broad income range, and detailed income range variables were missing. A fraction of respondents were asked to report exact income, but the amount of missing data for this subset of participating households was considerable (see table 7-3).

Table 7-3.—Missing data for socioeconomic status source variables

Variable	Number missing	Percent
Mother's education	886	5.6
Father's education	1,309	8.3
Mother's occupation	621	3.9
Father's occupation	556	3.5
Detailed income range	1,672	10.5
Exact Household income	1,076	33.0*

* The denominator is the number of households identified to report income to nearest $1,000

A two-stage procedure was used to impute missing values of each component of the SES composite variable. First, if a parent interview had been conducted in fall-first grade or kindergarten, missing values for the spring-first grade education, occupation, and detailed income range variables were filled in with previous round values. The rationale for using this approach was that the best source of data for an individual or a household was the data from the previous year. Thus using data from a previous round was the first method of imputing these items.

This first imputation stage was implemented as follows:

1. Education level was brought forward from the most recent previous round. This was done only if the same person was the parent figure both in spring-first grade and in the earlier round.

2. Occupation was brought forward only if the individual was in the labor force (i.e. was working at a paid job, on vacation from a paid job, or looking for a job). It was also required that the same person be the parent figure both in spring-first grade and in the earlier round. NOTE: Prestige scores were not assigned to individuals unless they were in the labor force, regardless of whether they reported an occupation

3. Detailed income category data were brought forward from the base year.

Second, data still missing after this initial step were imputed using a hot deck methodology. In hot deck imputation, the value reported by a respondent for a particular item is assigned or “donated” to a “similar” person who failed to respond to that question. Auxiliary information known for both donors and nonrespondents is used to form groups of persons having similar characteristics. These groups of similar respondents and nonrespondents are called “imputation cells.” The imputed value for a case with a missing value is taken from a randomly selected donor among the respondents within the cell.

Imputation cells were defined by respondent characteristics that were the best predictors of the variables to be imputed. These relationships had been determined previously by CHAID (Chi-squared Automatic Interaction Detector) analyses of the base year data. Missing values for the education, occupation, and detailed income range variables were imputed by the hot deck method for all households. In addition, an exact income value was imputed for households that met the size and income criteria presented in table 7-2. Because there were too few donors in seven imputation cells, it was impossible to impute exact household income by the hot deck method for 129 records. Instead, the midpoint of the detailed income range was imputed for these cases. Hot deck imputation was done in a sequential order, separately, by type of household (female single parent, male single parent, and both parents present). For households with both parents present, the mother’s and father’s variables were imputed separately. The new imputed values were used in the creation of the imputation cells if these values had been already imputed. If this was not the case, an “unknown” or missing category was created. As a rule, no imputed value was used as a donor. In addition, the same donor was not used more than two times. The order of hot deck imputation for all the variables was from the lowest percent missing to the highest, after filling in missing values with data from previous rounds.

Occupation imputation involved two steps. First, the labor force status of the parent was imputed (i.e., whether the parent was employed). Then the parent’s occupation was imputed only for those parents whose status was identified as employed either through the parent interview or the first imputation step. Likewise, detailed income range was imputed, then household size and detailed income range were used to determine whether exact household income would be imputed. The detailed income range was imputed in two steps: first for cases where the broad income range was known, and second for cases where it was unknown. For households meeting the size and income criteria presented in table 7-2, the variable for income was imputed last.

For households where both parents were present, the order of hot deck imputation was:

- Mother’s education;
- Father’s labor force status;
- Father’s occupation, if the father was employed or imputed as employed;
- Mother’s labor force status;
- Mother’s occupation, if the mother was employed or imputed as employed;
- Father’s education;
- Detailed income range, where the broad income range was known;
- Detailed income range, where the broad income range was unknown; and
- Exact household income value for households meeting the size and income criteria.

At this point, all of the missing values had been imputed. However an exact income value was still required to construct the SES composite for households that did not meet the size and income criteria for poverty. The midpoint of the detailed income range was assigned for this purpose to households at income levels of $50,000 or less. To derive the midpoint for higher income levels, households were ranked by kindergarten year exact income within the levels of the first-grade year detailed income range. The four upper levels of the detailed range were subdivided according to the quantiles of the exact income. The divisions were at the 33rd and 67th percentiles for levels of $50,001 to $75,000 and $75,001 to $100,000; they were at the 25th, 50th, and 75th percentiles for levels of $100,001 to $200,000 and $200,001 or more.

The log of exact household income was used to construct the SES composite. This value does not vary widely within the levels of the detailed income range, so the midpoint was a reasonable choice for constructing the SES composite. It was used only for this purpose and was not retained in the data file after the SES composite was computed.

All missing values of the SES components were imputed by the process described above. Tables 7-4 through 7-7 summarize the results. To assist in interpreting the tables, each number in the top half of table 7-5 is described here. The numbers in this part of the table refer to the imputation of mother’s occupation, for which 621 records had missing values. As explained earlier, imputation of occupation and household income was dependent on labor force status and detailed income range, respectively, which had also been imputed. The 621 records were resolved as follows:

- For 116 cases where occupation was missing, it was known that the mother was not in the labor force, that is, it was reported that she had not worked during the previous week, was not on vacation, or was not looking for work. Thus, occupation was inapplicable.
- For 105 cases where occupation was missing, labor force status was unknown and "not in the labor force" was filled in using data from previous rounds.
- For 24 cases where occupation was missing and labor force status was unknown, "not in the labor force" was imputed by hot deck.
- For 55 cases where occupation was missing and it was known that the mother was in the labor force, occupation was filled in using data from previous rounds.
- For 61 cases where occupation was missing and it was known that the mother was in the labor force, occupation was imputed by hot deck.
- For 192 cases where occupation was missing, "in the labor force" had been filled in using data from previous rounds and occupation was imputed by hot deck.
- Finally, for 68 cases where occupation was missing and "in the labor force" had been imputed by hot deck, occupation was imputed by hot deck.

The other tables of imputed variables describe similar processes for resolving cases.

Table 7-4.—Imputed education variables

SES component	Total missing	Number of values filled from previous rounds	Number of values imputed by hot deck	Number of cases resolved
Mother's education	886	512	374	886
Father's education	1,309	466	843	1,309

Table 7-5.—Imputed occupation variables

SES component	Number of values known	Number of values filled from previous rounds	Number of values imputed by hot deck	Number of cases resolved
Mother's occupation		Labor force status		
Not in labor force	116*	105	24	245
Filled from prev rounds	55			55
Hot deck imputed	61	192	68	321
Total missing				621
Father's occupation		Labor force status		
Not in labor force	74*	22	13	109
Filled from prev rounds	122			122
Hot deck imputed	77	166	82	325
Total missing				556

* Occupation was not imputed if "not in labor force" was known

Table 7-6.—Imputed detailed income range

SES component	Total missing	Number of values filled from previous rounds	Number of values imputed by hot deck		Number of cases resolved
			Broad income range		
			known	unknown	
Detailed income range	1,672	1,593	23	56	1,672

Table 7-7.—Imputed exact household income value

SES component	Number of values known	Number of values filled from previous rounds	Number of values imputed by hot deck	Number of cases resolved
Exact hh income value		Detailed income range		
Hot deck imputed	379	536	32	947
*Assigned midpoint	76	53		129
Total missing				1,076

* No donors could be found for these records; the midpoint of the detailed income range was assigned

Once the components of the SES variable were imputed, their corresponding z-scores or normalized values were computed. The expression of z-score z_{hi} for the h-th component in the i-th household is

$$z_{hi} = \frac{x_{hi} - \bar{x}_w}{se(\bar{x}_w)},$$

where

x_{hi}	is	the value of the h-th SES component for the i-th household;
w_i	is	the base weight for the i-th record;
$\bar{x}_w$	is	the weighted mean of x_{hi} ; and
$se(\bar{x}_w)$	is	the standard error of $\bar{x}_w$.

Thus, each component was converted to a z-score with mean of 0 and a standard deviation of one. For income, the component x_i is the logarithm of the income for i-th household. The logarithm of income was used because the distribution of the logarithm of income is less skewed than the direct income values. The SES value for the i-th household was then computed as

$$SES_i = \frac{\sum_{h=1}^{m_i} z_{hi}}{m_i},$$

where m_i is number of nonmissing SES components for the i-th household. W1SESL is the continuous variable for the SES composite that ranges from –2.96 to 2.88. As described, the SES composite is the average of up to five measures, each of which was standardized to have a mean of 0 and a standard deviation of 1, hence the negative valves. For analyses that require a continuous SES measure, such as multivariate regressions, W1SESL is the variable to use. A categorical SES variable (W1SESQ5) was created that contains the quintile for the value of the composite SES for the child. Quintile 1 represents the lowest SES category and quintile 5 represents the highest SES category. The quintiles were computed at the child level using the spring-first grade parent weights. For categorical analyses, use W1SESQ5 and the parent weight.

Note that for households with only one parent present, not all the components were defined. In these cases, the SES was computed averaging the available components.

The imputed income variable (W1INCOME) was also used to create a household-level poverty variable (W1POVRTY). Income was compared to preliminary census poverty thresholds for 1999, which vary by household size. Table 7-8 shows the broad income categories used in the ECLS-K parent interview for determining whether to ask a more detailed question about income to the nearest 1,000. For comparison, the table also shows weighted poverty thresholds from census.[3] Households whose income fell below the appropriate threshold were classified as poor (see table 7-8). For example, if a household contained two members, and the household income was lower than $11,214, then the household was considered to be below the poverty threshold.

Table 7-8.—ECLS-K and census poverty thresholds for 1999*

Household size	ECLS-K income categories	Census weighted average thresholds for 1999
2	Less than or equal to $15,000	$11,214
3	Less than or equal to $15,000	$13,290
4	Less than or equal to $20,000	$17,029
5	Less than or equal to $20,000	$20,127
6	Less than or equal to $25,000	$22,727
7	Less than or equal to $25,000	$25,912
8	Less than or equal to $30,000	$28,967
9+	Less than or equal to $35,000	$34,417

* U.S. Census Bureau, Current Population Survey. http://www.census.gov/hhes/poverty/threshld/thresh99.html.

Parent Education (W1PARED, W1DADED, and W1MOMED)

There are three parent education composites on the file. These are W1PARED (the highest level of education for the child's parents or nonparent guardians who reside in the household), W1DADED (father's highest level of education), and W1MOMED (mother's highest level of education). The variables include both parent (birth, adoptive, step, and foster) and nonparent guardians. For example, if the child had no parents but had a guardian, the education of the guardian and his/her spouse were used in the creation of the composites if the guardian was specified as such in the relationship variable or if the guardian was the respondent/respondent spouse and there were no other parent figures in the household.

[3] The ECLS-K provides an approximate but not exact measure of poverty. Income category thresholds used in the parent questionnaire are similar but not identical to those from weighted census averages.

In spring-first grade, parent education level was updated from the previous round if it was a household that had already been part of the study. Respondents were asked if they or their corresponding parent figures, if applicable, completed any additional grades of school or received any diplomas or degrees (PEQ.010). If so, PEQ.020 asked what grade the parent completed or what degree was received. If this education level was less than the education level reported in spring-kindergarten, the education level from spring-kindergarten was kept for the composite.

If both parents/guardians resided in the household, W1PARED was the highest value for education level from either the mother/guardian in W1MOMED or the father/guardian in W1DADED. If the household only had one parent or guardian, then W1PARED was equal to either W1MOMED or W1DADED depending on which parent or guardian resided with the child. If the education data for either of the parents was missing it was imputed, and the composite W1PARED was created based on both the reported and imputed data.

Parent Race-Ethnicity (P4HDASN, P4HDBLK, P4HDIND, P4HDPACI, P4HDWHT, P4HDHISP, P4HDMT1R, P4HMASN, P4HMBLK, P4HMIND, P4HMPACI, P4HMWHT, P4HMHISP, P4HMMT1R, P4HDRACE, and P4HMRACE)

The composites for race-ethnicity for the parents were calculated in the same way as those for the child, except that there is not a variable that supplements parent reported race-ethnicity with FMS data similar to the variable RACE for children. All data on parent race-ethnicity come from the parent interview. Race-ethnicity for parents is presented in the ECLS-K files in two ways—as dichotomous variables for each race-ethnicity category (for the father/male guardian the variables are P4HDASN, P4HDBLK, P4HDIND, P4HDPACI, P4HDWHT, P4HDHISP, P4HDMT1R, and for the mother/female guardian the variables are P4HMASN, P4HMBLK, P4HMIND, P4HMPACI, P4HMWHT, P4HMHISP, P4HMMT1R), and as a categorical race-ethnicity composite (for the father/male guardian it is P4HDRACE, and for the mother/female guardian it is P4HMRACE).

Respondents were allowed to indicate that they belonged to more than one of the five race categories (white, black or African American, American Indian or Alaska Native, Asian, Native Hawaiian or other Pacific Islander). From these responses, a series of five dichotomous race variables were created that indicated separately whether the respondent belonged to each of the five specified race groups. In addition one more dichotomous variable was created for those who had simply indicated that they were multiracial without specifying the race (e.g., biracial). The retention of the dichotomous variables on the file allows users to create different composites as needed.

Parent race-ethnicity is obtained for all parents and spouses of respondent parents. For example, in a family with a birth mother and stepfather the race-ethnicity of both parents would be obtained. However, in a family with a birth mother and the mother's boyfriend, if the mother's boyfriend was not identified as a spouse or partner of the mother, the race-ethnicity of the mother would be obtained but not that of the boyfriend.

7.4.3 Teacher Composite Variables

Details about how two of the teacher composites, child grade level and class type, were created are provided here. All of the teacher composites are listed and described in table 7-14.

Grade-Level Composite (T4GLVL)

To create the grade-level composite (T4GLVL), four possible sources of information were used: (1) teacher questionnaire part C, for kindergarten (Q16 T4KGRADE for grade level and Q17 T4KTYPCL for morning, full day, etc.) and for first grade (Q17 T4GRADE for grade level); (2) FMS information about grade level; and (3) student record abstract information about grade level (Q19 U4CRGRD), and (4) grade-level information from the special education questionnaire part B (Q24 E4GRADE).

If conflicts existed between these four sources, the grade level indicated by the majority of the nonmissing sources was used for T4GLVL. If equal numbers of sources were in conflict, the data from the teacher questionnaire C (TQC) was given highest priority followed by special education teacher data and then the FMS. Two exceptions were made. First, because the FMS grade-level information did not allow for ungraded classrooms, the FMS information was not considered in any case in which at least one source indicated "ungraded."

The second exception to using the grade level indicated by the majority of data sources involved a consideration of possible respondent error. The variables T4GRADE, T4KGRADE, E4GRADE, and U4CRGRD each offered the following choices to indicate grade level:

Kindergarten...	1
First grade	2
This is an ungraded classroom	3
Second grade	4

There was a high level of disagreement across data sources on this question, most likely the result of respondents circling “1” to indicate first grade, or “2” to indicate second grade. In these instances, if there were two sources of information, one of which was the FMS, and the FMS and the other source disagreed, the FMS data was used. If, however, there were more than two sources, T4GLVL was determined by that shown by the majority of variables.

Class Type (A4CLASS)

This variable (A4CLASS), which is based on data from the spring-first grade teacher questionnaire part A (TQA), indicates whether the TQA data attached to the child’s record are for a morning, afternoon, or full-day kindergarten class, or a nonkindergarten class (e.g., first or second grade). In spring-first grade, the categories for A4CLASS changed from those used for the kindergarten versions of this variable because A4CLASS is used to note TQA data for both kindergarten and nonkindergarten children (first and second graders). A4CLASS is defined by the TQA data that are linked to the child’s record. For the majority of children, A4CLASS and the child’s grade-level values (taken from the grade level composite T4GLVL) match. However, there are a small number of cases for which they do not. For example, a child may have been in a morning kindergarten class, but his teacher filled out only the afternoon section of the TQA instrument. In that case, A4CLASS was set to ‘kindergarten afternoon class’, and the teacher’s afternoon data were linked to that child. If the teacher had not responded to the TQA, the value of the A4CLASS was coded as blank; if the teacher had a TQA with two classes and neither matched the child’s class, then the value was coded as -9 (Not Ascertained).

7.4.4 School and Class Composite Variables

Variables on school and class characteristics were constructed from the teacher and school data and the sample frame. Details on how some of the variables were created follow.

School Type (S3SCTYP and S4SCTYP)

In fall-first grade the information for S3SCTYP was taken from the frame because there was no school administrator interview during that data collection period. In spring-first grade, S4SCTYP was created as follows. For new schools, questions 5 (S4PUBLIC) (whether school is public) and 7 (S4CATHOL, S4OTHREL) (type of private school) from the school administrator questionnaire, along with school sample frame data, were used to create the school type composite variable. If the response to question 5 (Is this a public school?) was “Yes,” then S4SCTYP was coded “public.” If the response to question 5 was “No,” and the response to question 7.a. (S4CATHOL) (Is your school a Catholic school) was “Yes,” then the school was coded as “Catholic.” If the responses to both questions 5 and 7.a. were “No,” and the response to question 7.b. (S4OTHEREL) (Is your school private with another religious affiliation?) was “Yes,” then S4KSCTYP was coded as “private, other religious.” Otherwise, if the school had a school administrator questionnaire with valid answers to questions 5 and 7, then S4KSCTYP was coded as “private, non-religious.”

For returning schools, the question about school type was not asked again because it was not expected that school type would have changed in a year. For these schools, the school type information was obtained either from the school sample data for originally sampled schools, the spring-kindergarten SAQ information, or the fall first-grade information if the school was part of the subsample in fall-first grade.

If there was no school administrator questionnaire, then school sample frame data were used to create the composite. If the school administrator information conflicted with the frame data, the school administrator data were used. If school type was “unknown” from both sources of data, the composite was coded as -9 (Not Ascertained). If the child was schooled at home, the composite was coded “Not Applicable.”

Public or Private School (S3PUPRI and S4PUPRI)

These variables are a less detailed version of school type (with only two categories—public and private) and are derived from the school type composites S3SCTYP and S4SCTYP described earlier. In fall-first grade the information was taken from the frame. In spring-first grade, it was created as follows. If S4SCTYP was 4 (public), then S4PUPRI was coded as public (1). If S4SCTYP was 1-3 (Catholic, other religious, other private) then S4PUPRI was coded as private (2). If S4SCTYP was coded as Not Ascertained (-9), then S4PUPRI was -9 (Not Ascertained). If S4SCTYP was coded "Not Applicable," then S4PUPRI was coded "Not Applicable."

School and Grade-Level Enrollment (S4ENRLK, S4ENRLF, and S4ENRLS)

There are three composite enrollment variables on the first grade file: total school enrollment (S4ENRLS), first grade enrollment (S4ENRLF), and kindergarten enrollment (S4ENRLK). Two of these variables were created somewhat differently than the enrollment variables in spring-kindergarten.

Total school enrollment was created in the same way as it was in spring-kindergarten using the total school enrollment variable from the school administrator questionnaire (S4ANUMCH). If this variable was missing, frame data were used. If frame data were also missing, the variable was coded -9 (Not Ascertained). The only change between spring-kindergarten and spring-first grade was that the frame data were updated. The frame data for private schools came from the 1999-2000 Private School Survey (PSS). The latest frame data for public schools came from the 1998-99 Common Core of Data (CCD) public school universe.

Kindergarten and first grade enrollment were calculated from a different source in spring-first grade than in spring-kindergarten. In spring-kindergarten, enrollment variables were based on data from the school administrator questionnaire (variables S4HDCHDN, S4FDCHDN, S4KCCHDN, S4TKCHDN and S4T1CHDN were summed and frame data were used if the variables were missing. In spring-first grade, questions about kindergarten enrollment were not asked so the data were taken from the frames. The same is true for first grade enrollment.

Percent Minority Students in the School (S4MINOR)

The composite variable that indicates the percentage of minority students in a school in spring-first grade (S4MINOR) is similar to the composite with the same information in spring-kindergarten (S2MINOR.) However, in spring-first grade, the school administrator questionnaire item that asked about the percent of minority students in the school had different response options than the one used in spring-kindergarten. In spring-kindergarten, the percent of minority students was derived from answers to the school administrator questionnaire by determining the percentage of children who were either of Hispanic or Latino origin (question 14) and the percentage of children who were American Indian or Alaska Native, Asian, black or African American, or Native Hawaiian or Other Pacific Islander (question 15) to create the percent minority composite. In spring-first grade, Hispanicity and race were included in the same question (question 11 for new schools and question 4 for returning schools). The combined race-ethnicity categories were Hispanic, regardless of race; black, not of Hispanic origin; white, not of Hispanic origin; Asian or Pacific Islander; American Indian or Native Alaskan; and other.

In spring-first grade, school administrators were allowed to report their answers as either numbers or percents, whereas in spring-kindergarten they were asked to report answers as percents. Because the composite for the percentage of minority students is reported as a percent, all answers recorded as numbers in spring-first grade were converted to percentages for the composite variable.

A flag for each individual race-ethnicity variable indicates whether the answer was reported as a number or a percent.[4] Because the composite is calculated as a percent, these flags will not be needed by users unless the analyst is interested in examining how answers were reported. If the flags (S4ASNFL, S4HSPFL, S4BLKFL, S4WHTFL, S4INDFL, and S4OTHFL) were equal to 1 for each of the race variables S4PCTASN, S4HISPPCT, S4BLKPCT, S4WHTPCT, S4INDPCT, S4OTHPCT, these 6 race-ethnicity variables were reported by the respondent as percentages. In this case, the five percentages for minorities (all categories but White, not of Hispanic origin) were added together to create the composite (S4PCTASN + S4HISPPCT + S4BLKPCT + S4INDPCT + S4OTHPCT). However, if the flags were equal to 2 for all the race-ethnicity variables, the answers were reported as numbers, and the race-

---

[4] There were also other questions in the school administrator questionnaire that allowed for answers to be recorded as either a number or percent. The flags for these variables are S4ADAFLG (average daily attendance reported as number/percent), S4ASINFL (question about Asian or Pacific Islander teachers reported as number or percent), S4HISPFL (question about Hispanic teachers reported as number or percent), S4BLAKFL (question about black teachers reported as number or percent), S4WHTEFL (question about white teachers reported as number or percent), S4AINDFL (question about American Indian or Native Alaskan teachers reported as number or percent), and S4TOTHFL (question about teachers of other races reported as number or percent). In all cases, the final variables related to these flags are reported as percentages, but the flags indicate how the answers were originally recorded by respondents.

ethnicity percentages were calculated by dividing the sum of numbers in the minority categories by the total enrollment variable S4TOTNUM.

In some cases, the composite could not be obtained from the data because of missing data or errors. If the composite could not be derived from the data,[5] percent minority was obtained from the school sample frame. If there was no value for this on the frame, the composite was coded -9 (Not Ascertained.)

School Instructional Level (S4SCLVL)

The purpose of this composite is to classify schools based on the highest grade taught in the school. For new schools, question 4 (with grade-level variable names S412TH, S411TH, S4TENTH, S4NINTH, S48TH, S47TH, S4SIXTH, S4FIFTH, S4FOURTH, S4THIRD, S4SECOND, S4GRADE1, S4PRE1, S4KINDER, S4TRANS, and S4PRKNDR) (grade levels included in the school) of the school administrator questionnaire was used to create this composite. The highest grade level circled on the form was determined, and the grade level was classified accordingly. If question 4 was left blank, question 10 (with grade-level variable names S4TWEL12, S4ELEV11, S4TEN10, S4NIN9TH, S4ATE8TH, S4SEV7TH, S4SIX6TH, S4FIF5TH, S4FOR4TH, S4THI3RD, S4SEC2ND, S4FIR2ST, S4PRE1ST, S4FIR2ST, S4PRE1ST, S4KINGAR, S4REDDYN, and S4PREK) from the school administrator questionnaire (grade levels that participate in special programs, if applicable) was used as a proxy for question 4. If the respondent did not answer questions 4 and 10, or there was no school administrator questionnaire, then school sample frame data were used to determine the value for the composite.

---

[5] Recoding rules for data with errors were the following:

1) If answers were reported as numbers and the total number of students was missing, the total from another question about total enrollment (Q3a S4ANUMCH) was used if the difference between the summed total and the reported Q3a total was within a 5 percent confidence interval (95-105 percent).
2) If the method of reporting was mixed (some in numbers, other in percents), the race percentages were -9 filled and the frame was checked for information.
3) If percentages were recorded, with none of the above errors, and the summed total across categories was within +/-5 percent of the reported total, any blanks were recoded to 0.
4) If the summed total was not 95 – 105 percent of the sum reported or not 95 – 105 percent of total enrollment from another question (Q3a S4ANUMCH), the individually reported percentages and numbers were made -9 and the frame was checked for information.
5) If numbers were reported, with none of the above errors, and the summed total across categories were within +/- 5 percent of the reported total, any blanks were recoded to 0. The composite came from the data in this case. If the summed total was outside these boundaries (e.g., the summed total was outside the 10 percent confidence interval) any blanks were filled with -9's and racial percentages were not calculated. In this case, the composite information had to come from the frame if available.

The grade-level questions were among those not asked of school administrators in returning schools in order to reduce the burden on school administrators. The value in spring-kindergarten for school instructional level (S2SCLVL) was copied to the spring-first grade variable (S4SCLVL).

7.4.5 Student Record Abstract and Field Management System Composite Variables

The composite variables created for the student record abstract data are described below.

School Year Start and End Dates (U4SCHBDD, U4SCHBMM, U4SCHBYY, U4SCHEDD, U4SCHEMM, and U4SCHEYY)

Variable names and descriptions of the composites that indicate the school year start and end dates are:

U4SCHBDD	U4 School Year Starting Date, Day
U4SCHBMM	U4 School Year Starting Date, Month
U4SCHBYY	U4 School Year Starting Date, Year
U4SCHEDD	U4 School Year Ending Date, Day
U4SCHEMM	U4 School Year Ending Date, Month
U4SCHEYY	U4 School Year Ending Date, Year

The composites were created for two purposes. The first purpose was to provide one set of school start and end dates for the *whole* school year. They are based on responses to multiple questions in the student record abstract form about the start and end dates for school terms. The number of these responses varied depending on the number of terms at the school. For example, if the school had two semesters, there were generally two sets of dates, one for the beginning and end of each term (e.g., one set indicated the year, month, and day of the start and end of term 1 taken from variables U4YYSTR1, U4MMSTR1, U4DDSTR1, U4YYETR1, U4MMETR1, and U4DDETR1, and the other set indicated the year, month, and day of the start and end of term 2 taken from variables U4YYSTR2, U4MMSTR2, U4DDSTR2, U4YYETR2, U4MMETR2, and U4DDETR2). If the school had trimesters, there were three sets of dates; if the school had four quarters there were four sets of dates. Some schools also had one term and one set of dates.

The second purpose for creating the composite variables was to assist users in handling inconsistencies in the ways that respondents reported the term dates. In most cases, students in the same school were all indicated to be in schools with the same start and end dates. In some cases, however, there appeared to be errors in the way that the school-level data were recorded on the abstract forms. For example, there were cases in which all but a few of the students in the same school had data indicating the same start and end dates for the school year. Usually, the data in error reflected when the individual student entered or left the school rather than the start and end dates of the school year for that student or was one or two days different from the dates given for other students in the same school. The composites were created by using the school start and end dates that the majority of children in a school had. A slightly different strategy was used for “year-round” schools. This is explained here.

Year-round Schools (F4YRRND)

This composite was created using data from the FMS. Some children were in year-round schools in which groups of children had different schedules and attended during segments of the year called “tracks.” For those schools, the starting date for the school year was the earliest date reported for a track of students. The latest date reported for a track was designated as the end of the school year. The values for the year-round school variable are 1 (Yes) and 2 (No).

Total Days Absent During School Year (U4ABSN)

A composite was created that indicates the total number of days the child was absent during the school year. It was derived as follows. First, a total was created by summing the number of absences reported by individual term. Second, a total was created based on the sum of the excused and unexcused absences for all applicable terms. The first total was used for the composite if the second total was not available because of missing data. If both totals were available, the first total was used for the composite value if the difference between total one and total two was equal to or less than 10. However, if the first total was 0 but there were data for excused or unexcused absences, the composite was set to -9 (Not

Ascertained). If the difference between the totals was greater than 10, the composite was also set to -9 (Not Ascertained).

Total Days Tardy During School Year (U4TARD)

A composite was also created that indicates the total number of days the child was tardy during the school year. It was derived in the same way as the composite for total absences. First, a total was created by summing the number of tardies reported by individual term. Second, a total was created based on the sum of the excused and unexcused tardies for all applicable terms. The first total was used for the composite if the second total was not available because of missing data. If both totals were available, the first total was used for the composite value if the difference between total one and total two was equal to or less than 10. However, if the first total was 0 but there were data for excused or unexcused tardies, the composite was set to -9 (Not Ascertained). If the difference between the totals was greater than 10, the composite was also set to -9 (Not Ascertained).

Indicator of Whether Child Received Special Education Services (F4SPECS)

The composite variable F4SPECS indicates whether or not the child received special education services in the spring of first grade, based on the presence or absence of a link to a special education teacher in the FMS in spring-kindergarten. Children were linked to special education teachers in the FMS. The values are 1 if the child received special education services, 2 if the child did not receive special education services, and -9 if the link was missing between the child and his or her teacher in the FMS.

Indicator of Whether Child Has an Individualized Education Plan (IEP) on Record at School (U4RIEP)

The variable U4RIEP indicates whether or not the child had an IEP or Individualized Family Service Plan (IFSP) on record at his/her school or another school in the spring of first grade. The values for the variable are 1 (child has an IEP/IFSP on record at his or her school, or at another school) and 2 (child does not have an IEP/IFSP on record at his or her school).

7.4.6 Base Year (Kindergarten Year) Composites on the First Grade File

There are several composite variables from the base year of the study that have been made available on the first grade file. These are described here.

The first two composites described below (HSATTEND and HSCHECK) can be used to indicate a child's Head Start attendance. These and other variables about Head Start are described in the ECLS-K Base Year Head Start Restricted-Use User's Manual (NCES 2001-025). Variables about Head Start were collected as part of the Head Start Verification Study. The study was designed to identify children in the ECLS-K sample who attended Head Start the year prior to kindergarten and to verify their attendance through contacts with the Head Start program they attended. Verification of Head Start attendance was needed because both parents and schools tend to overreport Head Start participation, with schools overreporting more than parents.

In the ECLS-K, parent, school, and center reports about Head Start participation were obtained. A Head Start directory (a computer database constructed to contain all Head Start centers in the ECLS-K primary sampling units) was used to confirm parent and school reports of Head Start participation and provide more accurate information for locating and contacting Head Start programs. In addition, Head Start center directors were given questionnaires to confirm a child's participation and in some cases contacted by telephone to confirm enrollment at the center. The study was funded by the Administration on Children, Youth, and Families (ACYF) of the U.S. Department of Health and Human Services.

Outcome of Head Start Verification (HSATTEND)

The composite variable HSATTEND is based on reports from centers during the base year of data collection, meaning that the Head Start center returned a questionnaire about the child's Head Start attendance or responded to a telephone contact. The variable was created, in part, from answers to question 1 in the Head Start questionnaire in which the Head Start director was asked to report (a) whether the child attended the program in the 1997-98 program year, (b) whether the child attended, but not in 1997-98, or (c) whether the child had never attended. The variable was also based on respondent comments on questionnaires. Although not asked directly, some respondents volunteered additional information that was relevant to children's classification as Head Start participants (e.g., information that the center was actually not a Head Start center, a note that the center had both Head Start and non-Head Start programs and the child attended the non-Head Start part of the program, etc.). Finally, the variable made use of information from telephone contacts to centers that were made to determine whether centers were Head Start. The additional information from respondent comments and telephone contacts with centers was used to add categories to the information in question 1.

Most of the categories for HSATTEND are self-explanatory. In addition to categories for nonlocatable centers and center nonresponse (categories 1 and 2), there are categories indicating that the child attended Head Start in 1997-98 (category 3); the child attended Head Start in a year other than 1997-98 (category 4); the child never attended the parent- or school-reported Head Start center (category 5); and that the center was not actually a Head Start center (category 6). There is also a category (category 7) for children who attended a non-Head Start program in 1997-1998 in a center that had both Head Start and non-Head Start programs. This outcome was obtained because children within the same classroom or center were funded by different sources. The non-Head Start parts of these programs were not supported by federal Head Start grants but were funded by other sources, such as state funds for children who are at-risk or have a disability.

Head Start Center Was Found in Directory (HSCHECK)

The other Head Start composite (HSCHECK) from the base year that is included in the file refers to whether the Head Start Center was found in a directory used during the parent interview. The Head Start directory was used during the parent interview to assist in obtaining accurate information about the Head Start center. As described above, the directory was in the form of a computer database that contained all Head Start centers in the ECLS-K primary sampling units (PSUs). It was designed to help confirm parent and school reports of Head Start participation and to provide more accurate information for locating and contacting Head Start programs. Parents who reported that their child had attended Head Start the year before kindergarten were asked for the name and address of the Head Start center and the name of the center's director.

The values for the composite are 1 (Yes) and 2 (No). A "Yes" answer on this composite variable indicates that the center was found in the directory. A "No" answer means that it was not. The variable can be used to compare the results of using a directory to confirm parent reports as compared to verifying these reports with centers.

Indicator of Whether Child Received Special Education Services (F2SPECS)

The composite variable F2SPECS indicates whether or not the child received special education services in the spring of the base year, based on the presence or absence of a link to a special education teacher in the FMS in spring-kindergarten. Children were linked to special education teachers in the FMS. The values are 1 if the child received special education services, 2 if the child did not receive special education services, and -9 if the link was missing between the child and his or her special education teacher in the FMS.

Indicator of Whether Child has an Individualized Education Plan (IEP) on Record at School (U2RIEP)

The variable U2RIEP indicates whether or not the child had an IEP or Individualized Family Service Plan (IFSP) on record at his/her school in the spring of the base year, according to information from the student record abstract (U2IEP). The values for the variable are 1 (child has an IEP/IFSP on record at his or her school) and 2 (child does not have an IEP/IFSP on record at his or her school).

It should be noted that this variable is slightly different from the composite in the spring-first grade, U4RIEP. The difference between U4RIEP and U2RIEP is that in spring-first grade respondents could indicate that the child had an IEP/IFSP on record at his/her school *or another school*. Both types of records were counted in value 1 in the composite U4RIEP.

Food Security Status (P2FSRAW, P2FSSCAL, and P2FSSTAT)

Food security status of the families in which children reside was assessed based on responses to the 18 food security questions (P2WORRFD through P2NOMONY) in the spring-kindergarten parent interview. The questions measure a wide range of food insecurity and reduced food intake. They are combined into a scale using statistical methods based on the Rasch measurement model. The items and the food security scale based on them have been validated using both ethnographic and statistical methods. Calculations of food security variables were carried out in accordance with the standard methods described in *Guide to Measuring Household Food Security, Revised 2000* (U.S. Department of

Agriculture, Food and Nutrition Service, Alexandria, VA, March 2000.) Analysis of the ECLS-K data using Rasch methods indicated that use of the standard benchmark household scores was appropriate.

P2FSSCAL is the scale score presentation of the food security items. It is an interval-level measure of food insecurity and is appropriate for linear models. This scale score is a Rasch transformation of the raw score (P2FSRAW). Valid values range from 1.4 to 13.0, with higher values indicating more severe food insecurity. Under Rasch-model assumptions, the scale score for families that affirm no items (raw score = 0) is indeterminate. It is less than the lowest measured value (1.4), but its precise value is unknown and may vary substantially among families. P2FSSCAL for such cases is assigned a value of -6. If these cases (a substantial majority of all cases) are included in linear models, appropriate methods must be used to take into account this indeterminacy.

P2FSSTAT is a categorical measure of food security status formed by dividing P2FSSCAL into four ordered categories: food secure, food insecure without hunger, food insecure with hunger (moderate), and food insecure with hunger (severe). P2FSSTAT is appropriate for comparing prevalence rates of food insecurity and hunger across subpopulations and can be used as a categorical variable in associative models. There are few cases in the most severe category, so for most prevalence reporting purposes, the two categories of food insecure with hunger (moderate and severe) should be collapsed and reported as a single category. When interpreting food security statistics, users should remember that food security status is a household-level characteristic. In most households classified as food insecure with hunger, the children in the household were not hungry.

The food security raw score, P2FSRAW, is a count of affirmative responses to the 18 items. This is an ordinal-level measure of food insecurity and is not recommended for use in analysis. Responses to items skipped because of screening are assumed to be negative. Families with no valid responses are coded as missing (-9). This includes families that were not interviewed in 1999 and a few who refused all of the food security questions or broke off the interview before them. Missing item responses of families with one or more valid responses were imputed as negatives (only 0.25 percent of the sample).

Total Days Absent During School Year (U2ABSN)

A composites was created that indicates for spring-kindergarten the total number of days the child was absent during the school year. This variable was created from student record abstract data and was derived in the same way that the variable for spring-first grade (U4ABSN) was created. First, a total was created by summing the number of absences reported by individual term. Second, a total was created based on the sum of the excused and unexcused absences for all applicable terms. The first total was used for the composite if the second total was not available because of missing data. If both totals were available, the first total was used for the composite value if the difference between total one and total two was equal to or less than 10. However, if the first total was 0 but there were data for excused or unexcused absences, the composite was set to -9 (Not Ascertained). Also, if the difference between the totals was greater than 10, the composite was also set to -9 (Not Ascertained).

Total Days Tardy During School Year (U2TARD)

A composite was also created that indicates for spring-kindergarten the total number of days the child was tardy during the school year. This variable was also created from student record abstract data. It was derived in the same way as the composite for total absences. First, a total was created by summing the number of tardies reported by individual term. Second, a total was created based on the sum of the excused and unexcused tardies for all applicable terms. The first total was used for the composite if the second total was not available because of missing data. If both totals were available, the first total was used for the composite value if the difference between total one and total two was equal to or less than 10. However, if the first total was 0 but there were data for excused or unexcused tardies, the composite was set to -9 (Not Ascertained). Also, if the difference between the totals was greater than 10, the composite was also set to -9 (Not Ascertained).

7.4.7 Parent Identifiers and Household Composition

The construction of parent identifiers and the household composition variables from the parent interview data was a two-step process. First, individuals identifying themselves as the child's mother/father were located within the household roster, and the type of their relationship to the child (biological, adoptive, foster, step, partner, or unknown) was established. For households containing more than one father or mother, a hierarchy was used to designate the "current," or residential, parent of each gender. The biological parent, if present, was always the current mother or father. In the absence of a biological parent, the current mother/father designation was given to the adoptive, step, foster/guardian, partner, or "unknown-type" parent. This information, along with household size and presence or absence of grandparents, siblings, and other relatives was used to construct the household composition variables P4HPARNT, P4HDAD, P4HMOM, and P4HFAMIL and parent type variables P4MOMTYP, and P4DADTYP.

After the actual residential parents were identified and the composite variables were constructed, in any household without a parent, the household respondent (and his or her spouse, if applicable) was assigned as a "parent figure." Parent demographic variables (including age, race-ethnicity, and education) were then constructed for all parents/parent "figures." It should be noted, however, that these parent "figures" were not defined as parents (meaning biological, step-, adoptive, or foster) in the construction of the household composition composite variables described earlier. For example, for P4HFAMIL, composite values are as follows:

- 1=two parents and sibling(s)
- 2=two parents, no siblings
- 3=one parent and sibling(s)
- 4=one parent, no siblings
- 5=other

Parent "figures" were placed in the "other" category for this composite. Likewise, for the composite P4HPARNT, parent "figures" were placed in categories 8 or 9 for related and unrelated guardians, respectively. Similarly, parent "figures" were included in the category 'no resident mother' for P4HMOM and 'no resident father' for P4HDAD.

It also should be noted that because the composite construction identifies only one resident mother or one resident father, same sex parents are not readily identified in the composites themselves. Two approaches can be used to identify these couples. First, the user should search the relationship variables (P3REL_1, P4REL_1, etc.) to identify households in which more than one person identifies himself or herself as a father/mother to the focal child. Second, since not all same-sex partners identify themselves as "mother" or "father" to the focal child, the user should also search for households in which the respondent (identified by P3PER_1, P4PER_1, etc.) is the child's parent and the respondent's spouse (identified from P3SPOUSE, P4SPOUSE) is the same sex as the respondent.

There are two sections in the parent interview that asked parent-figure-specific questions:

- PEQ Parent education
- EMQ Employment

Each of these sections may be completed during the parent interview about a different household member but about no more than two household members. Rather than reserve space (mostly unused) for these sections for each of the 17 household members listed in the FSQ section, these sections are repeated two times, with “pointer” or “foreign key” variables that hold the original number of the household member. For instance, if household member #3 is the first person to receive the PEQ section, then the pointer variable P4EDUP1 will hold the value “3,” and the actual PEQ variable P4NDEG_1 will hold the education information specified by the value of P4EDUP1. Table 7-9 identifies the pointer variables.

There were some errors in the way in which the pointers worked for education and employment in spring-first grade. There were two different sets of specifications governing the selection of the individuals who would be considered “parents” for the purpose of collecting education and employment information.

The education pointers were governed primarily by the relationship to the focal child. The mother/father was selected from the “relation to focal child” variable (P4REL_1 to P4REL_17, FSQ130) or the respondent and spouse were selected if there was no mother/father in the household. Consequently, any individual with a “mother” or “father” value for FSQ130 was identified as a parent; any individual not identified as a parent in FSQ130 was not recognized as a parent even if the respondent/parent listed that person as a spouse.

Table 7-9.—Pointers to parent figure questions

Person pointer		Interview item	
P4EDUP1	P4 PEQ010-060 HH PERSON POINTER 1	P4NDEG_1	P4 PEQ010 PERS 1 COMPLETED NEW DEGREE
		P4DEGT_1	P4 PEQ020 PERS 1 DEGREE TYPE COMPLETED
		P4ENR_1	P4 PEQ030 IF PERS 1 ENROLLED IN COURSES
		P4FPT_1	P4 PEQ040 PERS 1 COURSE FULL/PART TIME
		P4TRN_1	P4 PEQ050 IF PERS 1 GETS JOB TRAINING
		P4HRTR_1	P4 PEQ060 PERS 1 HR/WK SPEND ON TRAINING
P4EDUP2	P4 PEQ010-060 HH PERSON POINTER 2	P4NDEG_2	P4 PEQ010 PERS 2 COMPLETED NEW DEGREE
		P4DEGT_2	P4 PEQ020 PERS 2 DEGREE TYPE COMPLETED
		P4ENR_2	P4 PEQ030 PERS 2 ENROLLED IN COURSES
		P4FPT_2	P4 PEQ040 PERS 2 COURSE FULL/PART TIME
		P4TRN_2	PR PEQ050 IF PERS 2 GETS JOB TRAINING
		P4HTR_2	PR PEQ060 PERS 2 HR/WK SPEND ON TRAINING
P4EMPP1	P4 EMQ020-150 HH PERSON POINTER 1	P4PAY_1	P4 EMQ020 PERS 1 HAD PAID JOB LAST WEEK
		P4VAC_1	P4 EMQ030 IF PERS 1 ON LEAVE PAST WEEK
		P4JOB_1	P4 EMQ040 PERSON 1 NUMBER OF ALL JOBS
		P4HRS_1	P4 EMQ050 PERSON 1 HOURS/WK AT ALL JOBS
		P4LOK_1	P4 EMQ060 PERS 1 SOUGHT JOB LAST 4 WEEKS
		P4DO1_1	P4 EMQ070 PERS 1 CHKD W/PUB EMPL AGENCY

Table 7-9.—Pointers to parent figure questions (continued)

Person pointer		Interview item	
P4EMPP2	P4 EMQ020-150 HH PERSON POINTER 2	P4DO2_1	P4 EMQ070 PERS 1 CHKD W/PRIV EMP AGENCY
		P4DO3_1	P4 EMQ070 PERS 1 CHKD W/EMPLOYER DIRECTLY
		P4DO4_1	P4 EMQ070 PERS 1 CHKD W/FRIENDS & REL
		P4DO5_1	P4 EMQ070 PERS 1 PLACED OR ANSWERED ADS
		P4DO6_1	P4 EMQ070 PERS 1 READ WANT ADS
		P4DO7_1	P4 EMQ070 PERS 1 DID SOMETHING ELSE
		P4DOW_1	P4 EMQ080 WHAT PERSON 1 DOING LAST WEEK
		P4TAK_1	P4 EMQ100 PERS 1 JOB AVAILABLE LAST WEEK
		P4OCC_1	P4 EMQ130-50 1ST PERSON OCCUPATION CODE
		P4PAY_2	P4 EMQ020 PERS 2 HAD PAID JOB LAST WEEK
		P4VAC_2	P4 EMQ030 IF PERS 2 ON LEAVE PAST WEEK
		P4JOB_2	P4 EMQ040 PERSON 2 NUMBER OF ALL JOBS
		P4HRS_2	P4 EMQ050 PERSON 2 HOURS/WK AT ALL JOBS
		P4LOK_2	P4 EMQ060 PERS 2 SOUGHT JOB LAST 4 WEEKS
		P4DO1_2	P4 EMQ070 PERS 2 CHKD W/PUB EMPL AGENCY
		P4DO2_2	P4 EMQ070 PERS 2 CHKD W/PRIV EMP AGENCY
		P4DO3_2	P4 EMQ070 PERS 2 CHKD W/EMPLOYER DIRECTLY
		P4DO4_2	P4 EMQ070 PERS 2 CHKD W/FRIENDS & REL
		P4DO5_2	P4 EMQ070 PERS 2 PLACED OR ANSWERED ADS
		P4DO6_2	P4 EMQ070 PERS 2 READ WANT ADS
		P4DO7_2	P4 EMQ070 PERS 2 DID SOMETHING ELSE
		P4DOW_2	P4 EMQ080 WHAT PERSON 2 DOING LAST WEEK
		P4TAK_2	P4 EMQ100 PERS 2 JOB AVAILABLE LAST WEEK
		P4OCC_2	P4 EMQ130-50 2ND PERSON OCCUPATION CODE

Conversely, employment pointers were governed primarily by the relationship to the respondent. The respondent and spouse were selected if the respondent was a parent; otherwise the mother/father was selected. As a result, any individual who was listed as a "mother" or "father" in FSQ130, but was not listed as the spouse of the other parent (who is the respondent), was not be recognized as the other parent of the focal child. Whether the parent was still a current member of the household was not checked in the employment section.

The cases that follow are an illustration of the problems encountered with the pointers.

- In case 1, the mother was person 1, and her boyfriend was person 5. Because the mother reported the boyfriend as her spouse, the employment pointers were correctly set to 1 and 5. The education pointer was only set for the mother, since the relationship of her reported spouse to the focal child (as reported in FSQ130) was "partner's boyfriend" rather than "father."
- In case 2, the mother and father were persons 1 and 3 in the prior round but left the household as of spring-first grade. The grandparents (persons 7 and 8) were the new parent figures. Because the logic creating the employment pointers did not check to ensure that the former mother and father figures were still household members, the employment pointers were incorrectly set to 1 and 3. Because the algorithm that created the education pointers checked the household member status of the reported parents, the education pointers were correctly set to 7 and 8.
- In case 3, mother and father were persons 1 and 6. Because the father (who is the respondent) did not report person 1 (the mother) as his spouse, the employment pointer was set only for the father, even though FSQ130 indicated that person 1 was the focal child's mother. Because the education pointers relied on FSQ130 to indicate parent figure status, they were correctly set to 1 and 6.

These problems were handled by setting inappropriate data to missing. As explained in section 7.4.2, missing education and employment values were later imputed.

7.4.8 Supplemental Section in the Spring-First Grade Parent Interview

The Supplemental Section (SPQ) section of the parent interview was administered in spring-first grade to parents who had been previous round nonrespondents or who were added as a result of sample "freshening" (see chapter 4 for details on sample freshening). The SPQ was constructed to allow respondents participating for the first time in spring-first grade to provide data for some of the more vital items from the fall-kindergarten parent interview, rather than respond to the entire fall-kindergarten instrument. Table 7-10 shows the variable names for these items in each round of the study.

The topics covered in the spring first-grade SPQ section included child care in the year before kindergarten, Head Start participation, child health and health services received, marital status of the biological parents at the time of the child's birth, the primary language spoken in the home, country of the child's birth, and receipt of Women, Infants, and Children (WIC) benefits.

Relevant items from the SPQ section were used to construct composites for families new to the survey in spring-first grade. If the family was not new to the study in the spring-first grade, values derived for composites during the kindergarten year were copied forward into the spring-first grade composites. Composites that used the SPQ data included whether the child had child care before kindergarten (WKCAREPK), the primary language other than English spoken in the child's home (W1LANGST), whether the biological mother was married to the biological father at the time of the child's birth (W1MOMAR), and whether the mother worked between the child's birth and the start of kindergarten (W1HEARLY).

A few differences between composites from the kindergarten year of the study and those that are formed in first grade from variables in the SPQ section should be noted. One difference in definition concerns the composite WKCAREPK, which indicates whether the child was in child care with a relative, nonrelative, in Head Start, or in a center-based program the year before entering kindergarten. In fall-kindergarten, the composite variable was based on items within the child care section (CCQ025, 130, 215, or 280) that asked about nonparental care the child received on a regular basis before kindergarten. In spring-first grade, the data about prekindergarten arrangements were based on child care items in the SPQ section (SPQ010, SPQ020, SPQ030, and SPQ040). If the child was reported to be in child care on a regular basis with a relative (SPQ010), a nonrelative (SPQ020), Head Start (SPQ030) or in a day care center, nursery school or preschool (SPQ040) the year before kindergarten (the answer was "yes" to any of these questions), WKCAREPK was equal to 1 (Yes). Otherwise, if any of the four variables was unknown, WKCAREPK was coded as -9 (Not Ascertained). If the respondent indicated that the child had not received any of the four types of care in the year before kindergarten, then WKCAREPK was coded as 2 (No).

In fall-kindergarten, the composite was calculated somewhat differently than it was in spring-first grade because additional questions were used which asked if the child had ever had a particular type of care. These questions were CCQ120 (ever have nonrelative care), CCQ015 (ever had relative care), CCQ265 (ever had center care), and CCQ210 (ever in Head Start). In fall-kindergarten, if any of CCQ025, 130, 215, or 280 (questions on nonparental care on a regular basis) was equal to 1 (Yes), WKCAREPK was coded as 1 (Yes). Otherwise, if any of the four variables was unknown (8 or 9) and the corresponding "ever have care" question CCQ015 (ever had relative care) and CCQ265 (ever had center care), CCQ.210 (ever in Head Start)) was not equal to 2 (No), WKCAREPK was coded as -9 (Not Ascertained). Otherwise, WKCAREPK was coded as 2 (No). If the variable about whether the child was in a particular type of care the year before kindergarten was missing (e.g., CCQ130 about nonrelative care) and the variable about ever having been in a particular type of care (e.g., CCQ120) was not equal to 2 (No), WKCAREPK was coded as -9 (Not Ascertained).

Although the spring-first grade version of WKCAREPK and the fall-kindergarten version are very similar, it should be noted that the composite created in spring-first grade will not reflect the child care experiences of any new children added to the sample (the freshened sample) or prior round nonrespondents who received the SPQ section and did not have child care the year prior to kindergarten, but did have child care prior to the year before kindergarten. This is because the fall-kindergarten composite captured child care both the year before kindergarten and prior to that year, while the spring-first grade version of the composite did not.

Another variable that is calculated differently in first grade than in kindergarten is the variable for the primary language spoken at home (W1LANGST). It was formed from two variables in spring-first grade (SPQ155 "Is any language other than English spoken in your home" and SPQ157 "What is the primary language spoken in your home") rather than three variables as it was in fall-kindergarten (PLQ020 "Is any language other than English spoken in your home;" PLQ030 "Is English also spoken in your home," and PLQ040 "What is the primary language spoken in your home"). In fall-kindergarten, if PLQ020 was equal to 2 (no other language than English regularly spoken in home) then WKLANGST was 2 for English. If PLQ020 was missing then WKLANGST was missing. If PLQ020 was 1 (a language other than English was regularly spoken in the home) and PLQ030 was 2 (English was not also spoken in the home), then WKLANGST was 1 for non-English. If PLQ020 was 1 (a language other than English was regularly spoken in the home) and PLQ060 was 0 (primary language in home was English), then WKLANGST was 2 for English. Otherwise, if PLQ020 was 1 (a language other than English was regularly spoken in the home) and PLQ060 indicated any non-English language, then WKLANGST was 1 for non-English. The specification for the spring-first grade composite was the same

Table 7-10.—Supplemental section item matches to spring-first grade items

	Fall-kindergarten		Spring-kindergarten		Fall-first grade		Spring-first grade	
	Variable name	Item number	Variable name	Item number	Variable name	Item number	Variable name	Item number
Child care in year prior to kindergarten								
Relative care	P1RPREK	CCQ025	P2RPREK	SPQ010	P3RPREK	SPQ010	P4RCBEFK	SPQ010[1]
Nonrelative care	P1NPREK	CCQ130	P2NPREK	SPQ020	P3NPREK	SPQ020	P4NCBEFK	SPQ020[1]
Head Start	P1HSPREK	CCQ215	P2HSPREK	SPQ030	P3HSPREK	SPQ030	P4HSBEFK	SPQ030[1]
Center care	P1CPREK	CCQ280	P2CPREK	SPQ040	P3CPREK	SPQ040	P4DBEFK	SPQ040[1]
Status at birth								
Birth weight: pounds	P1WEIGHP	CHQ005	P2WEIGHP	SPQ060	P3WEIGHP	SPQ060	P4WEIGHP	SPQ060
Birth weight: ounces	P1WEIGHO	CHQ005	P2WEIGHO	SPQ065	P3WEIGHO	SPQ065	P4WEIGHO	SPQ065
More than 5.5 pounds	P1WEIGH5	CHQ010	P2WEIGH5	SPQ070	P3WEIGH5	SPQ070	P4WEIGH5	SPQ070
More than 3 pounds	P1WEIGH6	CHQ015	P2WEIGH6	SPQ080	P3WEIGH6	SPQ080	P4WEIGH6	SPQ080
More than 2 weeks early	P1PREMAT	CHQ025	P2PREMAT	SPQ090	P3PREMAT	SPQ090	P4PREMAT	SPQ090
How premature: number	P1EARLY	CHQ030	P2EARLY	SPQ100	P3EARLY	SPQ100	P4EARLY	SPQ100
How premature: unit	P1EARDAY	CHQ030	P2EARDAY	SPQ105	P3EARDAY	SPQ105	P4EARDAY	SPQ105
Vision and hearing difficulties								
Difficulty seeing far obj.	P1SIGHT	CHQ285	P2SIGHT	SPQ107	P3SIGHT	SPQ107	P4SIGHT	CHQ270
Vision correctable	P1CORREC	CHQ315	P2CORREC	SPQ108	P3CORREC	SPQ108	P4CORREC	CHQ315
Difficulty hearing speech	P1DIFFHR	CHQ230	P2DIFFHR	SPQ109	P3DIFFHR	SPQ109	P4DIFFHR	CHQ190
Receipt of services in year prior to kindergarten								
Receipt of any services	P1THERAP	CHQ340	P2THERAP	SPQ110	P3THERAP	SPQ110	P4THERAP	SPQ110
Speech therapy	P1THERA2	CHQ345A	P2THERA2	SPQ120A	P3THERA2	SPQ120A	P4THERA2	SPQ120A
Occupational therapy	P1THERA3	CHQ345B	P2THERA3	SPQ120B	P3THERA3	SPQ120B	P4THERA3	SPQ120B
Vision services	P1THERA5	CHQ345D	P2THERA5	SPQ120D	P3THERA5	SPQ120D	P4THERA5	SPQ120D
Social work services	P1THERA6	CHQ345E	P2THERA6	SPQ120E	P3THERA6	SPQ120E	P4THERA6	SPQ120E
Psychological services	P1THERA7	CHQ345F	P2THERA7	SPQ120F	P3THERA7	SPQ120F	P4THERA7	SPQ120F
Home visits	P1THERA8	CHQ345G	P2THERA8	SPQ120G	P3THERA8	SPQ120G	P4THERA8	SPQ120G
Parent support	P1THERA9	CHQ345H	P2THERA9	SPQ120H	P3THERA9	SPQ120H	P4THERA9	SPQ120H
Special needs	P1THER10	CHQ345I	P2THER10	SPQ120I	P3THER10	SPQ120I	P4THER10	SPQ120I
Private tutoring	P1THER11	CHQ345J	P2THER11	SPQ120J	P3THER11	SPQ120J	P4THER11	SPQ120J
Braille	P1THER12	CHQ345K	P2THER12	SPQ120K	P3THER12	SPQ120K	P4THER12	SPQ120K
Sign language	P1THER13	CHQ345L	P2THER13	SPQ120L	P3THER13	SPQ120L	P4THER13	SPQ120L
Age therapy began	P1THER14	CHQ375	P2THER14	SPQ130A	P3THER14	SPQ130A	P4THER16	SPQ130A[2]
Age therapy began: Unit	P1THER15	CHQ375U	P2THER15	SPQ130B	P3THER15	SPQ130B	P4THER17	SPQ130B[2]
Family background								
Biolog. parents married	P1BIOLOG	MHQ060	P2BIOLOG	SPQ150	P3BIOLOG	SPQ150	P4BIOLOG	SPQ150
Other language at home	P1ANYLNG	PLQ020	P2ANYLNG	SPQ155	P3ANYLNG	SPQ155	P4ANYLNG	SPQ155
Primary lang. at home	P1PRMLNG	PLQ060	P2PRMLNG	SPQ157	P3PRMLNG	SPQ157	P4PRMLNG	SPQ157
Resp/bio mom job for pay	P1MOMWRK	EMQ170	P2MOMWRK	SPQ200	P3MOMWRK	SPQ200	P4MOMWRK	SPQ200
WIC benefits: pregnant	P1WICMOM	WPQ030	P2WICMOM	SPQ210	P3WICMOM	SPQ210	P4WICMOM	SPQ210
WIC benefits for child	P1WICCHD	WPQ040	P2WICCHD	SPQ220	P3WICCHD	SPQ220	P4WICCHD	SPQ220

[1] The text of SPQ010-040 was changed slightly in round 4; the phrase “the year before (he/she) started kindergarten” was replaced by “the year before (he/she) started kindergarten/in 1998”.

[2] SPQ130A-130B Asks the earliest age at which each service (120A-120L) began but does not have the same answer categories for responses. In fall-kindergarten, spring-kindergarten, and fall first grade, SPQ130A asked for entry of age and SPQ130B asked for entry of the unit (months or years). In spring-first grade, SPQ130A asked for entry in years and SPQ13OB asked for entry of months.

(SPQ155 and SPQ157 are equivalent to PLQ020 and PLQ060), however, because there was no equivalent question in SPQ for PLQ30 (is English also spoken in your home) the definition was slightly different. Despite the differences in question structure, the kindergarten and first grade composites capture the same information.

Another change in the SPQ derived composites involves W1MOMAR (biological mother was married to the biological father at the time of the child's birth). The kindergarten versions of this composite were WKHMOMAR and WKBMOMAR. These variables indicated whether or not the biological mother was married to the biological father at the time of the child's birth. WKHMOMAR was coded if the biological mother resided in the house at the time of the interview; WKBMOMAR was coded if the biological mother resided outside the house at the time of the interview. The variables were mutually exclusive; that is, if there was a nonmissing value for WKHMOMAR, the variable WKBMOMAR would be set to inapplicable (-1), and the converse was also true. Only one composite was used in spring-kindergarten to represent this information because with four rounds of data it did not make sense to define what round would determine the presence or absence of the biological mother. In order to create two separate variables based on the presence or absence of the biological mother in the household during spring first grade, a user can use the variable P4MOMTYP in conjunction with W1MOMAR.

Another difference between the variable W1MOMAR and the WKHMOMAR/ WKBMOMAR combination is that, while many different marital status variables were used to create the composites in fall-kindergarten (marital information from MHQ020, MHQ120, MHQ125, MHQ060, MHQ025, MHQ130, MHQ065; and the child's date of birth CHILDDOB), only one question (SPQ150) was used in the SPQ sections in spring-first grade to determine the same information (whether his/her biological mother and father were married when the child was born) as the question was asked directly. In fall-kindergarten, the date of marriage was determined and then compared to the date of the child's birth. If the date of marriage preceded the birth, the composite was coded as 1 (Yes). If the date of birth preceded the year of marriage, the composite was coded as 2 (No). If information for either date was missing, the composite was missing. However, the differences in items used to create the composite did not lead to any differences in how the composite was defined.

Finally, the composite W1HEARLY (mother worked between the child's birth and kindergarten) was also calculated somewhat differently in the first grade based on how the question was asked. In the base year of the study, the question about the mother working was only asked if there was a mother in the household. If there was no mother in the household, WKHEARLY was set to -1. In spring-first grade, it did not matter for the purposes of collecting retrospective data if there was a mother in the household at the time of spring-first grade. Thus, if there was an answer to the SPQ question in spring-first grade (P4MOMWRK, SPQ200), it was used regardless of whether there was a mother in the household. However, if a respondent indicated while answering SPQ200 that there was no mother in the household between the child's birth and kindergarten, then W1HEARLY was coded -1.

(The "cross-round" composites from the SPQ section are useful because they provide an opportunity to more easily examine some 'key' items from nonrespondents and maximize the number of cases with these data. For example, if a user wanted to examine child care in the year before kindergarten for children who entered in fall-kindergarten and children in the freshened sample in spring-first grade, he or she could look at questions about child care by using the variables shown in table 7-10. As shown in the first category heading in table 7-10 (child care in year prior to kindergarten), the variables for relative care, nonrelative care, Head Start, and center-based child care are P1RPREK, P1NPREK, P1HSPREK, P1CPREK (fall-kindergarten variables) and P4RCBEFK, P4NCBEFK, P4HSBEFK, and P4DBEFK (spring-first grade variables). The fall-kindergarten variables about child care before kindergarten were from the child care section of the kindergarten questionnaire (CCQ) and the spring-kindergarten questions were from the supplemental section (SPQ) of the first grade questionnaire.)

7.4.9 Industry and Occupation Codes Used in ECLS-K

This section describes the aggregated categories that were used for coding occupation in the ECLS-K.

- **1. Handlers, Equipment Cleaners, Helpers, and Laborers**

  This category includes occupations that involve helping other workers and performing routine nonmachine tasks. A wide variety of helpers, handlers, etc., are included in this category. Examples include construction laborers, freight, stock, and material movers, garage and service station related occupations, parking lot attendants, and vehicles washers and equipment cleaners.

- **2. Production Working Occupations**

  This category includes occupations concerned with setting up, operating, and tending of machines and hand production work usually in a factory or other fixed place of business.

- **3. Service Occupations**

  The category includes occupations providing personal and protective services to individuals, and current maintenance and cleaning for building and residences. Some examples include food service, health service (e.g., aides or assistants), cleaning services other than household, and personal services.

- **4. Agricultural, Forestry, and Fishing Occupations**

  This category is concerned with the production, propagation (breeding/growing), gathering, and catching of animals, animal products, and plant products (timber, crop, and ornamental); the provision of services associated with agricultural production; and game farms, fisheries, and wildlife conservation. "Other agricultural and related occupations" include occupations concerned with the production and propagation of animals, animal products, plants and products (crops and ornamental).

- **5. Marketing and Sales Occupations**

  This category includes occupations involving selling goods or services, purchasing commodities and property for resale, and conducting wholesale or retail business.

- **6. Transportation and Material Moving Occupations**

  This category includes occupations concerned with operating and controlling equipment used to facilitate the movement of people or materials and the supervising of those workers.

- **7. Precision Production Occupations**

  Precision production includes occupations concerned with performing production tasks that require a high degree of precision or attainment of rigid specification and operating plants or large systems. Examples are tool and die makers, pattern and model makers, machinists, jewelers, engravers, and so on. Also included are some food-related occupations including butchers and bakers. Plant and system operators include water and sewage, gas, power, chemical, petroleum, and other plant or system operators.

- **8. Administrative Support Occupations, including Clerks**

  This category includes occupations involving preparing, transcribing, transferring, systematizing, and preserving written communications and records; collecting accounts; gathering and distributing information; operating office machines and data processing equipment; operating switchboards; distributing mail and messages; and other support and clerical duties such as bank teller, data entry keyer, etc.

- **9. Mechanics and Repairers**

  Mechanics and repairers are persons who do adjustment, maintenance, part replacement, and repair of tools, equipment, and machines. Installation may be included if installation is usually done in conjunction with other duties of the repairers.

- **10. Construction and Extractive Occupations**

  This category includes occupations that normally are performed at a specific site, which will change over time, in contrast to production workers, where the work is usually at a fixed location. Construction workers include those in overall construction, brickmasons, stonemasons, carpenters, electricians, drywall installers, paperhangers and painters, etc. Extractive occupations include oil well drillers, mining machine operators, and so on.

- **11. Technologists and Technicians, except Health**

  This category includes those providing technical assistance in engineering and scientific research, development, testing, and related activities, as well as operating and programming technical equipment and systems.

- **12. Writers, Artists, Entertainers, and Athletes**

  This category includes occupations concerned with creating and executing artistic works in a personally interpreted manner by painting, sculpturing, drawing, engraving, etching, and other methods; creating designs for products and interior decorations; designing and illustrating books, magazines, and other publications; writing; still, motion picture and television photography/filming; producing, directing, staging, acting, dancing, singing in entertainment; and participating in sports and athletics as competitor or player and administering and directing athletic programs.

- **13. Executive, Administrative, and Managerial Occupations**

  This category includes senior-level and middle management occupations and occupations that directly support management. Senior-level managers are persons concerned with policy making, planning, staffing, directing, and/or controlling activities. Middle managers include persons who plan, organize, or direct and/or control activities at the operational level. Legislators are also included in this category. Workers in this category are not directly concerned with the fabrication of products or with the provision of services. Other officials and administrators include consultants, library directors, customhouse builders, and location managers.

- **14. Health Technologists and Technicians**

  This category includes occupations concerned with providing technical assistance in the provision of health care. For example, clinical laboratory technologists and technicians, dental hygienists, radiologic technicians, licensed practical nurses (LPNs), and other health technologists are included here.

- **15. Social Scientists, Social Workers, Religious Workers, and Lawyers**

  This category includes occupations concerned with the social needs of people and in basic and applied research in the social sciences.

- **16. Registered Nurses, Pharmacists, Dieticians, Therapists, and Physician's Assistants**

  This category includes occupations concerned with the maintenance of health, the prevention of illness and the care of the ill through the provision and supervision of nursing care; compounding drugs, planning food service or nutritional programs; providing assistance to physicians; and the provision of therapy and treatment as directed by physicians.

- **17. Natural Scientists and Mathematicians**

  This category includes those engaged primarily in the application of scientific principles to research and development. Natural scientists are those in the physical sciences (e.g., chemistry, physics) and the life sciences (e.g., biology, agriculture, medicine). In addition, this category includes those in computer science, mathematics (including statistics), and operations research.

- **18. Teachers, except Postsecondary Institution**

  This category includes prekindergarten and kindergarten teachers, elementary and secondary teachers, special education teachers, instructional coordinators, and adult education teachers (outside postsecondary).

- **19. Engineers, Surveyors, and Architects**

  The category includes occupations concerned with applying principles of architecture and engineering in the design and construction of buildings, equipment and processing systems, highways and roads, and land utilization.

- **20. Teachers: College, University, and Other Postsecondary Institution; Counselors, Librarians, and Archivists**

  This category includes those who teach at higher education institutions and at other postsecondary (after high school) institutions, such as vocational institutes. In addition, vocational and educational counselors, librarians, and archivists are included here.

- **21. Physicians, Dentists, and Veterinarians**

  This category includes health care professionals who diagnose and treat patients. In addition to physicians, dentists, and veterinarians, this category includes optometrists, podiatrists, and other diagnosing and treating professionals, such as optometrists, podiatrists, chiropractors, hypnotherapists, and acupuncturists.

7.5 Relationship Among Composite Variables on the Child Catalog

The child catalog contains all child, teacher, and school composites, but it should be noted that some of the teacher variables on the child catalog are slightly altered when they are brought from the teacher questionnaire to the child level for children retained in kindergarten in the 1999-2000 school year. While some teacher-level variables are class specific when collected, i.e., for the kindergarten version of TQA), there are three versions of each of these class-specific variables (AM, PM, and all day), the child file will contain data from only one of each of the three variables—depending on whether the child attends an AM, PM, or all day class. For example, there are three variables (one for AM, one for PM, and one for all day classes) that represent percent of African Americans in the teacher's class. The teacher may provide information for more than one type of class if she teachers more than one class. Because the child can only be in one class, there is only one variable on the child record for percent African American in class, that which is appropriate for the class he/she attended. The same is true for percent Hispanic and percent minority variables. The variable A4CLASS is used to identify which of the teacher-reported classes was used for the child. See section 7.4.3 for details on the A4CLASS variable.

7.6 Children Who Changed Schools

There are several variables on the file that can be used to determine if a child moved to a different school across rounds. The variables S1_ID, S2_ID, S3_ID and S4_ID are school identification numbers that indicate what school the child was in at the end of a round of data collection. If we did not know where the child was at the end of the round of data collection, we used the scheme shown in table 7-11 for assigning ID numbers.

By comparing these variables, users can determine whether the child physically moved from one school to another and in which round. Some children will have moved more than once over the course of the first 2 years. These variables store the school at the end of the round only: if a child moved from School A to School B and then to School C in the same round, the school ID variable will contain School C; the brief stay in School B is not represented in the data.

Table 7-11.—Case status and identification numbers for children not followed or located

Case status	Sn_ID
Unlocatable. (Field staff were unable to locate a transfer student in his/her new school.)	9995
End of field period. (Information on the transfer student's new school was identified too late in the field period for the case to be re-fielded for the assessment.)	9996
Moved to non-sampled PSU. (The transfer student enrolled in a school that was outside of ECLS-K's sampled PSUs--field staff did not attempt to collect the assessment, but did attempt to collect the parent interview.)	9997
Do not follow. (The transfer student was flagged by the statisticians as "do not follow" because of subsampling of transfer students due to cost constraints- if the child moved from his/her original school, field staff did not "follow" him/her" to the new school and did not collect a child assessment or parent interview.)	9998

NOTE: S3_ID was set to "-1" (Not Applicable) if the child was not included in the fall-first grade subsample.

Children moved between schools for a variety of reasons, but one factor was that a school terminated at kindergarten and most of the students went to first grade at another particular school. This is known as a "destination school" and the move is known as a "destination move." Destination schools were schools where it was determined before data collection that at least four ECLS-K children would move into them from a kindergarten-terminal school or a school that had closed. The variables on the file that indicate destination moves are:

- R3DEST "MOVED TO FALL-FIRST GRADE DESTINATION SCHOOL"
- R4DEST "MOVED TO SPRING-FIRST GRADE DESTINATION SCHOOL"

It should be noted that the destination school may also have been an originally sampled school; in this case, the school was a destination school only for the new students, not for the originally sampled students. Because only 30 percent of all students were included in the fall-first grade data collection, a destination move could have occurred between spring-kindergarten and fall-first grade (for the 30 percent included in fall-first grade), between fall-first grade and spring-first grade (for the same 30 percent in fall-first grade), or between spring-kindergarten and spring-first grade (for the other 70 percent). The variable R3DEST was set to 1 if the child moved between spring-kindergarten and fall-first grade and the move was considered to be a destination move, and the variable R4DEST was set to 1 if a move between spring-kindergarten and spring-first grade, or between fall-first grade and spring-first grade, occurred and was considered to be a destination move.

Other variables on the file that will be of interest to users examining school changes are R4R2SCHG (school type change between spring-kindergarten and spring-first grade) and R4R3SCHG (school type change between fall-first grade and spring-first grade). These indicate whether the child changed schools and, if so, what the school type was in the previous and new school (e.g., whether the change was from public to private school, private to private school, etc.).

7.7 Merging Base Year School-Level Data with the First Grade Child-Level Data

As discussed in chapter 2, section 2.7, two versions of the school administrator questionnaire (SAQ) were distributed in the spring-first grade round. Schools participating in the base year were asked to complete the “returning” school questionnaire. In order to reduce the burden on returning school administrators within the first grade school year, the “returning” school questionnaire collected fewer details on school and student characteristics, school facilities and resources, school policies and programs, and school and community connections. According to the data collection plan, schools that were new to the ECLS-K study, because sampled children transferred into them, or that had been in the study but had not completed an SAQ previously would be asked to complete the “new” school questionnaire. Because these schools were new to the study, the “new” school questionnaire asked for more details about the topics described above and was similar to the spring-kindergarten SAQ.

In practice, the “new” and “returning” school questionnaires were distributed on the basis of whether the school was an original school or a new one to which students had transferred. It should be noted that there were some errors in this process. Out of a total of 2,132 responding schools, 9 original schools completed the “new” SAQ and 11 new transfer schools completed the “returning” SAQ. In addition, the “new” school questionnaire was not systematically administered to schools that did not complete the spring-kindergarten SAQ (112 of these schools received the “returning” school questionnaire). There were also a few cases in which there was a spring-kindergarten SAQ, but the “new” school questionnaire was completed (6 cases).

Because of the differences between the “new” and “returning” school questionnaires, there are variables on the data file that contain data from the new schools, but not from the returning schools. In order to use data from these variables in analysis for all of the schools, not just the new schools, data from returning schools collected in the base year must be pulled forward. To conduct analyses using all school data, analysts must merge the base year school variables that were not collected in spring-first grade with the first grade school variables for returning schools**.** The base year school catalog has been included with the first grade child catalog to facilitate merging the SAQ variables from both rounds of data collection. Table 7-12 lists the variables from the spring-kindergarten school questionnaire that do not appear in the spring-first grade returning school questionnaire.

The SAQ includes school administrator characteristics. School administrators changed in some returning schools between the kindergarten and first grade school years. In the first grade returning SAQ, if the school administrator was the same as in kindergarten, then the school administrator skipped the school administrator characteristics in order to reduce burden. If the returning school administrator was new in first grade, then he or she answered the school administrator characteristic items. Thus, analysts must decide which school administrator characteristics to use (i.e., those for the kindergarten school administrator or those for the first grade school administrator). First grade school composites were created for all schools regardless of which questionnaire was administered. The school administrator characteristic variables appear in table 7-12 that follows under the heading "School Administrator Characteristics."

Data from the spring-kindergarten questions listed in the table 7-12 are in the school catalog on the base year ECB.[6] To merge these base year school data with the first grade child-level school data, analysts can:

- Select the variables to be analyzed from the base year school catalog and the variable S_ID. This creates a "working taglist" (see section 8.4 for more detail on how to create a working taglist);
- Use the Extract function to create the extraction program. (refer to chapter 8, section 8.5);
- Run the program generated after extraction to create a school-level data set (DATA1);
- Rename the S_ID to S4_ID in order to merge the data;
- Use the first grade child catalog to select the variables to be analyzed and the variable S4_ID;
- Use the Extract function to create the extraction program. (refer to chapter 8, section 8.5);

[6] ECLS-K Base Year Public-Use User's Manual (NCES 2001-029), October 2000.

Table 7-12.—Variables from the spring-kindergarten school administrator questionnaire that were included in the first grade new school administrator questionnaire but not the first grade returning school administrator questionnaire

Variable name	Variable label
S2UNGRAD	S2 Q4 GRADE LEVEL-UNGRADED
S2SPLNDS	S2 Q4 GRADE LEVEL-SPECIAL NEEDS
S2PRKNDR	S2 Q4 GRADE LEVEL-PREKINDERGARTEN
S2TRANS	S2 Q4 GRADE LEVEL-TRANSITIONAL
S2KINDER	S2 Q4 GRADE LEVEL-KINDERGARTEN
S2PRE1	S2 Q4 GRADE LEVEL-PREFIRST
S2GRADE1	S2 Q4 GRADE LEVEL-FIRST GRADE
S2SECOND	S2 Q4 GRADE LEVEL-SECOND GRADE
S2THIRD	S2 Q4 GRADE LEVEL-THIRD GRADE
S2FOURTH	S2 Q4 GRADE LEVEL-FOURTH GRADE
S2FIFTH	S2 Q4 GRADE LEVEL-FIFTH GRADE
S2SIXTH	S2 Q4 GRADE LEVEL-SIXTH GRADE
S27TH	S2 Q4 GRADE LEVEL-SEVENTH GRADE
S28TH	S2 Q4 GRADE LEVEL-EIGHTH GRADE
S2NINTH	S2 Q4 GRADE LEVEL-NINTH GRADE
S2TENTH	S2 Q4 GRADE LEVEL-TENTH GRADE
S211TH	S2 Q4 GRADE LEVEL-ELEVENTH GRADE
S212TH	S2 Q4 GRADE LEVEL-TWELFTH GRADE
S2PUBLIC	S2 Q5 IS IT A PUBLIC SCHOOL
S2REGSKL	S2 Q6A IS IT REGULAR PUBLIC SCHOOL
S2MAGSKL	S2 Q6B IS IT A MAGNET SCHOOL
S2CHCESK	S2 Q6C IS IT SCHOOL OF CHOICE
S2BIASKL	S2 Q6D IS IT A TRIBAL SCHOOL
S2SPEDSK	S2 Q6E IS IT A SPECIAL ED SCHOOL
S2EARLCH	S2 Q6F IS IT AN EARLY CHILDHOOD CTR
S2CATHOL	S2 Q7A1 IS IT A CATHOLIC SCHOOL
S2DIOCES	S2 Q7A2 IS IT A DIOCESAN SCHOOL
S2PARISH	S2 Q7A3 IS IT A PARISH SCHOOL
S2PRIVRD	S2 Q7A4 IS IT A PRIVATE ORDER
S2OTHREL	S2 Q7B PRIVATE, OTHER RELIGIOUS
S2NAISKL	S2 Q7C PRIVATE-ACCREDITED BY NAIS
S2OTHPRI	S2 Q7D IS IT OTHER PRIVATE
S2PVTSPD	S2 Q7E IS IT SPECIAL EDUCATION
S2PVTEAR	S2 Q7F IS IT AN EARLY CHILDHOOD CENTER
S2FOCUS	S2 Q8 DOES SCHOOL HAVE A FOCUS
S2EMPHAS	S2 Q9 WHAT'S THE SCHOOL EMPHASIS
S2PREK	S2 Q10 PRE-K PARTICIPATES
S2REDDYN	S2 Q10 TRANSITIONAL PARTICIPATES
S2KINGAR	S2 Q10 KINDERGARTEN PARTICIPATES
S2PRE1ST	S2 Q10 PREFIRST PARTICIPATES
S2FIR1ST	S2 Q10 1ST GRADE PARTICIPATES
S2SEC2ND	S2 Q10 2ND GRADE PARTICIPATES
S2THI3RD	S2 Q10 3RD GRADE PARTICIPATES

Table 7-12.—Variables from the spring-kindergarten school administrator questionnaire that were included in the first grade new school administrator questionnaire but not the first grade returning school administrator questionnaire (continued)

Variable name	Variable label
S2FOR4TH	S2 Q10 4TH GRADE PARTICIPATES
S2FIF5TH	S2 Q10 5TH GRADE PARTICIPATES
S2SIX6TH	S2 Q10 6TH GRADE PARTICIPATES
S2SEV7TH	S2 Q10 7TH GRADE PARTICIPATES
S2ATE8TH	S2 Q10 8TH GRADE PARTICIPATES
S2NIN9TH	S2 Q10 9TH GRADE PARTICIPATES
S2TEN10	S2 Q10 10TH GRADE PARTICIPATES
S2ELEV11	S2 Q10 11TH GRADE PARTICIPATES
S2TWEL12	S2 Q10 12TH GRADE PARTICIPATES
S2AFUND	S2 Q23A FROM STATE COMPENSATORY FUNDS
S2BFUND	S2 Q23B FROM COMMUNITY FUND RAISING
S2CFUND	S2 Q23C FROM PTA FUND RAISING
S2DFUND	S2 Q23D FROM LOCAL/NATIONAL BUSINESS
S2EFUND	S2 Q23E FROM SPECIAL ED PROGRAM/AGENCY
S2FFUND	S2 Q23F FROM AUXILIARY/AFFILIATED SER
S2GFUND	S2 Q23G FROM MEDICAID
S2HFUND	S2 Q23H FROM IMPACT AID
S2IFUND	S2 Q23I FROM BILINGUAL AID
S2JFUND	S2 Q23J FROM MIGRANT AID
S2KFUND	S2 Q23K FROM OTHER GRANTS
S2CHLDNM	S2 Q24 # OF CHILDREN SITE ACCOMMODATES
S2CMNITY	S2 Q30 COMMUNITY SCHOOL IS LOCATED IN
S2NOTEST*	S2 Q37 NO GRADE TESTED
S2TESTPK*	S2 Q37 PRE-KINDERGARTEN TESTED
S2TESTR*	S2 Q37 TRANSITIONAL-K TESTED
S2TESTK*	S2 Q37 KINDERGARTEN TESTED
S2TESTP1*	S2 Q37 PRE-FIRST TESTED
S2TEST1*	S2 Q37 FIRST GRADE TESTED
S2TEST2*	S2 Q37 2ND GRADE TESTED
S2TEST3*	S2 Q37 THIRD GRADE TESTED
S2TEST4*	S2 Q37 FOURTH GRADE TESTED
S2TEST5*	S2 Q37 FIFTH GRADE TESTED
S2TEST6*	S2 Q37 SIXTH GRADE TESTED
S2TEST7*	S2 Q37 SEVENTH GRADE TESTED
S2TEST8*	S2 Q37 EIGHTH GRADE TESTED
S2TEST9*	S2 Q37 NINTH GRADE TESTED
S2TEST10*	S2 Q37 TENTH GRADE TESTED
S2TEST11*	S2 Q37 ELEVENTH GRADE TESTED
S2TEST12*	S2 Q37 TWELFTH GRADE TESTED
S2B4SCH*	S2 Q41A OFFER BEFORE-SCHOOL CARE
S2SUMMER*	S2 Q41G OFFER SUMMER SCHOOL PROGRAMS
S2MIGRSY*	S2 Q41H SCHOOL YR MIGRANT PROGRAMS
S2MIGRSM*	S2 Q41I SUMMER MIGRANT PROGRAMS

Table 7-12.—Variables from the spring-kindergarten school administrator questionnaire that were included in the first grade new school administrator questionnaire but not the first grade returning school administrator questionnaire (continued)

Variable name	Variable label
S2HRVIS*	S2 Q41J OFFER HEARING/VISION TESTS
S2CHCARE*	S2 Q41K CHILDCARE FOR PARENT EVENT
S2PTAMT*	S2 Q43A FREQUENCY PTA/PTO MEETINGS
S2NWSHME*	S2 Q43B FREQUENCY NEWS SENT HOME
S2RPRTCD*	S2 Q43C FREQUENCY OF REPORT CARDS
S2PTCONF*	S2 Q43D FREQ OF PARENT-TCHR CONFERENCE
S2HVISIT*	S2 Q43E FREQUENCY OF HOME VISITS
S2INVITE*	S2 Q43F FREQ OF PERFORMANCES FOR PARENTS
S2CLASPR*	S2 Q43G FREQ OF CLASSROOM PROGRAMS
S2FUNDRS*	S2 Q43H FREQUENCY OF FUNDRAISERS
S2WRKSHP*	S2 Q43I FREQUENCY OF WORKSHOPS
S2ONIEP	S2 Q53A SPECIAL EDUCATION ON IEP
S2ON504	S2 Q53B SPECIAL EDUCATION ON 504 PLAN
S2IEP504	S2 Q53C SPECIAL ED ON IEP & 504
S2REGCUR	S2 Q55A IEP IN REG CURRICULUM/1 SUBJECT
S2REGMLA	S2 Q55B IEP IN REG MATH/LANGUAGE ARTS
S2EVLIEP	S2 Q55C IEPS SAME EVAL STANDARDS
S2TSTIEP	S2 Q55D IEPS TAKE STANDARD TESTING
S2NOGFT*	S2 Q56 NO GIFTED/TALENTED PROGRAM
S2GFTR*	S2 Q56 GIFTED/TALENTED IN TRANSITIONAL K
S2GFTK*	S2 Q56 GRADE K GIFTED/TALENTED
S2GFTP*	S2 Q56 PREFIRST GIFTED/TALENTED
S2GFT1ST*	S2 Q56 1ST GRADE GIFTED/TALENTED
S2GFT2ND*	S2 Q56 2ND GRADE GIFTED/TALENTED
S2GFT3RD*	S2 Q56 3RD GRADE GIFTED/TALENTED
S2GFT4TH*	S2 Q56 4TH GRADE GIFTED/TALENTED
S2GFT5TH*	S2 Q56 5TH GRADE GIFTED/TALENTED
S2GFTNBR	S2 Q57 NUMBER GIFTED/TALENTED
S2LOSLRY	S2 Q59 LOWEST ANNUAL BASE SALARY
S2HISLRY	S2 Q60 HIGHEST ANNUAL BASE SALARY
School Administrator Characteristics**	
S2GNDER	S2 Q70 GENDER OF PRINCIPAL
S2BRTHYR	S2 Q71 YEAR PRINCIPAL WAS BORN
S2ORIGIN	S2 Q72 PRINCIPAL IS HISPANIC/LATINO
S2RACE1	S2 Q73A PRINCIPAL IS AMERICAN INDIAN
S2RACE2	S2 Q73B PRINCIPAL IS ASIAN
S2RACE3	S2 Q73C PRINCIPAL IS BLACK
S2RACE4	S2 Q73D PRINCIPAL IS HAWAIIAN OR PAC IS
S2RACE5	S2 Q73E PRINCIPAL IS WHITE
S2YSTCH	S2 Q74A NUMBER OF YRS TEACHING
S2TOTPRI	S2 Q74B NUMBER OF YRS AS PRINCIPAL

Table 7-12.—Variables from the spring-kindergarten school administrator questionnaire that were included in the first grade new school administrator questionnaire but not the first grade returning school administrator questionnaire (continued)

Variable name	Variable label
S2PRINHR	S2 Q74C NUMBER YRS A PRINCIPAL HERE
S2YRPREK	S2 Q75A YRS TEACHING PREK/HEADSTART
S2YRK	S2 Q75B YEARS TEACHING KINDERGARTEN
S2YR1ST	S2 Q75C YEARS TEACHING FIRST GRADE
S2YR2TO5	S2 Q75D YRS TEACHING SECOND-FIFTH GRADE
S2YR6	S2 Q75E YEARS TEACHING SIXTH OR HIGHER
S2YRESL	S2 Q75F YEARS TEACHING ESL PROGRAMS2 Q75G YEARS
S2YRBILG	TEACHING BILINGUAL ED
S2YRSPED	S2 Q75H YEARS TEACHING SPECIAL ED
S2YRPHED	S2 Q75I YEARS TEACHING PHYSICAL ED
S2YRART	S2 Q75J YEARS TEACHING ART OR MUSIC
S2CRSECE	S2 Q76A COURSES IN EARLY CHILDHOOD ED
S2CRSELE	S2 Q76B COURSES IN ELEMENTARY ED
S2CRSSPE	S2 Q76C COURSES IN SPECIAL EDUCATION
S2CRSESL	S2 Q76D COURSES IN ESL
S2CRSCDV	S2 Q76E COURSES IN CHILD DEVELOPMENT
S2CRSMTR	S2 Q76F COURSES IN READING METHODS
S2CRSMTM	S2 Q76G COURSES IN MATH METHODS
S2CRSSCI	S2 Q76H COURSES IN SCIENCE METHODS
S2CRSADM	S2 Q76I COURSES IN ADMINISTRATION
S2EDLVL	S2 Q77 HIGHEST LEVEL OF EDUCATION
S2INSTRU	S2 Q78A HOURS/WEEK WORKING W/TEACHERS
S2INRMGT	S2 Q78B HOURS/WEEK FOR SCHOOL MANAGEMENT
S2DISCAT	S2 Q78C HRS/WEEK FOR DISCIPLINE/ABSENCE
S2MONITR	S2 Q78D HRS/WEEK MONITORING SCHOOL AREAS
S2TEEECH	S2 Q78E HOURS/WEEK SPENT TEACHING
S2TALKPT	S2 Q78F HOURS/WEEK MEETING WITH PARENTS
S2STUDNT	S2 Q78G HOURS/WEEK MEETING WITH STUDENTS
S2PPRWRK	S2 Q78H HOURS/WEEK W/REQUIRED PAPERWORK
S2KNWNME	S2 Q79 ESTIMATE PERCENT OF NAMES KNOWN

* Slightly different version of the question than in the spring-first grade SAQ questionnaire for new schools.

** Questions about school administrator characteristics were in a separate section of the questionnaire and could only be completed by the school administrator.

- Run the program generated after extraction to create a child-level data set (DATA2);
- Sort DATA1 and DATA2 by S4_ID; and
- Merge DATA1 and DATA2 by S4_ID to obtain a merged child-level file.

Once the spring-kindergarten school data has been merged to the spring-first grade child data, the user may need to create new variables combining data from the spring-kindergarten SAQ and the spring-first grade SAQ. For example, if the analyst wants to limit the schools in an analysis to only those offering second grade, the variables S4SECOND (from the child level spring-first grade data) and S2SECOND (from the spring-kindergarten school catalog) will need to be combined, because the S4SECOND variable exists only for schools the received the SAQ for new schools in spring-kindergarten.

7.8 Composite Tables

Tables 7-13 and 7-14 describe the composite and derived variables that are on the ECLS-K child catalog. Please note that a few of the variables specified in the “derived from” column are intermediary variables that were not included in the final data sets. An example of an intermediary variable is the child gender variable from the parent questionnaire, CHILDGEN. If this variable were missing, or had conflicting information across rounds of the study, information about gender was used from the field management system (FMS). The variable CHILDGEN is not included in the final data set, but the composite GENDER is. Other intermediary variables are taken from either the FMS or the sample frame. Items from these are not delivered on the file.

The “derived from” column also contains the item numbers from the questionnaire, which help in identifying the items that were used in the creation of these composites and derived variables. This information allows a user to decide if he or she would like to use the composite based on how it was defined.

Table 7-13.—Fall-first grade composite variables

Variable name	Category	Description	Derived from	Values
R3_AGE	Child	Child's age at the time the direct child assessment occurred. This age will be in months and years, not days.	DOBMO, DOBDA, DOBYR, C3ASMTMM, C3ASMTDD, C3ASMTYY	Continuous
GENDER	Child	Child's gender	CHILDGEN (not on file) and FMS (not on file)	1=Male; 2=Female
DOBMO	Child	Child date of birth month	CHILDDOB (not on file) from first data collection in which reported in parent interview, FMS date of birth variable (not on file)	1-12
DOBDD	Child	Child's date of birth day	CHILDDOB (not on file) from first data collection in which reported in parent interview, and FMS date of birth variable (not on file)	1-31
DOBYY	Child	Child's date of birth year	CHILDDOB (not on file) from first data collection in which reported in parent interview, and FMS date of birth variable (not on file)	
W1RACETH	Child	Race and ethnicity of the focal child	P4RC1_1 through P4RC6_1 up to "_25" (FSQ195)	1=White, 2=Black or African American, 3=Hispanic, race specified, 4=Hispanic, no race specified, 5=Asian, 6=Native Hawaiian or other Pacific Islander, 7=American Indian or Alaska Native, 8=More than 1 race, non-Hispanic
RACE	Child	Child race and ethnicity	WKRACETH, W1RACETH, FMS race variable (not on file)	1=White, 2=Black or African American, 3=Hispanic, race specified, 4=Hispanic, no race specified, 5=Asian, 6=Native Hawaiian or other Pacific Islander, 7=American Indian or Alaska Native, 8=More than 1 race, non-Hispanic
W1AMERIN	Child	Child is Native American	P4RC1_1 through P4RC6_1 up to "_25" (FSQ195)	1=Yes, 2=No
W1ASIAN	Child	Child is Asian	P4RC1_1 through P4RC6_1 up to "_25" (FSQ195)	1=Yes, 2=No
W1PACISL	Child	Child is Pacific Islander	P4RC1_1 through P4RC6_1 up to "_25" (FSQ195)	1=Yes, 2=No

Table 7-13.—Fall-first grade composite variables (continued)

Variable Name	Category	Description	Derived from	Values
W1BLACK	Child	Child is African American	P4RC1_1 through P4RC6_1 up to "_25" (FSQ195)	1=Yes, 2=No
W1WHITE	Child	Child is White	P4RC1_1 through P4RC6_1 up to "_25" (FSQ195)	1=Yes, 2=No
W1HISP	Child	Child is Hispanic	P4HSP_1 through P4HSP_25 (FSQ190)	1=Yes, 2=No
W1MT1RAC	Child	Child is more than one race	FSQ195 and the other specify text (not on file) collected in FSQ198 when FSQ195 = 91 (another race)	1=Yes, 2=No
C3BMI	Child	Fall-first grade Body Mass Index	C3HEIGHT, C3WEIGHT	Continuous
C3HEIGHT	Child	Child's fall-first grade composite height	C3HGT1, C3HGT2	Continuous
C3WEIGHT	Child	Child's fall-first grade composite weight	C3WGT1, C3HGT2	Continuous
P3CARNOW	Child	Focal child received nonparental care during the summer	P3RELCAR (CCQ020), P3NRLCAR (CCQ120), P3DAYCAR (CCQ220)	1 = Yes, 2 = No
P3HRSNOW	Child	Total number of hours per week the focal child spent in all primary, nonparental care during the summer	P3RELCAR (CCQ020), P3RELREG (CCQ055), P3RELNMH (CCQ070), P3RELNUM (CCQ030), P3RELOTH (CCQ110), P3NRLCAR (CCQ120), P3NRLREG (CCQ150), P3NRLNMH (CCQ170), P3HRLNUM (CCQ130), P3NRMOTH (CCQ210), P3DAYCAR (CCQ220), P3DCREG (CCQ230), P3DCHRS (CCQ250), P3DCNUMB (CCQ225), and P3DCOTHR (CCQ290).	Continuous
P3NUMNOW	Child	Total number of all types of care arrangements the focal child had on a regular basis during the summer	P3RELNUM (CCQ030), P3RELCAR (CCQ020), P3NRLNUM (CCQ130), P3NRLCAR (CCQ120), P3DCNUMB (CCQ225), P3DAYCAR (CCQ220)	Continuous
P3PRIMNW	Child	Primary, nonparental arrangement in which the child spent the most hours per week during the summer	P3RELNMH (CCQ070),P3NRLNMH (CCQ170), P3DCHRS (CCQ250),P3RELPLA (CCQ050),P3NRLPLA (CCQ140)	0 = No nonparental care, 1 = Relative care in child's home, 2 = Relative care in another home, 3 = Nonrelative care in child's home, 4 = Nonrelative care in another home, 5 = Center-based program, 6 = 2 or more programs, 7 = location of care varies

Table 7-13.—Fall-first grade composite variables (continued)

Variable Name	Category	Description	Derived from	Values
P3SUMVD	Child	Length of summer school vacation in days	P3LTOTMO (TUQ010), P3LTOTDA (TUQ015), P3STRTMO (TUQ020), P3STRTDA (TUQ025)	1 = 15 days or less, 2 = 16-109 days, and 3 = 110 or more days.
P3SUMSH	Child	Length of summer school program in hours	P3SUMSCH (HEQ220), P3PRMSTM (HEQ230A), P3PRMSTD (HEQ230B), P3PRMEDM (HEQ240A), P3PRMEDD (HEQ240B), P3NDYPRM (HEQ250), P3NHRPRM (HEQ260)	1 = 30 hours or less, 2 = 31-359 hours, and 3 = 360 hours or more.
R3R2SCHG	Child	Child made a school change from public to private or vice versa between spring-kindergarten and fall-first grade.	S2KPUPRI, S3PUPRI	1 = Child did not change schools 2 = Child transferred from public to public 3 = Child transferred from private to private 4 = Child transferred from public to private 5 = Child transferred from private to public 6 = Child transferred, other
R3DEST	Child	Moved to fall-first grade destination school	FMS (variable not on file)	1=Yes 2=No
P3KSHCHG	Family/ HH	Household had roster change. A dichotomous variable set to 1 if a household experienced a true change between rounds. In this context, true change means that at least one individual entered or left the household between rounds, and this change did not involve interviewer corrections for fall-first grade roster errors.	P3REL (FSQ130), P3UNR (FSQ180), P3JOI, P3REASL (FSQ015)	0 = False, 1 = True

Table 7-13.—Fall-first grade composite variables (continued)

Variable Name	Category	Description	Derived from	Values
R3REGION	School	Indicates the geographic region of the child's school	Sampling Frame (variable not on file)	1=Northeast: CT, ME, MA, NH, RI, VT, NJ, NY, PA; 2=Midwest: IL, IN, MI, OH, WI, IA, KS, MN, MO, NE, ND, SD; 3=South: DE, DC, FL, GA, MD, NC, SC, VA, WV, AL, KY, MS, TN, AR, LA, OK, TX; 4=West: AZ, CO, ID, MT, NV, NM, UT, WY, AK, CA, HA, OR, WA; -1 = schooled at home
R3URBAN	School	Location type for school—7 category version	Sample frame (variable not on file)	1=Large city - a central city of Consolidated Metropolitan Statistical Area (CMSA) with a pop. Greater to or equal to 250,000; 2=Mid-size city - a central city of a CMSA or Metropolitan Statistical Area (MSA) with a pop. Less than 250,000; 3= Urban fringe of large city - any incorporated place, Census Designated Place, or nonplace territory within a CMSA or MSA of a large city and defined as urban by the Census Bureau; 4 = Urban Fringe of mid-size city - any incorporated place, Census Designated Place, or nonplace territory within a CMSA or MSA of a mid-size city and defined as urban by the Census Bureau; 5= Large town - an incorporated place or Census Designated Place with a pop. Greater than or equal to 25,000 and located outside a CMSA or MSA; 6=Small town - an incorporated place or Census Designated Place with a pop. Less than 25,000 and greater than 2,500 - located outside a CMSA or MSA; 7=Rural - any incorporated place, Census Designated Place, or nonplace territory designated as rural by the Census Bureau

Table 7-13.—Fall-first grade composite variables (continued)

Variable Name	Category	Description	Derived from	Values
R3LOCALE	School	Location type for school—8 category version	Sample frame (variable not on file)	1=Large city - a central city of Consolidated Metropolitan Statistical Area (CMSA) with a pop. Greater to or equal to 250,000; 2=Mid-size city - a central city of a CMSA or Metropolitan Statistical Area (MSA) with a pop. Less than 250,000; 3= Urban fringe of large city - any incorporated place, Census Designated Place, or nonplace territory within a CMSA or MSA of a large city and defined as urban by the Census Bureau; 4 = Urban Fringe of mid-size city - any incorporated place, Census Designated Place, or nonplace territory within a CMSA or MSA of a mid-size city and defined as urban by the Census Bureau; 5= Large town - an incorporated place or Census Designated Place with a pop. Greater than or equal to 25,000 and located outside a CMSA or MSA; 6=Small town - an incorporated place or Census Designated Place with a pop. Less than 25,000 and greater than 2,500 - located outside a CMSA or MSA; 7 = non-MSA Rural - any incorporated place, Census Designated Place, or nonplace territory designated as rural by the Census Bureau that is not within a MSA; 8 = MSA Rural - any incorporated place, Census Designated Place, or nonplace territory designated as rural by the Census Bureau that is within a MSA
S3PUPRI	School	Public or private school	S3SCTYP	1 = Public, 2 = Private

Table 7-14.—Spring-first grade composite variables

Variable name	Category	Description	Derived from	Values
CHILDID	Child	Child identification number	OCHILDID (not on file)	
R4_AGE	Child	Child's age at the time the direct child assessment occurred. This age will be in month and years, not days.	DOBMO, DOBDA, DOBYR, C4ASMTMM, C4ASMTDD, C4ASMTYY	Continuous
GENDER	Child	Child's gender	CHILDGEN (not on file) and FMS (variable not on file)	1=Male; 2=Female
DOBMO	Child	Child date of birth month	CHILDDOB (not on file) from first data collection in which reported in parent interview, FMS date of birth variable (not on file)	1-12
DOBDD	Child	Child's date of birth day	CHILDDOB (not on file) from first data collection in which reported in parent interview, and FMS date of birth variable (not on file)	1-31
DOBYY	Child	Child's date of birth year	CHILDDOB (not on file) from first data collection in which reported in parent interview, and FMS date of birth variable (not on file)	
W1RACETH	Child	Race and ethnicity of the focal child	P4RC1_1 through P4RC6_1 up to "_25" (FSQ195)	1=White, 2=Black or African American, 3=Hispanic, race specified, 4=Hispanic, no race specified, 5=Asian, 6=Native Hawaiian or other Pacific Islander, 7=American Indian or Alaska Native, 8=More than 1 race, non-Hispanic

Table 7-14.—Spring-first grade composite variables (continued)

Variable name	Category	Description	Derived from	Values
RACE	Child	Child race and ethnicity	WKRACETH, W1RACETH, FMS race variable (not on file)	1=White, 2=Black or African American, 3=Hispanic, race specified, 4=Hispanic, no race specified, 5=Asian, 6=Native Hawaiian or other Pacific Islander, 7=American Indian or Alaska Native, 8=More than 1 race, non-Hispanic
W1AMERIN	Child	Child is Native American	P4RC1_1 through P4RC6_1 up to "_25" (FSQ195)	1=Yes, 2=No
W1ASIAN	Child	Child is Asian	P4RC1_1 through P4RC6_1 up to "_25" (FSQ195)	1=Yes, 2=No
W1PACISL	Child	Child is Pacific Islander	P4RC1_1 through P4RC6_1 up to "_25" (FSQ195)	1=Yes, 2=No
W1BLACK	Child	Child is African American	P4RC1_1 through P4RC6_1 up to "_25" (FSQ195)	1=Yes, 2=No
W1WHITE	Child	Child is White	P4RC1_1 through P4RC6_1 up to "_25" (FSQ195)	1=Yes, 2=No
W1HISP	Child	Child is Hispanic	P4HSP_1 through P4HSP_25 (FSQ190)	1=Yes, 2=No
W1MT1RAC	Child	Child is more than one race	FSQ195 and the other specify text (not on file) collected in FSQ198 when FSQ195 = 91 (another race)	1=Yes, 2=No
C4BMI	Child	Spring-first grade body mass index	C4HEIGHT, C4WEIGHT	Continuous
C4HEIGHT	Child	Child's spring-first grade composite height	C4HGT1, C4HGT2	Continuous
C4WEIGHT	Child	Child's spring-first grade composite weight	C4WGT1, C4WGT2	Continuous

Table 7-14.—Spring-first grade composite variables (continued)

Variable name	Category	Description	Derived from	Values
P4DISABL	Child	Child currently has a disability	P4DIAGNO (CHQ050), P4PROFFD (CHQ110), P4COMMU2 (CHQ170), P4DIFFH3 (CHQ210), P4VISIO2 (CHQ300), P4RSVTSY (CHQ340)	1=Yes, 2=No
P4CARNOW	Child	Focal child is currently receiving any nonparental care	P4RELNOW (CCQ010), P4NRNOW (CCQ150), P4CTRNOW (CCQ260)	1=Yes, 2=No
WKCAREPK	Child	Focal child received any nonparental care the year before entering kindergarten	P4RCBEFK (SPQ.010), P4NCBEFK (SPQ.020), P4HSBEFK (SPQ.030), P4DBEFK (SPQ.040)	1=Yes, 2=No
P4HRSNOW	Child	Total number of hours per week the focal child currently spends in all primary, nonparental child care	P4RHRS (CCQ090), P4NHRS (CCQ190), P4CHRS (CCQ355), P4RELNOW (CCQ010), P4RELNUM (CCQ060), P4RHROTH (CCQ140), P4NRNOW (CCQ150), P4NRNUM (CCQ165), P4NHROTH (CCQ250), P4CTRNOW (CCQ260), P4CTRNUM (CCQ325), P4CHROTH (CCQ403), P4RWEEK (CCQ080), P4NWEEK (CCQ180), P4CWEEK (CCQ340)	Continuous
P4NUMNOW	Child	Total number of all types of nonparental care arrangements the focal child currently has on a regular basis	P4RELNUM (CCQ060), P4NRNUM (CCQ165), P4CTRNUM (CCQ325), P4RELNOW (CCQ010), P4NRNOW (CCQ150), P4CTRNOW (CCQ260)	Continuous
P4PRIMNW	Child	Primary nonparental child care arrangement in which the child currently spends the most hours per week	P4HRSNOW (composite), P4RHRS (CCQ090), P4NHRS (CCQ190), P4RPLACE (CCQ070), P4NPLACE (CCQ170), P4CHRS (CCQ355)	0=No nonparental care, 1=Relative care in child's home, 2=Relative care in another home 3=Nonrelative care in child's home, 4=Nonrelative care in another home, 5=Center-based program, 6=2 or more programs, 7 = Location of care varies

Table 7-14.—Spring-first grade composite variables (continued)

Variable name	Category	Description	Derived from	Values
CPSOLDS	Child	Round in which the child passed the English Oral Language Development Scale (OLDS)	C1SCTOT, C2SCTOT, C3SCTOT, C4SCTOT	0 = Child not identified as needing the OLDS 1 = Child passed OLDS in fall-kindergarten 2 = Child passed OLDS in spring-kindergarten 3 = Child passed OLDS in fall-first grade 4 = Child passed OLDS in spring-first grade 9 = Child did not pass the OLDS by spring-first grade
F4SPECS	Child	This variable indicates whether or not the child received special education services based on the presence or absence of a link to a special education teacher in the FMS	Child-teacher link in the FMS (variable not on file)	1 = Child got special education services in spring-first grade 2 = Child did not get special education services in spring-first grade
U4RIEP	Child	This variable indicates whether or not the child has an Individualized Education Program (IEP) or Individualized Family Service Plan (IFSP) on record at his/her school or at another school according to information from the student record abstract.	U4IEP	1 = Child has IEP/IFSP on record at his/her school or another school 2 = Child does not have an IEP/IFSP
U4TARD	Child	Total days tardy during school year	U4TRDYT1, U4TRDYE1, U4TRDYU1, U4TRDYT2, U4TRDYE2, U4TRDYU2, U4TRDYT3, U4TRDYE3, U4TRDYU3, U4TRDYT4, U4TRDYE4, U4TRDYU4	Continuous
U4ABSN	Child	Total number of days absent during the school year	U4ABSTO1, U4ABSEX1, U4ABSUN1, U4ABSTO2, U4ABSEX2, U4ABSUN2, U4ABSTO3, U4ABSEX3, U4ABSUN3, U4ABSTO4, U4ABSEX4, U4ABSUN4	Continuous

Table 7-14.—Spring-first grade composite variables (continued)

Variable name	Category	Description	Derived from	Values
A4PLBK	Child	Percent of blacks in class—child level data	If child is in first grade: A4FPBLK (variable not on file); if child is in kindergarten: either A4APBLK, A4PPBLK, or A4DPBLK (variables not on file)	0-100
A4PHIS	Child	Percent of Hispanics in class—child level data	If child is in first grade: A4FPHIS (variable not on file); if child is in kindergarten: either A4APHIS, A4PPHIS, or A4DPHIS (variables not on file)	0-100
A4PMIN	Child	Percent of minorities in class—child level data	If child is in first grade: A4FPMIN (variable not on file); if child is in kindergarten: either A4APMIN, A4PPMIN, or A4DPMIN (variables not on file)	0-100
R4DEST	Child	Moved to spring-first grade destination school	FMS (variable not on file)	1=Yes 2=No
F2SPECS	Child	Whether or not the child received special education services based on the presence or absence of a link to a special education teacher in the FMS	Child-teacher link in the FMS (variable not on file)	1 = Child got special education services in spring-kindergarten 2 = Child did not get special education in spring-kindergarten
U2RIEP	Child	Whether or not the child has an IEP or IFSP on record at his/her school according to information from the Student Record Abstract	U2IEP	1 = Child has IEP/IFSP on record at his/her school 2 = Child does not have an IEP/IFSP on record
U2TARD	Child	Total days tardy during the school year	U2TRDYT1, U2TRDYE1, U2TRDYU1, U2TRDYT2, U2TRDYE2, U2TRDYU2, U2TRDYT3, U2TRDYE3, U2TRDYU3, U2TRDYT4, U2TRDYE4, U2TRDYU4	Continuous

Table 7-14.—Spring-first grade composite variables (continued)

Variable name	Category	Description	Derived from	Values
U2ABSN	Child	Total number of days absent during the school year	U2ABSTO1, U2ABSEX1, U2ABSUN1, U2ABSTO2, U2ABSEX2, U2ABSUN2, U2ABSTO3, U2ABSEX3, U2ABSUN3, U2ABSTO4, U2ABSEX4, U2ABSUN4	Continuous
A4CLASS	Child flag	This flag indicates the TQA data that is attached to a child's record and whether the data is for a morning, afternoon, or full-day kindergarten class, or a non-kindergarten class (e.g., first or second grade).	A4ACLASS, A4DCLASS, A4PCLASS, and the presence of TQA data for non-kindergartners	1= spring 2000 kindrgrtn qnr-AM class 2= spring 2000 kindrgrtn qnr-PM class 3= spring 2000 kindrgrtn qnr-AD class 4= spring 2000 qstnr (non-kindergarten)
R4R2TCHG	Child flag	Child changed teachers between spring-kindergarten and spring-first grade	R4T_ID, R2T_ID, R4R2SCHG	0=No change, 1=Changed teachers
R4R2SCHG	Child flag	School type change between spring-kindergarten and spring-first grade	school ID	1 = child did not change schools 2 = child transferred from public to public 3 = child transferred from private to private 4 = child transferred from public to private 5 = child transferred from private to public 6 = child transferred, other
R4R3SCHG	Child flag	School type change between fall-first grade and spring-first grade	school ID, S3SCTYP, S4SCTYP	1 = child did not change schools 2 = child transferred from public to public 3 = child transferred from private to private 4 = child transferred from public to private 5 = child transferred from private to public 6 = child transferred, other
RNDFLG	Child flag	Round that child joined study	C1CPTW0, C1PW0, C1CW0, C2CPTW0, C2CW0, C2PW0	1 = child joined in fall-kindergarten 2 = child joined in spring-kindergarten 3 = child joined in spring-first grade
P4MOMID	Family/HH	Household roster number of resident mother	P4REL_1 to P4REL_25 (FSQ130), P4UNR_1 to P4UNR_25 (FSQ180)	1-17

Table 7-14.—Spring-first grade composite variables (continued)

Variable name	Category	Description	Derived from	Values
P4DADID	Family/HH	Household roster number of resident father	P4REL_1 to P4REL_25 (FSQ130), P4UNR_1 to P4UNR_25 (FSQ180)	1-17
P4HPARNT	Family/HH	Classification of the focal child's parents who reside in the household	P4REL (FSQ130), P4UNR (FSQ180), P4HMOM, P4HDAD (composites)	1=Biological mother and biological father, 2=Biological mother and other father (step-, adoptive, foster), 3=Biological father and other mother (step-, adoptive, foster), 4=Biological mother only, 5=Biological father only, 6=Two adoptive parents, 7=Single adoptive parent or adoptive parents and stepparent, 8=Related guardian(s), 9=Unrelated guardian(s)
P4HFAMIL	Family/HH	Family type categories using both parent and sibling information	P4REL (FSQ130), P4UNR (FSQ180), P4HMOM, P4HDAD, P4NUMSIB (composites)	1=Two parents and sibling(s), 2=Two parents, no siblings, 3=One parent and sibling(s), 4=One parent, no siblings, 5=Other
P4NUMSIB	Family/HH	Total number of siblings with whom the focal child lives, including anyone reporting him/herself as the child of the focal child's foster parent/guardian	P4REL_1 to P4REL_25 (FSQ130), P4UNR_1 to P4UNR_25 (FSQ180)	Continuous
P4LESS18	Family/HH	Total number of household members younger than 18 years old	P4REL_1 to P4REL_25 (FSQ.130), P4UNR_1 to P4UNR_25 (FSQ.180), P4AGE_1 to P4AGE_25 (FSQ.030)	Continuous
P4OVER18	Family/HH	Total number of household members age 18 or older	P4REL_1 to P4REL_25 (FSQ.130), P4UNR_1 to P4UNR_25 (FSQ.180), P4AGE_1 to P4AGE_25 (FSQ.030)	Continuous
P4HTOTAL	Family/HH	Total number of household members	P4REL_1 to P4REL_25 (FSQ130), P4UNR_1 to P4UNR_25 (FSQ180)	Continuous

Table 7-14.—Spring-first grade composite variables (continued)

Variable name	Category	Description	Derived from	Values
P4TWIN	Family/HH	Household has sampled twins	P4PER_1 to P4PER_25 (person type in FSQ roster)	0=No twin in HH, 1=Twin in HH
W1POVRTY	Family/HH	Poverty indicator	W1INCOME, P4HTOTAL (composites), and Census-defined thresholds	1=Below poverty threshold, 2=At or above poverty threshold
W1INCOME	Family/HH	Household income	P4HILOW (PAQ100), P4INCCAT (PAQ110), P4INCOME (PAQ120)	Continuous
W1SESL	Family/HH	Socioeconomic scale	W1INCOME, W1MOMED, W1DADED, W1MOMSCR, W1DADSCR (all composites)	Continuous
W1SESQ5	Family/HH	Quintile indicator for W1SESL	Derived from W1SESL at the child level using spring-first grade P4 weights	1=First quintile (lowest), 2=Second quintile, 3=Third quintile, 4=Fourth quintile, 5=Fifth quintile (highest)
W1PARED	Family/HH	Highest level of education for the child's parents or nonparental guardians who reside in the household. If only one parent or guardian resides in the household, W1PARED reflects that parent's education level.	W1MOMED, W1DADED (composites)	1=8th grade or below, 2=9th to 12th grades, 3=High school diploma/equivalent, 4=Voc/Tech program, 5=Some college, 6=Bachelor's Degree, 7=Graduate/professional school/no degree, 8=Master's degree, 9=Doctorate or professional degree

Table 7-14.—Spring-first grade composite variables (continued)

Variable name	Category	Description	Derived from	Values
W1MOMSCR	Family/HH	Mother's occupation GSS prestige score	1989 GSS prestige scores, EMQ120 (not on file), EMQ130 (not on file) , and EMQ140 (not on file).	29.6 Handler, Equip, Cleaner, Helpers, Labor; 33.42 Production Working Occupation; 34.95 Service Occupations; 35.63 Agriculture, Forestry, Fishing Occupations; 35.78 Marketing & Sales Occupation; 35.92 Transportation, Material Moving; 37.67 Precision Production Occupation; 38.18 Administrative Support, Including Clerk; 39.18 Mechanics & Repairs; 39.2 Construction & Extractive Occupations; 48.69 Technologists, Except Health; 52.54 Writers, Artists, Entertainers, Athletes; 53.5 Executive, Admin, Managerial Occupation; 57.83 Health Technologists & Technicians; 59 Social Scientist/Workers, Lawyers; 61.56 Registered Nurses, Pharmacists; 62.87 Natural Scientists & Mathematicians; 63.43. Teacher, Except Postsecondary; 64.89 Engineers, Surveyors, & Architects; 72.1 Teachers; College, Postsecondary Counselors, Librarians; 77.5 Physicians, Dentists, Veterinarians

Table 7-14.—Spring-first grade composite variables (continued)

Variable name	Category	Description	Derived from	Values
W1DADSCR	Family/HH	Father's occupation GSS prestige score	1989 GSS prestige scores, EMQ120 (not on file)EMQ130 (not on file), and EMQ140 (not on file).	29.6 Handler, Equip, Cleaner, Helpers, Labor; 33.42 Production Working Occupation; 34.95 Service Occupations; 35.63 Agriculture, Forestry, Fishing Occupations; 35.78 Marketing & Sales Occupation; 35.92 Transportation, Material Moving; 37.67 Precision Production Occupation; 38.18 Administrative Support, Including Clerk; 39.18 Mechanics & Repairs; 39.2 Construction & Extractive Occupations; 48.69 Technologists, Except Health; 52.54 Writers, Artists, Entertainers, Athletes; 53.5 Executive, Admin, Managerial Occupation; 57.83 Health Technologists & Technicians; 59 Social Scientist/Workers, Lawyers; 61.56 Registered Nurses, Pharmacists; 62.87 Natural Scientists & Mathematicians; 63.43. Teacher, Except Postsecondary; 64.89 Engineers, Surveyors, & Architects; 72.1 Teachers; College, Postsecondary Counselors, Librarians; 77.5 Physicians, Dentists, Veterinarians
W1LANGST	Family/HH	The primary language other than English spoken in the child's home	P4ANYLNG (SPQ155), P4PRMLNG (SPQ157)	1=Non-English language, 2=English language
P4MARSTA	Family/HH	Marital status of parent(s) in household	P4CURMAR (FSQ300)	1=Married, 2=Separated, 3=Divorced, 4=Widowed, 5=Never married, 7=No biological or adoptive parents in household

Table 7-14.—Spring-first grade composite variables (continued)

Variable name	Category	Description	Derived from	Values
P4ERRFLG	Family/HH flag	Household roster has clear errors	P4REL_1 to P4REL_25 (FSQ130), P4UNR_1 to P4UNR_25 (FSQ180), P4JOI_1 to P4JOI1 (round joined study), P4RDP_1 to P4RDP_25 (round departed study), P4REAS_1 to P4REAS_25 (reason left household)	0=False, 1=True
P4SHCHG	Family/HH	Household roster had a change between rounds.	P4REL_1 to P4REL_25 (FSQ130), P4UNR_1 to P4UNR_25 (FSQ180), P4JOI_1 to P4JOI1 (round joined study), P4RDP_1 to P4RDP_25 (round departed study), P4REAS_1 to P4REAS_25 (reason left household)	0=False, 1=True
P4SPQDAT	Family/HH	Record in section SPQ	PARENTID	
P4DADTYP	Family/HH	Type of resident father	P4REL(FSQ130), P4UNR (FSQ180)	1=Biological, 2=Other, 3=None
P4HDAD	Family/HH	Indicates whether the birth, adoptive, step or foster father of the focal child resides in the household with the focal child	P4REL(FSQ130), P4UNR (FSQ180), P4PARTNR (FSQ110), P4SPOUSE (FSQ120)	1=Birth father, 2=Adoptive father, 3=Step father, 4=Foster father, 5=Partner father, 6=Father, unknown type, 7= No resident father
P4HDAGE	Family/HH	Age of resident father	P4AGE_1 through P4AGE_25, P4DADID	Continuous
P4HDRACE	Family/HH	Race and ethnicity of the father or male guardian in the household	P4RC1_1 through P4RC6_1 up to " _25," and P4HSP_1 through P4HSP_25 (FSQ190 and FSQ195).	1=White, 2=Black or African American, 3=Hispanic, race specified, 4=Hispanic, no race specified, 5=Asian, 6=Native Hawaiian or other Pacific Islander, 7=American Indian or Alaska Native, 8=More than 1 race, non-Hispanic
P4HDASN	Family/HH	Father or male guardian in the household is Asian	P4RC1_1 through P4RC6_1 up to "_25" (FSQ195)	1=Yes, 2=No

Table 7-14.—Spring-first grade composite variables (continued)

Variable name	Category	Description	Derived from	Values
P4HDBLCK	Family/HH	Father or male guardian in the household is African American	P4RC1_1 through P4RC6_1 up to "_25" (FSQ195)	1=Yes, 2=No
P4HDIND	Family/HH	Father or male guardian in the household is Native American	P4RC1_1 through P4RC6_1 up to "_25" (FSQ195)	1=Yes, 2=No
P4HDPACI	Family/HH	Father or male guardian in the household is Pacific Islander	P4RC1_1 through P4RC6_1 up to "_25" (FSQ195)	1=Yes, 2=No
P4HDWHT	Family/HH	Father or male guardian in the household is White	P4RC1_1 through P4RC6_1 up to "_25" (FSQ195)	1=Yes, 2=No
P4HDHISP	Family/HH	Father or male guardian in the household is Hispanic	P4HSP_1 through P4HSP_25 (FSQ190)	1=Yes, 2=No
P4HDMT1R	Family/HH	Father or male guardian in the household is more than one race	FSQ195 and the other specify text (not on file) collected in FSQ198 when FSQ195 = 91 (another race).	1=Yes, 2=No
W1DADED	Family/HH	Father's highest level of education	P4HIG_1 through P4HIG_4 (PEQ020), P4HIS_1 through P4HIS_4 (PEQ030)	1=8th grade or below, 2=9th to 12th grades, 3=High school diploma/equivalent, 4=Voc/Tech program, 5=Some college, 6=Bachelor's Degree, 7=Graduate/professional school/no degree, 8=Master's degree, 9=Doctorate or professional degree
P4HDEMP	Family/HH	The work status of the father/male guardian in the household.	P4HRS (EMQ050), P4PAY_1 (EMQ020), P4VAC (EMQ 030), P4LOK (EMQ060), P4DO1 (EMQ070), P4DO2 (EMQ070), P4DO3 (EMQ070), P4DO4 (EMQ070), P4DO5 (EMQ070), P4DO6 (EMQ070), P4D07 (EMQ070)	1=35 hours or more per week (EMQ.050), 2=Less than 35 hours per week (EMQ.050), 3=Looking for work (EMQ.060 & EMQ.070 does not =6), 4=Not in the labor force

Table 7-14.—Spring-first grade composite variables (continued)

Variable name	Category	Description	Derived from	Values
P4DADOCC	Family/HH	Father's occupation	Combination of EMQ120 (not on file), EMQ130 (not on file) , and EMQ140 (not on file)	1 Handler, Equip, Cleaner, Helpers, Labor; 2 Production Working Occupation; 3 Service Occupations; 4 Agriculture, Forestry, Fishing Occupations; 5 Marketing & Sales Occupation; 6 Transportation, Material Moving; 7 Precision Production Occupation; 8 Administrative Support, Including Clerk; 9 Mechanics & Repairs; 10 Construction & Extractive Occupations; 11 Technologists, Except Health; 12 Writers, Artists, Entertainers, Athletes; 13 Executive, Admin, Managerial Occupation; 14 Health Technologists & Technicians; 15 Social Scientist/Workers, Lawyers; 16 Registered Nurses, Pharmacists; 17 Natural Scientists & Mathematicians; 18 Teacher, Except Postsecondary; 19 Engineers, Surveyors, & Architects; 20 Teachers; college, Postsecondary Counselors, Librarians; 21 Physicians, Dentists, Veterinarians; 22 Unemployed, retired
P4MOMTYP	Family/HH	Type of resident mother	P4REL (FSQ130), P4UNR (FSQ180)	1=Biological, 2=Other, 3=None
P4HMOM	Family/HH	Indicates whether the birth, adoptive, step, or foster mother of the focal child resides in the household with the focal child	P4REL (FSQ130), P4UNR (FSQ180)	1=Birth mother; 2=Adoptive mother, 3=Step mother, 4=Foster mother, 5=Parter mother, 6=Mother, unknown type, 7=No resident mother
P4HMAGE	Family/HH	Age of resident mother	P4AGE_1 through P4AGE_25	Continuous

Table 7-14.—Spring-first grade composite variables (continued)

Variable name	Category	Description	Derived from	Values
P4HMRACE	Family/HH	Race and ethnicity of the mother or female guardian in the household	P4RC1_1 through P4RC6_1 up to " _25," and P4HSP_1 through P4HSP_25 (FSQ190 and FSQ195).	1=White, 2=Black or African American, 3=Hispanic, race specified, 4=Hispanic, no race specified, 5=Asian, 6=Native Hawaiian or other Pacific Islander, 7=American Indian or Alaska Native, 8=More than 1 race, non-Hispanic
P4HMASN	Family/HH	Mother or female guardian in the household is Asian	P4RC1_1 through P4RC6_1 up to "_25" (FSQ195)	1=Yes, 2=No
P4HMBLCK	Family/HH	Mother or female guardian in the household is African American	P4RC1_1 through P4RC6_1 up to "_25" (FSQ195)	1=Yes, 2=No
P4HMIND	Family/HH	Mother or female guardian in the household is Native American	P4RC1_1 through P4RC6_1 up to "_25" (FSQ195)	1=Yes, 2=No
P4HMPACI	Family/HH	Mother or female guardian in the household is Pacific Islander	P4RC1_1 through P4RC6_1 up to "_25" (FSQ195)	1=Yes, 2=No
P4HMWHT	Family/HH	Mother or female guardian in the household is White	P4RC1_1 through P4RC6_1 up to "_25" (FSQ195)	1=Yes, 2=No
P4HMHISP	Family/HH	Mother or female guardian in the household is Hispanic	P4HSP_1 through P4HSP_25 (FSQ190)	1=Yes, 2=No
P4HMMT1R	Family/HH	Mother or female guardian in the household is more than one race	FSQ195 and the other specify text (not on file) collected in FSQ198 when FSQ195 = 91 (another race)	1=Yes, 2=No

Table 7-14.—Spring-first grade composite variables (continued)

Variable name	Category	Description	Derived from	Values
W1MOMED	Family/HH	Mother's highest level of education	P4HIG_1 through P4HIG_4 (PEQ020), P4HIS_1 through P4HIS_4 (PEQ030)	1=8th grade or below, 2=9th to 12th grades, 3=High school diploma/equivalent, 4=Voc/Tech program, 5=Some college, 6=Bachelor's Degree, 7=Graduate/professional school/no degree, 8=Master's degree, 9=Doctorate or professional degree
P4HMEMP	Family/HH	The work status of the mother or female guardian in the household	P4PAY_1 (EMQ020), P4HRS (EMQ050), P4VAC (EMQ 030), P4LOK (EMQ060), P4DO1 (EMQ070), P4DO2 (EMQ070), P4DO3 (EMQ070), P4DO4 (EMQ070), P4DO5 (EMQ070), P4DO6 (EMQ070), P4D07 (EMQ070)	1=35 hours or more per week (EMQ.050), 2=Less than 35 hours per week (EMQ.050), 3=Looking for work (EMQ.060 & EMQ.070 does not =6), 4=Not in the labor force

Table 7-14.—Spring-first grade composite variables (continued)

Variable name	Category	Description	Derived from	Values
P4MOMOCC	Family/HH	Mother's occupation	Combination of EMQ120 (not on file), EMQ130 (not on file) , and EMQ140 (not on file)	1 Handler, Equip, Cleaner, Helpers, Labor; 2 Production Working Occupation; 3 Service Occupations; 4 Agriculture, Forestry, Fishing Occupations; 5 Marketing & Sales Occupation; 6 Transportation, Material Moving; 7 Precision Production Occupation; 8 Administrative Support, Including Clerk; 9 Mechanics & Repairs; 10 Construction & Extractive Occupations; 11 Technologists, Except Health; 12 Writers, Artists, Entertainers, Athletes; 13 Executive, Admin, Managerial Occupation; 14 Health Technologists & Technicians; 15 Social Scientist/Workers, Lawyers; 16 Registered Nurses, Pharmacists; 17 Natural Scientists & Mathematicians; 18 Teacher, Except Postsecondary; 19 Engineers, Surveyors, & Architects; 20 Teachers; college, Postsecondary Counselors, Librarians; 21 Physicians, Dentists, Veterinarians; 22 Unemployed, retired
W1HEARLY	Family/HH	Mother worked between the child's birth and the start of kindergarten	P4MOMWRK (SPQ200)	1=Yes, 2=No
W1MOMAR	Family/HH	Biological mother was married to the biological father at the time of the child's birth	P4BIOLOG (SPQ150)	1=Yes, 2=No
P4ABSDAD	Family/HH	Type of nonresident father	P4REL (FSQ130),) P4CTP_N1-P4CTP_N4 (NRQ100)	1=Biological only, 2=Both biological and adoptive
P4BDVITL	Family/HH	The vital status of the nonresident biological father	P4CTP_N1-P4CTP_N4 (NRQ100)	1=Living, 2=Dead

Table 7-14.—Spring-first grade composite variables (continued)

Variable name	Category	Description	Derived from	Values
P4ABSMOM	Family/HH	Type of nonresident mother	P4REL (FSQ130), P4CTP_N1-P4CTP_N4 (NRQ100)	1=Biological only, 2=Both biological and adoptive
P4BMVITL	Family/HH	The vital status of the nonresident biological mother	P4CTP_N1-P4CTP_N4 (NRQ100)	1=Living, 2=Dead
P2FSRAW	Family/HH	Food security raw score, a simple count of the number of food security items affirmed by the parent.	P2WORRFD (FDQ130A), P2FDLAST (FDQ130B), P2BLMEAL (FDQ130C), P2LOWCST (FDQ130D), P2NOBAL (FDQ130E), P2CANTAF (FDQ130F), P2EVCUT2 (FDQ140), P2EVCUT (FDQ150), P2EATLES (FDQ160), P2HUNGRY (FDQ170), P2LOSEWT (FDQ180), P2NOTEAT (FDQ190), P2NOTEA2 (FDQ200), P2CUTML (FDQ210), P2CHSKIP (FDQ220), P2OFTCUT (FDQ230), P2CHIEVR (FDQ240), P2NOMONY (FDQ250) Calculations were carried out by USDA in accordance with the standard methods described in Measuring Household Food Security (Bickel et al, 2000, available on the Food and Nutrition Service Web Site, http://www.fns.usda.gov/oane/).	Continuous

Table 7-14.—Spring-first grade composite variables (continued)

Variable name	Category	Description	Derived from	Values
P2FSSCAL	Family/HH	A food security scale score. This is a measure of the severity of food insecurity or hunger experienced in the household in the previous 12 months.	P2WORRFD (FDQ130A), P2FDLAST (FDQ130B), P2BLMEAL (FDQ130C), P2LOWCST (FDQ130D), P2NOBAL (FDQ130E), P2CANTAF (FDQ130F), P2EVCUT2 (FDQ140), P2EVCUT (FDQ150), P2EATLES (FDQ160), P2HUNGRY (FDQ170), P2LOSEWT (FDQ180), P2NOTEAT (FDQ190), P2NOTEA2 (FDQ200), P2CUTML (FDQ210), P2CHSKIP (FDQ220), P2OFTCUT (FDQ230), P2CHIEVR (FDQ240), P2NOMONY (FDQ250) Calculations were carried out in accordance with the standard methods described in Measuring Household Food Security (Bickel et al, 2000, available on the Food and Nutrition Service Web Site, http://www.fns.usda.gov/oane/).	Continuous

Table 7-14.—Spring-first grade composite variables (continued)

Variable name	Category	Description	Derived from	Values
P2FSSTAT	Family/HH	A categorical measure of food security status that identifies households as food secure, food insecure without hunger, food insecure with hunger (moderate), and food insecure with hunger (severe).	P2WORRFD (FDQ130A), P2FDLAST (FDQ130B), P2BLMEAL (FDQ130C), P2LOWCST (FDQ130D), P2NOBAL (FDQ130E), P2CANTAF (FDQ130F), P2EVCUT2 (FDQ140), P2EVCUT (FDQ150), P2EATLES (FDQ160), P2HUNGRY (FDQ170), P2LOSEWT (FDQ180), P2NOTEAT (FDQ190), P2NOTEA2 (FDQ200), P2CUTML (FDQ210), P2CHSKIP (FDQ220), P2OFTCUT (FDQ230), P2CHIEVR (FDQ240), P2NOMONY (FDQ250) Calculations were carried out in accordance with the standard methods described in Measuring Household Food Security (Bickel et al, 2000, available on the Food and Nutrition Service Web Site, http://www.fns.usda.gov/oane/).	1 = Food secure; 2 = Food insecure without hunger; 3 = Food insecure with hunger (moderate); 4 = food insecure with hunger (severe)
B4AGE	Teacher	Teacher's age	B4YRBORN	21-78
T4GLVL	Teacher	Grade level of child	T4GRADE, T4KGRADE, U4CRGRD, T4KTYPCL, FMS teacher type (variable not on file)	1 = Kindergarten, half day 2 = Kindergarten, full day 3 = Kindergarten, half or full day unknown 4 = First Grade 5 = Second Grade 6 = Ungraded classroom
A4ACLASS	Teacher flag	Presence of spring-first grade TQA AM data-Kindergarten	A4ANEW	0=False, 1=True
A4DCLASS	Teacher flag	Presence of spring-first grade TQA all day data-Kindergarten	A4DNEW	0=False, 1=True

Table 7-14.—Spring-first grade composite variables (continued)

Variable name	Category	Description	Derived from	Values
A4PCLASS	Teacher flag	Presence of spring-first grade TQA PM data-Kindergarten	A4PNEW	0=False, 1=True
A4TQUEX	Teacher flag	Presence of spring-first grade TQA data-first or higher grades	Presence or absence of data	0=False, 1=True
B4TQUEX	Teacher flag	Presence of spring-first grade TQB data-Kindergarten	Presence or absence of data	0=False, 1=True
T4TQCDAT	Teacher flag	Presence of spring-first grade TQC data	Presence or absence of data	0=False, 1=True
R4REGION	School	Indicates the geographic region of the child's school	Sampling Frame (variable not on file)	1=Northeast: CT, ME, MA, NH, RI, VT, NJ, NY, PA; 2=Midwest: IL, IN, MI, OH, WI, IA, KS, MN, MO, NE, ND, SD; 3=South: DE, DC, FL, GA, MD, NC, SC, VA, WV, AL, KY, MS, TN, AR, LA, OK, TX; 4=West: AZ, CO, ID, MT, NV, NM, UT, WY, AK, CA, HA, OR, WA;-1 = schooled at home

Table 7-14.—Spring-first grade composite variables (continued)

Variable name	Category	Description	Derived from	Values
R4URBAN	School	Location type for school—7 category version	Sample frame (variable not on file)	1=Large city - a central city of Consolidated Metropolitan Statistical Area (CMSA) with a pop. Greater to or equal to 250,000; 2=Mid-size city - a central city of a CMSA or Metropolitan Statistical Area (MSA) with a pop. Less than 250,000; 3= Urban fringe of large city - any incorporated place, Census Designated Place, or nonplace territory within a CMSA or MSA of a large city and defined as urban by the Census Bureau; 4 = Urban Fringe of mid-size city - any incorporated place, Census Designated Place, or nonplace territory within a CMSA or MSA of a mid-size city and defined as urban by the Census Bureau; 5= Large town - an incorporated place or Census Designated Place with a pop. Greater than or equal to 25,000 and located outside a CMSA or MSA; 6=Small town - an incorporated place or Census Designated Place with a pop. Less than 25,000 and greater than 2,500 - located outside a CMSA or MSA; 7=Rural - any incorporated place, Census Designated Place, or nonplace territory designated as rural by the Census Bureau

Table 7-14.—Spring-first grade composite variables (continued)

Variable name	Category	Description	Derived from	Values
R4LOCALE	School	Location type for school—8 category version	Sample frame (variable not on file)	1=Large city - a central city of Consolidated Metropolitan Statistical Area (CMSA) with a pop. Greater to or equal to 250,000; 2=Mid-size city - a central city of a CMSA or Metropolitan Statistical Area (MSA) with a pop. Less than 250,000; 3= Urban fringe of large city - any incorporated place, Census Designated Place, or nonplace territory within a CMSA or MSA of a large city and defined as urban by the Census Bureau; 4 = Urban Fringe of mid-size city - any incorporated place, Census Designated Place, or nonplace territory within a CMSA or MSA of a mid-size city and defined as urban by the Census Bureau; 5= Large town - an incorporated place or Census Designated Place with a pop. Greater than or equal to 25,000 and located outside a CMSA or MSA; 6=Small town - an incorporated place or Census Designated Place with a pop. Less than 25,000 and greater than 2,500 - located outside a CMSA or MSA; 7 = non-MSA Rural - any incorporated place, Census Designated Place, or nonplace territory designated as rural by the Census Bureau that is not within a MSA; 8 = MSA Rural - any incorporated place, Census Designated Place, or nonplace territory designated as rural by the Census Bureau that is within a MSA

Table 7-14.—Spring-first grade composite variables (continued)

Variable name	Category	Description	Derived from	Values
S4SCTYP	School	School type from the school administrator questionnaire	S4PUBLIC, S4CATHOL, S4OTHEREL, S4OTHPRI, CS_TYPE2, school sample frame (variable not on file), S3SCTYP, S2KSCTYP, transfer school file	1=Catholic, 2=Other Religious, 3=Other Private, 4=Public
S4PUPRI	School	Public or private school	S4SCTYP	1=Public, 2=Private
S4ENRLK	School	Total kindergarten enrollment	Sampling Frame (variable not on file)	Continuous
S4ENRLF	School	Total first grade enrollment	Sampling Frame (variable not on file)	Continuous
S4ENRLS	School	Total school enrollment	S4ANUMCH (SAQ), Sampling Frame (variable not on file)	1=0-149 students; 2=150-299 students; 3=300-499 students; 4=500-749 students; 5= 750 and above students
S4MINOR	School	Percentage of minority students in school	Sampling Frame (variable not on file), S4PCTASN, S4HISPPCT, S4BLKPCT, S4WHTPCT, S4INDPCT, S4OTHPCT,	Continuous
S4GFTED	School	Percentage of gifted/talented students in school	S4GFTNBR, S4ANUMCH, S2KGFTED	Continuous
S4FLNCH	School	Percentage of students eligible for free lunch in school	S4ELILNC, S4ANUMCH	Continuous
S4RLNCH	School	Percent of students eligible for reduced price lunch in school	S4ELIRED, S4ANUMCH	Continuous

Table 7-14.—Spring-first grade composite variables (continued)

Variable name	Category	Description	Derived from	Values
S4SCLVL	School	School instructional level from SAQ	S412TH, S411TH, S4TENTH, S4NINTH, S48TH, S47TH, S4SIXTH, S4FIFTH, S4FOURTH, S4THIRD, S4SECOND, S4GRADE1, S4PRE1, S4KINDER, S4TRANS, S4PRKNDR, S4TWEL12, S4ELEV11, S4TEN10, S4NIN9TH, S4ATE8TH, S4SEV7TH, S4SIX6TH, S4FIF5TH, S4FOR4TH, S4THI3RD, S4SEC2ND, S4FIR2ST, S4PRE1ST, S4KINGAR, S4REDDYN, S4PREK, S2SCLVL, Sampling Frame (variable not on file)	1=Less than 1st grade, 2=Primary school, 3=Elementary school, 4=Combined school
U4SCHBDD	School	School year starting date day	R4RDDSTR1, R4RDDSTR2, R4RDDSTR3, R4RDDSTR4, and whether a year round school (from FMS variable not on file)	1-31
U4SCHBMM	School	School year starting date month	R4MMSTR1, R4MMSTR2, R4MMSTR3, R4MMSTR4, and whether a year round school (from FMS variable not on file)	1-12
U4SCHBYY	School	School year starting date year	R4YYSTR1, R4YYSTR2, R4YYSTR3, R4YYSTR4, and whether a year round school (from FMS variable not on file)	1999-2000
U4SCHEDD	School	School year ending date day	R4RDDSTR1, R4RDDSTR2, R4RDDSTR3, R4RDDSTR4, and whether a year round school (from FMS variable not on file)	1-31
U4SCHEMM	School	School year ending date month	R4MMSTR1, R4MMSTR2, R4MMSTR3, R4MMSTR4, and whether a year round school (from FMS variable not on file)	1-12

Table 7-14.—Spring-first grade composite variables (continued)

Variable name	Category	Description	Derived from	Values
U4SCHEYY	School	School year ending date year	R4YYSTR1, R4YYSTR2, R4YYSTR3, R4YYSTR4, and whether a year round school (from FMS variable not on file)	1999-2000
F4YRRND	School	Year round school	FMS (variable not on file)	1 = Year round school, 2 = Not year round school
K4INFAC	School flag	Presence or absence of facilities checklist data	Presence or absence of data	0=ID not in facilities checklist, 1=ID in facilities checklist
S4INSAC	School flag	Presence or absence of SAQ data	Presence or absence of data	0=ID not in SAQ, 1=ID in SAQ
U4SRABS	School flag	Presence of spring-first grade SRA data	Presence or absence of data	0 = False, 1 = True

7.9 Masked Variables

All the variables from the ECLS-K restricted-use file are included in the same order on the ECLS-K public-use file. New variables created during the masking process are added to the end of the files. For some of the variables, certain categories were modified. The value labels for those masked variables were updated from the restricted-use variables to reflect the new categories that were created during the masking process.

Variables on the restricted-use files were modified in different ways based on the disclosure analysis NCES conducted in order to protect the identity of the respondents and children. There are several types of modifications on the public-use files.

1. Outliers are top- or bottom- coded to prevent identification of unique schools, teachers, parents and children without affecting overall data quality.
2. Certain schools identified as at risk for disclosure have a 5 to 10 percent noise introduced in those variables that pose a risk for disclosure.
3. Variables with too few cases and a sparse distribution are suppressed in the public-use files.
4. Certain continuous variables are modified into categorical variables, and certain categorical variables have their categories collapsed in the public-use file. While this protects the cases from a disclosure risk, these variables can still be used in all different kinds of analysis such as regression analysis.

There is a comment field in the variable frequency distribution view screen of the electronic code book that displays a comment for each masked variable indicating whether the variable from the restricted-use file has been recoded or suppressed in the public-use file. Variables that were recoded in any way during the data masking process display the comment, "This data recoded for respondent confidentiality." Variables that were suppressed on the public-use file for protection of the respondent or child from identification display the comment, "This data suppressed for respondent confidentiality" and all values for the variable are set to equal –2 for that variable.

Table 7-15 presents the list of the masked variables. The tables display the variable name, variable label and the comment displayed in the electronic code book indicating if the variable was recoded or suppressed. The table is sorted sequentially by the variable Field ID (see section 8.3.1.1 for how to use the variable Field ID.)

All variables from the special education teacher questionnaire part A (i.e., all variables with the prefix D4), from the special education teacher questionnaire part B (i.e., all variables with the prefix E4), and from the student record abstract (i.e., all variables with the prefix U4) have been suppressed on the first grade public-use file. For brevity, these variables are not included in table 7-15.

Table 7-15.—Recoded and suppressed data on the ECLS-K First-Grade Year Public-Use File

FieldID	FieldName	FieldLabel	Comment
10	CS_TYPE2	TYPE OF SCHOOL IN BASE YEAR SAMPLE FRAME	This data recoded for respondent confidentiality.
12	R3URBAN	R3 LOCATION TYPE - 7 CATEGORIES	This data recoded for respondent confidentiality.
13	R3LOCALE	R3 LOCATION TYPE - 8 CATEGORIES	This data suppressed for respondent confidentiality.
15	R3FIPSST	R3 SCHOOL FIPS STATE CODE	This data suppressed for respondent confidentiality.
16	R3FIPSCT	R3 SCHOOL FIPS COUNTY CODE	This data suppressed for respondent confidentiality.
17	R3CCDLEA	R3 CCD LEA\SCHOOL DIST ID (PUBLIC)	This data suppressed for respondent confidentiality.
18	R3CCDSID	R3 CCD SCHOOL ID (PUBLIC)	This data suppressed for respondent confidentiality.
19	R3STSID	R3 STATE SCHOOL ID (PUBLIC)	This data suppressed for respondent confidentiality.
20	R3SCHZIP	R3 SCHOOL ZIP CODE	This data suppressed for respondent confidentiality.
21	R3SCHPIN	R3 SCHOOL PIN (PRIVATE)	This data suppressed for respondent confidentiality.
22	R4REGION	R4 CENSUS REGION	This data recoded for respondent confidentiality.
23	R4URBAN	R4 LOCATION TYPE - 7 CATEGORIES	This data recoded for respondent confidentiality.
24	R4LOCALE	R4 LOCATION TYPE - 8 CATEGORIES	This data suppressed for respondent confidentiality.
26	R4FIPSST	R4 SCHOOL FIPS STATE CODE	This data suppressed for respondent confidentiality.
27	R4FIPSCT	R4 SCHOOL FIPS COUNTY CODE	This data suppressed for respondent confidentiality.
28	R4CCDLEA	R4 CCD LEA\SCHOOL DIST ID (PUBLIC)	This data suppressed for respondent confidentiality.
29	R4CCDSID	R4 CCD SCHOOL ID (PUBLIC)	This data suppressed for respondent confidentiality.
30	R4STSID	R4 STATE SCHOOL ID (PUBLIC)	This data suppressed for respondent confidentiality.
31	R4SCHZIP	R4 SCHOOL ZIP CODE	This data suppressed for respondent confidentiality.
32	R4SCHPIN	R4 SCHOOL PIN (PRIVATE)	This data suppressed for respondent confidentiality.
37	DOBYY	CHILD COMPOSITE DOB YEAR	This data recoded for respondent confidentiality.
43	R4AGE	R4 COMPOSITE CHILD ASSESSMENT AGE(MNTHS)	This data recoded for respondent confidentiality.
47	R4R2SCHG	R4 CHILD SCH CHANGE TYPE BTWN RNDS 2 & 4	This data recoded for respondent confidentiality.
48	R4R3SCHG	R4 CHILD SCH CHANGE TYPE BTWN RNDS 3 & 4	This data recoded for respondent confidentiality.
345	A4PBLK	A4 PERCENT OF BLACKS IN CLASS	This data recoded for respondent confidentiality.
346	A4PHIS	A4 PERCENT OF HISPANICS IN CLASS	This data recoded for respondent confidentiality.
347	A4PMIN	A4 PERCENT OF MINORITIES IN CLASS	This data recoded for respondent confidentiality.
348	B4AGE	B4 TEACHER'S AGE	This data recoded for respondent confidentiality.
350	S3SCTYP	S3 SCHOOL TYPE FROM FMS	This data recoded for respondent confidentiality.
352	S4SCTYP	S4 SCHOOL TYPE FROM THE SCH ADMIN QUEST	This data recoded for respondent confidentiality.
353	S4ENRLK	S4 TOTAL SCHOOL K ENROLLMENT	This data recoded for respondent confidentiality.
354	S4ENRLF	S4 TOTAL SCHOOL FIRST GRADE ENROLLMENT	This data recoded for respondent confidentiality.
355	S4ENRLS	S4 TOTAL SCHOOL ENROLLMENT	This data recoded for respondent confidentiality.
356	S4MINOR	S4 PERCENT MINORITY STUDENTS	This data recoded for respondent confidentiality.
357	S4FLNCH	S4 PCT FREE LUNCH ELIGIBLE STUDENTS	This data recoded for respondent confidentiality.
358	S4RLNCH	S4 PCT REDUCED LUNCH ELIGIBLE STUDENTS	This data recoded for respondent confidentiality.
360	S4GFTED	S4 PERCENT GIFTED/TALENTED STUDENTS	This data recoded for respondent confidentiality.
361	S4PUPRI	S4 PUBLIC OR PRIVATE SCHOOL	This data recoded for respondent confidentiality.
398	P3HOMZIP	P3 HOME ZIP CODE	This data suppressed for respondent confidentiality.
734	P3PRMLNG	P3 SPQ157 WHAT PRIMARY LANGUAGE AT HOME	This data suppressed for respondent confidentiality.
902	P4HOMZIP	P4 HOME ZIP CODE	This data suppressed for respondent confidentiality.
1350	P4THER12	P4 SPQ120K INSTRUCTION IN BRAILLE	This data suppressed for respondent confidentiality.
1351	P4THER13	P4 SPQ120L INSTRUCTION IN SIGN LANGUAGE	This data suppressed for respondent confidentiality.
1352	P4THER16	P4 SPQ130A AGE AT THERAPY BEGINNING-YR	This data suppressed for respondent confidentiality.
1353	P4THER17	P4 SPQ130B AGE AT THERAPY BEGINNING-MTH	This data suppressed for respondent confidentiality.
1356	P4PRMLNG	P4 SPQ157 WHAT PRIMARY LANGUAGE AT HOME	This data recoded for respondent confidentiality.
1357	P4CHPLAC	P4 SPQ160 CHILD BORN IN THIS COUNTRY	This data suppressed for respondent confidentiality.
1359	P4YRCOME	P4 SPQ180 YEAR COME TO UNITED STATES	This data suppressed for respondent confidentiality.
1360	P4CITIZN	P4 SPQ190 CHILD A U.S. CITIZEN	This data suppressed for respondent confidentiality.

Table 7-15.—Recoded and suppressed data on the ECLS-K First-Grade Year Public-Use File (continued)

FieldID	FieldName	FieldLabel	Comment
1607	P4DGNATT	P4 CHQ060 DIAGNOSIS - ATTENTION	This data suppressed for respondent confidentiality.
1609	P4YYDIAG	P4 CHQ075 YR AT 1ST DIAGNOSIS-LRN ABLTY	This data suppressed for respondent confidentiality.
1613	P4PROFFD	P4 CHQ110 IF ACTIVITY PROBLEM DIAGNOSED	This data suppressed for respondent confidentiality.
1614	P4DGNACT	P4 CHQ120 DIAGNOSIS - ACTIVITY	This data suppressed for respondent confidentiality.
1616	P4YYDIA2	P4 CHQ135 YR AT 1ST DIAGNOSIS-ACTIVITY	This data suppressed for respondent confidentiality.
1621	P4YYDIA4	P4 CHQ185 YEAR AT 1ST DIAGNOSIS-SPEECH	This data suppressed for respondent confidentiality.
1626	P4YYDIA5	P4 CHQ225 YR AT 1ST DIAGNOSIS-HEARING	This data suppressed for respondent confidentiality.
1627	P4HEARS	P4 CHQ230 DEGREE OF CHILD'S DEAFNESS	This data suppressed for respondent confidentiality.
1628	P4HEARAI	P4 CHQ240 IF CHILD WEARS HEARING AID	This data suppressed for respondent confidentiality.
1629	P4COCHLE	P4 CHQ250 IF CHILD HAS COCHLEAR IMPLANTS	This data suppressed for respondent confidentiality.
1630	P4HEARS2	P4 CHQ260 DEVICE EFFECT ON CHD'S HEARING	This data suppressed for respondent confidentiality.
1635	P4DIA6YY	P4 CHQ313 YR AT 1ST DIAGNOSIS-VISION	This data suppressed for respondent confidentiality.
1636	P4CORREC	P4 CHQ315 IF CHD'S VISION IS CORRECTABLE	This data suppressed for respondent confidentiality.
1637	P4BESTEY	P4 CHQ320 WHAT CAN CHILD BEST SEE	This data suppressed for respondent confidentiality.
1640	P4SPECIL	P4 CHQ335 IF CHD USES SPECIAL EQUIPMENT	This data suppressed for respondent confidentiality.
1714	P4HOWPAY	P4 PAQ137 HOW MUCH PAID IN TUITION ($)	This data recoded for respondent confidentiality.
1731	A4KGRD1T	A4K Q1B FIRST GRADE	This data suppressed for respondent confidentiality.
1732	A4KGRD2T	A4K Q1C SECOND GRADE	This data suppressed for respondent confidentiality.
1733	A4KGRD3T	A4K Q1D THIRD GRADE	This data suppressed for respondent confidentiality.
1734	A4KGRD4T	A4K Q1E FOURTH GRADE	This data suppressed for respondent confidentiality.
1735	A4KGRD5T	A4K Q1F FIFTH GRADE	This data suppressed for respondent confidentiality.
1736	A4KREGK	A4K Q2A TCH REGULAR 1-YR KINDERGARTEN	This data suppressed for respondent confidentiality.
1737	A4K2YR1	A4K Q2B TEACHES 1ST YR OF 2-YR K	This data suppressed for respondent confidentiality.
1738	A4K2YR2	A4K Q2C TEACHES 2ND YR OF 2-YR K	This data suppressed for respondent confidentiality.
1739	A4KTRNK	A4K Q2D TCH TRANSITIONAL KINDERGARTEN	This data suppressed for respondent confidentiality.
1740	A4KPR1S	A4K Q2E TEACHES PRE-1ST GRADE AFTER K	This data suppressed for respondent confidentiality.
1741	A4KUNGR	A4K Q2F TEACHES UNGRADED CLASS	This data suppressed for respondent confidentiality.
1742	A4KMULG	A4K Q2G TEACHES MULTIGRADE CLASS	This data suppressed for respondent confidentiality.
1743	A4KTPRE	A4K Q3A TCH PREKINDERGARTEN LEVELS	This data suppressed for respondent confidentiality.
1744	A4KTTRN	A4K Q3B TCH TRANSITIONAL KINDERGARTEN	This data suppressed for respondent confidentiality.
1745	A4KTREG	A4K Q3C TEACHES REGULAR KINDERGARTEN	This data suppressed for respondent confidentiality.
1746	A4KTPR1	A4K Q3D TCH PRE-1ST GRADE LEVEL	This data suppressed for respondent confidentiality.
1747	A4KT1ST	A4K Q3E TCH 1ST GRADE LEVEL	This data suppressed for respondent confidentiality.
1748	A4KT2ND	A4K Q3F TCH 2ND GRADE LEVEL	This data suppressed for respondent confidentiality.
1749	A4KT3RD	A4K Q3G TCH 3RD GRADE OR HIGHER LEVEL	This data suppressed for respondent confidentiality.
1750	A4KAONLY	A4K Q4A MORNING CLASS ONLY	This data suppressed for respondent confidentiality.
1751	A4KPONLY	A4K Q4B AFTERNOON CLASS ONLY	This data suppressed for respondent confidentiality.
1756	A4K3YRO	A4K Q7A HOW MANY 3-YEAR-OLDS IN CLASS	This data recoded for respondent confidentiality.
1757	A4K4YRO	A4K Q7B HOW MANY 4-YEAR-OLDS IN CLASS	This data recoded for respondent confidentiality.
1758	A4K5YRO	A4K Q7C HOW MANY 5-YEAR-OLDS IN CLASS	This data recoded for respondent confidentiality.
1759	A4K6YRO	A4K Q7D HOW MANY 6-YEAR-OLDS IN CLASS	This data recoded for respondent confidentiality.
1760	A4K7YRO	A4K Q7E HOW MANY 7-YEAR-OLDS IN CLASS	This data recoded for respondent confidentiality.
1761	A4K8YRO	A4K Q7F HOW MANY 8-YEAR-OLDS IN CLASS	This data recoded for respondent confidentiality.
1762	A4K9YRO	A4K Q7G HOW MANY 9-YEAR-OLDS IN CLASS	This data recoded for respondent confidentiality.
1763	A4KTAGE	A4K Q7H TOTAL CLASS ENROLLMENT (AGE)	This data recoded for respondent confidentiality.
1764	A4KASIA	A4K Q8A # OF ASIAN/PACIFIC ISLANDERS	This data recoded for respondent confidentiality.
1765	A4KHISP	A4K Q8B # OF HISPANICS (ALL RACES)	This data recoded for respondent confidentiality.
1766	A4KBLAC	A4K Q8C # OF NON-HISPANIC BLACKS	This data recoded for respondent confidentiality.

Table 7-15.—Recoded and suppressed data on the ECLS-K First-Grade Year Public-Use File (continued)

FieldID	FieldName	FieldLabel	Comment
1767	A4KWHIT	A4K Q8D # OF NON-HISPANIC WHITES	This data recoded for respondent confidentiality.
1768	A4KAMRI	A4K Q8E # OF AMERICAN INDIANS	This data recoded for respondent confidentiality.
1769	A4KRACE	A4K Q8F # OF STUDENTS OF OTHER RACES	This data recoded for respondent confidentiality.
1770	A4KTOTR	A4K Q8 TOTAL CLASS ENROLLMENT (RACES)	This data recoded for respondent confidentiality.
1771	A4KBOYS	A4K Q9 NUMBER OF BOYS IN CLASS	This data recoded for respondent confidentiality.
1772	A4KGIRL	A4K Q9 NUMBER OF GIRLS IN CLASS	This data recoded for respondent confidentiality.
1782	A4KCVTN	A4K Q16B STUDENTS SPEAK VIETNAMESE	This data suppressed for respondent confidentiality.
1783	A4KCCHN	A4K Q16C STUDENTS SPEAK CHINESE	This data suppressed for respondent confidentiality.
1784	A4KCJPN	A4K Q16D STUDENTS SPEAK JAPANESE	This data suppressed for respondent confidentiality.
1785	A4KCKRN	A4K Q16E STUDENTS SPEAK KOREAN	This data suppressed for respondent confidentiality.
1786	A4KCFLP	A4K Q16F STUDENTS SPEAK FILIPINO LNG	This data suppressed for respondent confidentiality.
1787	A4KOTAS	A4K Q16G STUDENT SPEAK OTHR ASIAN LNG	This data suppressed for respondent confidentiality.
1788	A4KOTLN	A4K Q16H STUDENTS SPEAK ANOTHER LNG	This data suppressed for respondent confidentiality.
1789	A4KLANO	A4K Q16H SPECIFY STUDENTS' OTHER LANG	This data suppressed for respondent confidentiality.
1791	A4KNUML	A4K Q18 NUMBER LEP STUDENTS IN CLASS	This data suppressed for respondent confidentiality.
1792	A4KNOES	A4K Q19 LEP STUDENTS GET NO ESL	This data suppressed for respondent confidentiality.
1793	A4KESLR	A4K Q19 LEP STUDENTS GET IN-CLASS ESL	This data suppressed for respondent confidentiality.
1794	A4KESLO	A4K Q19 LEP STUDENTS GET OUTSIDE ESL	This data suppressed for respondent confidentiality.
1795	A4KTNOO	A4K Q20A TCHRS SPEAK ONLY ENGLISH	This data suppressed for respondent confidentiality.
1796	A4KTSPN	A4K Q20B TCHRS SPEAK SPANISH	This data suppressed for respondent confidentiality.
1797	A4KTVTN	A4K Q20C TCHRS SPEAK VIETNAMESE	This data suppressed for respondent confidentiality.
1798	A4KTCHN	A4K Q20D TCHRS SPEAK CHINESE	This data suppressed for respondent confidentiality.
1799	A4KTJPN	A4K Q20E TCHRS SPEAK JAPANESE	This data suppressed for respondent confidentiality.
1800	A4KTKRN	A4K Q20F TCHRS SPEAK KOREAN	This data suppressed for respondent confidentiality.
1801	A4KTFLP	A4K Q20G TCHRS SPEAK A FILIPINO LNG	This data suppressed for respondent confidentiality.
1802	A4KTOTA	A4K Q20H TCHRS SPEAK OTHER ASIAN LNG	This data suppressed for respondent confidentiality.
1803	A4KTOTL	A4K Q20I TCHRS SPEAK ANOTHER LANGUAGE	This data suppressed for respondent confidentiality.
1804	A4KLEPO	A4K Q201SPECIFY TCHRS OTHER LANGUAGE	This data suppressed for respondent confidentiality.
1805	A4KNONE	A4K Q21 DAILY TIME TCHR SPEAK NON-ENG	This data suppressed for respondent confidentiality.
1808	A4KGIFT	A4K Q23A # CLASSIFIED GFTED/TALENTED	This data recoded for respondent confidentiality.
1809	A4KPRTG	A4K Q23B # TAKE PART GIFTED/TALENTED	This data recoded for respondent confidentiality.
1814	A4KDISA	A4K Q24 NUMBER WITH DISABILITIES	This data recoded for respondent confidentiality.
1815	A4KIMPA	A4K Q25A COMMUNICATION IMPAIRMENTS	This data recoded for respondent confidentiality.
1816	A4KLRND	A4K Q25B LEARNING DISABILITIES	This data recoded for respondent confidentiality.
1817	A4KEMPR	A4K Q25C SERIOUS EMOTIONAL PROBLEMS	This data recoded for respondent confidentiality.
1818	A4KRETA	A4K Q25D MENTAL RETARDATION	This data suppressed for respondent confidentiality.
1819	A4KDELA	A4K Q25E DEVELOPMENTAL DELAY	This data recoded for respondent confidentiality.
1820	A4KVIS	A4K Q25F VISION IMPAIRMENTS	This data recoded for respondent confidentiality.
1821	A4KHEAR	A4K Q25G HEARING IMPAIRMENTS	This data recoded for respondent confidentiality.
1822	A4KORTH	A4K Q25H ORTHOPEDIC IMPAIRMENTS	This data recoded for respondent confidentiality.
1823	A4KOTHE	A4K Q25I OTHER HEALTH IMPAIRMENTS	This data recoded for respondent confidentiality.
1824	A4KMULT	A4K Q25J MULTIPLE DISABILITIES	This data suppressed for respondent confidentiality.
1825	A4KAUTS	A4K Q25K AUTISM	This data suppressed for respondent confidentiality.
1826	A4KTRAU	A4K Q25L TRAUMATIC BRAIN INJURIES	This data suppressed for respondent confidentiality.
1827	A4KDEAF	A4K Q25M DEAFNESS AND BLINDNESS	This data suppressed for respondent confidentiality.
1828	A4KOTDI	A4K Q25N OTHER SPECIFY DISABILITIES	This data suppressed for respondent confidentiality.
1829	A4KSPCI	A4K Q26A SPECIAL DISABILITY SERVICES	This data recoded for respondent confidentiality.
1830	A4KIEP	A4K Q26B IEP FOR CHLDRN W/DISABILITY	This data recoded for respondent confidentiality.

Table 7-15.—Recoded and suppressed data on the ECLS-K First-Grade Year Public-Use File (continued)

FieldID	FieldName	FieldLabel	Comment
1831	A4KSC50	A4K Q26C CHILDREN W/SECTION 504 PLAN	This data recoded for respondent confidentiality.
1834	A4KISPN	A4K Q27B INSTRUCTION - SPANISH	This data suppressed for respondent confidentiality.
1837	A4KIJPN	A4K Q27E INSTRUCTION - JAPANESE	This data suppressed for respondent confidentiality.
1842	A4KLNGO	A4K Q27I LANGUAGE OF INSTRUCTION-OTHR	This data suppressed for respondent confidentiality.
2075	A4TMULG	A4 Q1 CURRENTLY TEACH MULTIGRADE CLASS	This data suppressed for respondent confidentiality.
2076	A4TPREK	A4 Q2A TCH PREKINDERGARTEN LEVELS	This data suppressed for respondent confidentiality.
2077	A4TTRNK	A4 Q2B TCH TRANSITIONAL KINDERGARTEN	This data suppressed for respondent confidentiality.
2078	A4TREGK	A4 Q2C TEACHES REGULAR KINDERGARTEN	This data suppressed for respondent confidentiality.
2079	A4TPRE1	A4 Q2D TCH PRE-1ST GRADE LEVEL	This data suppressed for respondent confidentiality.
2080	A4T1ST	A4 Q2E TCH 1ST GRADE LEVEL	This data suppressed for respondent confidentiality.
2081	A4T2ND	A4 Q2F TCH 2ND GRADE LEVEL	This data suppressed for respondent confidentiality.
2082	A4T3RD	A4 Q2G INCLUDES 3RD GRADE	This data suppressed for respondent confidentiality.
2083	A4T4TH	A4 Q2H INCLUDES 4TH GRADE	This data suppressed for respondent confidentiality.
2084	A4T5TH	A4 Q2I INCLUDES 5TH GRADE	This data suppressed for respondent confidentiality.
2085	A45YRSLS	A4 Q3A # 5 YEARS OLD OR LESS IN CLASS	This data recoded for respondent confidentiality.
2086	A46YROL	A4 Q3B HOW MANY 6-YEAR-OLDS IN CLASS	This data recoded for respondent confidentiality.
2087	A47YROL	A4 Q3C HOW MANY 7-YEAR-OLDS IN CLASS	This data recoded for respondent confidentiality.
2088	A48YROL	A4 Q3D HOW MANY 8-YEAR-OLDS IN CLASS	This data recoded for respondent confidentiality.
2089	A49YROL	A4 Q3E HOW MANY 9-YEAR-OLDS IN CLASS	This data recoded for respondent confidentiality.
2090	A410YRMO	A4 Q3F # 10-YEAR OLDS & OLDER IN CLASS	This data recoded for respondent confidentiality.
2091	A4TOTAG	A4 Q3G TOTAL CLASS ENROLLMENT (AGE)	This data recoded for respondent confidentiality.
2092	A4ASIAN	A4 Q4A # OF ASIAN/PACIFIC ISLANDERS	This data recoded for respondent confidentiality.
2093	A4HISP	A4 Q4B # OF HISPANICS (ALL RACES)	This data recoded for respondent confidentiality.
2094	A4BLACK	A4 Q4C # OF NON-HISPANIC BLACKS	This data recoded for respondent confidentiality.
2095	A4WHITE	A4 Q4D # OF NON-HISPANIC WHITES	This data recoded for respondent confidentiality.
2096	A4AMRIN	A4 Q4E # OF AMERICAN INDIANS	This data recoded for respondent confidentiality.
2097	A4RACEO	A4 Q4F # OF STUDENTS OF OTHER RACES	This data recoded for respondent confidentiality.
2098	A4TOTRA	A4 Q4G TOTAL CLASS ENROLLMENT (RACES)	This data recoded for respondent confidentiality.
2099	A4BOYS	A4 Q5A NUMBER OF BOYS IN CLASS	This data recoded for respondent confidentiality.
2100	A4GIRLS	A4 Q5B NUMBER OF GIRLS IN CLASS	This data recoded for respondent confidentiality.
2106	A4GIFT	A4 Q8A # CLASSIFIED AS GIFTED/TALENTED	This data recoded for respondent confidentiality.
2107	A4PRTGF	A4 Q8B # TAKE PART IN GIFTED/TALENTED	This data recoded for respondent confidentiality.
2118	A4DISAB	A4 Q11 NUMBER WITH DISABILITIES	This data recoded for respondent confidentiality.
2119	A4IMPAI	A4 Q12A COMMUNICATION IMPAIRMENTS	This data recoded for respondent confidentiality.
2120	A4LRNDI	A4 Q12B LEARNING DISABILITIES	This data recoded for respondent confidentiality.
2121	A4EMPRB	A4 Q12C SERIOUS EMOTIONAL PROBLEMS	This data recoded for respondent confidentiality.
2122	A4RETAR	A4 Q12D MENTAL RETARDATION	This data suppressed for respondent confidentiality.
2123	A4DELAY	A4 Q12E DEVELOPMENTAL DELAY	This data recoded for respondent confidentiality.
2124	A4VIS	A4 Q12F VISION IMPAIRMENTS	This data recoded for respondent confidentiality.
2125	A4HEAR	A4 Q12G HEARING IMPAIRMENTS	This data recoded for respondent confidentiality.
2126	A4ORTHO	A4 Q12H ORTHOPEDIC IMPAIRMENTS	This data recoded for respondent confidentiality.
2127	A4OTHER	A4 Q12I OTHER HEALTH IMPAIRMENTS	This data recoded for respondent confidentiality.
2128	A4MULTI	A4 Q12J MULTIPLE DISABILITIES	This data suppressed for respondent confidentiality.
2129	A4AUTSM	A4 Q12K AUTISM	This data suppressed for respondent confidentiality.
2130	A4TRAUM	A4 Q12L TRAUMATIC BRAIN INJURIES	This data recoded for respondent confidentiality.
2131	A4DEAF	A4 Q12M DEAFNESS AND BLINDNESS	This data suppressed for respondent confidentiality.
2132	A4OTDIS	A4 Q12N OTHER SPECIFY DISABILITIES	This data suppressed for respondent confidentiality.
2133	A4SPCIA	A4 Q13A SPECIAL DISABILITY SERVICES	This data recoded for respondent confidentiality.
2134	A4IEP	A4 Q13B IEP FOR CHILDREN W/ DISABILITY	This data recoded for respondent confidentiality.

Table 7-15.—Recoded and suppressed data on the ECLS-K First-Grade Year Public-Use File (continued)

FieldID	FieldName	FieldLabel	Comment
2135	A4SC504	A4 Q13C CHILDREN W/ SECTION 504 PLAN	This data recoded for respondent confidentiality.
2139	A4CVTNM	A4 Q15B STUDENTS SPEAK VIETNAMESE	This data suppressed for respondent confidentiality.
2140	A4CCHNS	A4 Q15C STUDENTS SPEAK CHINESE	This data suppressed for respondent confidentiality.
2141	A4CJPNS	A4 Q15D STUDENTS SPEAK JAPANESE	This data suppressed for respondent confidentiality.
2142	A4CKRN	A4 Q15E STUDENTS SPEAK KOREAN	This data suppressed for respondent confidentiality.
2143	A4CFLPN	A4 Q15F STUDENTS SPEAK FILIPINO LNG	This data suppressed for respondent confidentiality.
2144	A4OTASN	A4 Q15G STUDENTS SPEAK OTHR ASIAN LNG	This data suppressed for respondent confidentiality.
2145	A4OTLNG	A4 Q15H STUDENTS SPEAK ANOTHER LNG	This data suppressed for respondent confidentiality.
2146	A4LANOS	A4 Q15H SPECIFY STUDENTS' OTHER LANG	This data suppressed for respondent confidentiality.
2148	A4NUMLE	A4 Q17 NUMBER LEP STUDENTS IN CLASS	This data recoded for respondent confidentiality.
2149	A4NOESL	A4 Q18A LEP STUDENTS GET NO ESL	This data recoded for respondent confidentiality.
2150	A4ESLRE	A4 Q18B LEP STUDENTS GET IN-CLASS ESL	This data recoded for respondent confidentiality.
2151	A4ESLOU	A4 Q18C LEP STUDENTS GET OUTSIDE ESL	This data recoded for respondent confidentiality.
2154	A4TVTNM	A4 Q19C TCHRS SPEAK VIETNAMESE	This data suppressed for respondent confidentiality.
2155	A4TCHNS	A4 Q19D TCHRS SPEAK CHINESE	This data suppressed for respondent confidentiality.
2156	A4TJPNS	A4 Q19E TCHRS SPEAK JAPANESE	This data suppressed for respondent confidentiality.
2158	A4TFLPN	A4 Q19G TCHRS SPEAK A FILIPINO LNG	This data suppressed for respondent confidentiality.
2159	A4TOTAS	A4 Q19H TCHRS SPEAK OTHER ASIAN LNG	This data suppressed for respondent confidentiality.
2160	A4TOTLG	A4 Q19I TCHRS SPEAK ANOTHER LANGUAGE	This data suppressed for respondent confidentiality.
2161	A4LEPOS	A4 Q19I SPECIFY TCHRS OTHER LANGUAGE	This data suppressed for respondent confidentiality.
2163	A4ISPNH	A4 Q20B INSTRUCTION - SPANISH	This data suppressed for respondent confidentiality.
2164	A4IVTNM	A4 Q20C INSTRUCTION - VIETNAMESE	This data suppressed for respondent confidentiality.
2165	A4ICHNS	A4 Q20D INSTRUCTION - CHINESE	This data suppressed for respondent confidentiality.
2166	A4IJPNS	A4 Q20E INSTRUCTION - JAPANESE	This data suppressed for respondent confidentiality.
2167	A4IKRN	A4 Q20F INSTRUCTION - KOREAN	This data suppressed for respondent confidentiality.
2168	A4IFLPN	A4 Q20G INSTRUCTION - FILIPINO LNG	This data suppressed for respondent confidentiality.
2169	A4IOTAS	A4 Q20H INSTRUCTION - OTHER ASIAN LNG	This data suppressed for respondent confidentiality.
2170	A4IOTLN	A4 Q20I INSTRUCTION - OTHER LNG	This data suppressed for respondent confidentiality.
2171	A4LNGOS	A4 Q20I LANGUAGE OF INSTRUCTION-OTHER	This data suppressed for respondent confidentiality.
2512	B4KTGEND	B4K Q16 TEACHER'S GENDER	This data suppressed for respondent confidentiality.
2513	B4KYRBOR	B4K Q17 TEACHER'S YEAR OF BIRTH	This data recoded for respondent confidentiality.
2514	B4KHISP	B4K Q18 HISPANIC OR LATINO	This data suppressed for respondent confidentiality.
2515	B4KRACE1	B4K Q19 AMERICAN INDIAN OR ALASKA NATIVE	This data suppressed for respondent confidentiality.
2516	B4KRACE2	B4K Q19 ASIAN	This data suppressed for respondent confidentiality.
2518	B4KRACE4	B4K Q19 NATIVE HAWAIIAN OR OTHER PAC IS	This data suppressed for respondent confidentiality.
2520	B4KYRSTC	B4K Q20 NUMBER YEARS BEEN SCHOOL TEACHER	This data recoded for respondent confidentiality.
2521	B4KYRSPR	B4K Q21A YRS TEACHER TAUGHT PRESCHOOL	This data recoded for respondent confidentiality.
2522	B4KYRSKI	B4K Q21B YRS TEACHER TAUGHT KINDERGARTEN	This data recoded for respondent confidentiality.
2523	B4KYRSFS	B4K Q21C YRS TEACHER TAUGHT FIRST GRADE	This data recoded for respondent confidentiality.
2524	B4KYRS2T	B4K Q21D YRS TEACHER TAUGHT 2 TO 5 GRADE	This data recoded for respondent confidentiality.
2525	B4KYRS6P	B4K Q21E YRS TCHER TAUGHT 6 GRADE OR UP	This data recoded for respondent confidentiality.
2526	B4KYRSES	B4K Q21F YRS TEACHER TAUGHT ESL	This data recoded for respondent confidentiality.
2527	B4KYRSBI	B4K Q21G YRS TEACHER TAUGHT BILINGUAL ED	This data recoded for respondent confidentiality.
2528	B4KYRSSP	B4K Q21H YRS TEACHER TAUGHT SPECIAL ED	This data recoded for respondent confidentiality.
2529	B4KYRSPE	B4K Q21I YRS TEACHER TAUGHT PHYSICAL ED	This data recoded for respondent confidentiality.
2530	B4KYRSAR	B4K Q21J YRS TEACHER TAUGHT ART OR MUSIC	This data recoded for respondent confidentiality.
2531	B4KYRSCH	B4K Q22 YRS TCHER TAUGHT AT THIS SCHOOL	This data recoded for respondent confidentiality.
2532	B4KHGHST	B4K Q23 HIGHEST ED LEVEL TCHER ACHIEVED	This data recoded for respondent confidentiality.

Table 7-15.—Recoded and suppressed data on the ECLS-K First-Grade Year Public-Use File (continued)

FieldID	FieldName	FieldLabel	Comment
2598	B4TGEND	B4 Q11 TEACHER'S GENDER	This data suppressed for respondent confidentiality.
2599	B4YRBORN	B4 Q12 TEACHER'S YEAR OF BIRTH	This data recoded for respondent confidentiality.
2601	B4RACE1	B4 Q14 AMERICAN INDIAN OR ALASKA NATIVE	This data suppressed for respondent confidentiality.
2602	B4RACE2	B4 Q14 ASIAN	This data suppressed for respondent confidentiality.
2604	B4RACE4	B4 Q14 NATIVE HAWAIIAN OR OTHER PAC IS	This data suppressed for respondent confidentiality.
2606	B4YRSTC	B4 Q15 NUMBER YEARS BEEN SCHOOL TEACHER	This data recoded for respondent confidentiality.
2607	B4YRSPRE	B4 Q16A YRS TEACHER TAUGHT PRESCHOOL	This data recoded for respondent confidentiality.
2608	B4YRSKIN	B4 Q16B YRS TEACHER TAUGHT KINDERGARTEN	This data recoded for respondent confidentiality.
2609	B4YRSFST	B4 Q16C YRS TEACHER TAUGHT FIRST GRADE	This data recoded for respondent confidentiality.
2610	B4YRS2T5	B4 Q16D YRS TEACHER TAUGHT 2 TO 5 GRADE	This data recoded for respondent confidentiality.
2611	B4YRS6PL	B4 Q16E YRS TEACHER TAUGHT 6 GRADE OR UP	This data recoded for respondent confidentiality.
2612	B4YRSESL	B4 Q16F YRS TEACHER TAUGHT ESL	This data recoded for respondent confidentiality.
2613	B4YRSBIL	B4 Q16G YRS TEACHER TAUGHT BILINGUAL ED	This data recoded for respondent confidentiality.
2614	B4YRSSPE	B4 Q16H YRS TEACHER TAUGHT SPECIAL ED	This data recoded for respondent confidentiality.
2615	B4YRSPE	B4 Q16I YRS TEACHER TAUGHT PHYSICAL ED	This data recoded for respondent confidentiality.
2616	B4YRSART	B4 Q16J YRS TEACHER TAUGHT ART OR MUSIC	This data recoded for respondent confidentiality.
2617	B4YRSCH	B4 Q17 YRS TEACHER TAUGHT AT THIS SCHOOL	This data recoded for respondent confidentiality.
2618	B4HGHSTD	B4 Q18 HIGHEST ED LEVEL TEACHER ACHIEVED	This data recoded for respondent confidentiality.
2666	T4KFLBHN	T4K Q1 FELL BEHIND DUE TO HEALTH	This data suppressed for respondent confidentiality.
2671	T4KPLLES	T4K Q2E PULL-OUT ESL PROGRAM	This data suppressed for respondent confidentiality.
2672	T4KINCES	T4K Q2F IN-CLASS ESL	This data suppressed for respondent confidentiality.
2673	T4KGFTTA	T4K Q2G GIFTED AND TALENTED PROGRAM	This data suppressed for respondent confidentiality.
2674	T4KBEHPR	T4K Q2H PROGRAM FOR BEHAVIORAL PROBLEMS	This data suppressed for respondent confidentiality.
2677	T4KTT1EN	T4K Q3C TITLE 1 ENGLISH/LANGUAGE ARTS	This data suppressed for respondent confidentiality.
2678	T4KTT1CM	T4K Q3D TITLE 1 COMBINED SUBJECTS	This data suppressed for respondent confidentiality.
2679	T4KTT1ES	T4K Q3E TITLE 1 ESL/BILINGUAL	This data suppressed for respondent confidentiality.
2680	T4KTT1SP	T4K Q3F TITLE 1 HANDICAPPED/SPECIAL ED	This data suppressed for respondent confidentiality.
2697	T4KLNGTM	T4K Q15 LENGTH OF TIME IN CLASSROOM	This data suppressed for respondent confidentiality.
2698	T4KGRADE	T4K Q16 GRADE CHILD ENROLLED	This data suppressed for respondent confidentiality.
2699	T4KTYPCL	T4K Q17 TYPE OF CLASS	This data suppressed for respondent confidentiality.
2729	T4PLLESL	T4 Q1E PULL-OUT ESL PROGRAM	This data suppressed for respondent confidentiality.
2730	T4INCESL	T4 Q1F IN-CLASS ESL	This data suppressed for respondent confidentiality.
2731	T4GFTRD	T4 Q1G GIFTED PROGRAM IN READING	This data suppressed for respondent confidentiality.
2732	T4GFTMTH	T4 Q1H GIFTED PROGRAM IN MATHEMATICS	This data suppressed for respondent confidentiality.
2733	T4BEHPRB	T4 Q1I PROGRAM FOR BEHAVIORAL PROBLEMS	This data suppressed for respondent confidentiality.
2736	T4TT1ENG	T4 Q2C TITLE 1 ENGLISH/LANGUAGE ARTS	This data suppressed for respondent confidentiality.
2737	T4TT1CMB	T4 Q2D TITLE 1 COMBINED SUBJECTS	This data suppressed for respondent confidentiality.
2738	T4TT1ES	T4 Q2E TITLE 1 ESL/BILINGUAL	This data suppressed for respondent confidentiality.
2739	T4TT1SP	T4 Q2F TITLE 1 HANDICAPPED/SPECIAL ED	This data suppressed for respondent confidentiality.
2750	T4FLBHND	T4 Q8 FELL BEHIND DUE TO HEALTH	This data suppressed for respondent confidentiality.
2761	T4LNGTM	T4 Q16 LENGTH OF TIME IN CLASSROOM	This data suppressed for respondent confidentiality.
2762	T4GRADE	T4 Q17 GRADE CHILD ENROLLED	This data suppressed for respondent confidentiality.
2766	S4NUMDAY	S4 Q1 NUMBER OF DAYS MUST ATTEND	This data recoded for respondent confidentiality.
2767	S4ADA	S4 Q2 % AVERAGE DAILY ATTENDANCE FOR YR.	This data recoded for respondent confidentiality.
2769	S4ANUMCH	S4 Q3A # ENROLLED AROUND 10/1/1999	This data recoded for respondent confidentiality.
2770	S4BNUMCH	S4 Q3B # ENROLLED SINCE 10/1/1999	This data recoded for respondent confidentiality.
2771	S4CNUMCH	S4 Q3C # WHO LEFT - DIDN'T RETURN	This data recoded for respondent confidentiality.
2772	S4UNGRAD	S4 Q4 GRADE LEVEL-UNGRADED	This data suppressed for respondent confidentiality.

Table 7-15.—Recoded and suppressed data on the ECLS-K First-Grade Year Public-Use File (continued)

FieldID	FieldName	FieldLabel	Comment
2775	S4TRANS	S4 Q4 GRADE LEVEL-TRANSITIONAL	This data suppressed for respondent confidentiality.
2777	S4PRE1	S4 Q4 GRADE LEVEL-PREFIRST	This data suppressed for respondent confidentiality.
2778	S4GRADE1	S4 Q4 GRADE LEVEL-FIRST GRADE	This data suppressed for respondent confidentiality.
2779	S4SECOND	S4 Q4 GRADE LEVEL-SECOND GRADE	This data suppressed for respondent confidentiality.
2786	S4NINTH	S4 Q4 GRADE LEVEL-NINTH GRADE	This data suppressed for respondent confidentiality.
2787	S4TENTH	S4 Q4 GRADE LEVEL-TENTH GRADE	This data suppressed for respondent confidentiality.
2788	S411TH	S4 Q4 GRADE LEVEL-ELEVENTH GRADE	This data suppressed for respondent confidentiality.
2789	S412TH	S4 Q4 GRADE LEVEL-TWELFTH GRADE	This data suppressed for respondent confidentiality.
2792	S4MAGSKL	S4 Q6B IS IT A MAGNET SCHOOL	This data suppressed for respondent confidentiality.
2794	S4BIASKL	S4 Q6D IS IT A TRIBAL SCHOOL	This data suppressed for respondent confidentiality.
2795	S4SPEDSK	S4 Q6E IS IT A SPECIAL ED SCHOOL	This data suppressed for respondent confidentiality.
2800	S4PRIVRD	S4 Q7A3 IS IT A PRIVATE ORDER	This data suppressed for respondent confidentiality.
2802	S4NAISKL	S4 Q7C PRIVATE-ACCREDITED BY NAIS	This data suppressed for respondent confidentiality.
2804	S4PVTSPD	S4 Q7E IS IT SPECIAL EDUCATION	This data suppressed for respondent confidentiality.
2807	S4EMPHAS	S4 Q9 WHAT'S THE SCHOOL EMPHASIS	This data suppressed for respondent confidentiality.
2808	S4PREK	S4 Q10 PRE-K PARTICIPATES	This data suppressed for respondent confidentiality.
2809	S4REDDYN	S4 Q10 TRANSITIONAL PARTICIPATES	This data suppressed for respondent confidentiality.
2810	S4KINGAR	S4 Q10 KINDERGARTEN PARTICIPATES	This data suppressed for respondent confidentiality.
2811	S4PRE1ST	S4 Q10 PREFIRST PARTICIPATES	This data suppressed for respondent confidentiality.
2812	S4FIR1ST	S4 Q10 1ST GRADE PARTICIPATES	This data suppressed for respondent confidentiality.
2813	S4SEC2ND	S4 Q10 2ND GRADE PARTICIPATES	This data suppressed for respondent confidentiality.
2814	S4THI3RD	S4 Q10 3RD GRADE PARTICIPATES	This data suppressed for respondent confidentiality.
2815	S4FOR4TH	S4 Q10 4TH GRADE PARTICIPATES	This data suppressed for respondent confidentiality.
2816	S4FIF5TH	S4 Q10 5TH GRADE PARTICIPATES	This data suppressed for respondent confidentiality.
2817	S4SIX6TH	S4 Q10 6TH GRADE PARTICIPATES	This data suppressed for respondent confidentiality.
2818	S4SEV7TH	S4 Q10 7TH GRADE PARTICIPATES	This data suppressed for respondent confidentiality.
2819	S4ATE8TH	S4 Q10 8TH GRADE PARTICIPATES	This data suppressed for respondent confidentiality.
2820	S4NIN9TH	S4 Q10 9TH GRADE PARTICIPATES	This data suppressed for respondent confidentiality.
2821	S4TEN10	S4 Q10 10TH GRADE PARTICIPATES	This data suppressed for respondent confidentiality.
2822	S4ELEV11	S4 Q10 11TH GRADE PARTICIPATES	This data suppressed for respondent confidentiality.
2823	S4TWEL12	S4 Q10 12TH GRADE PARTICIPATES	This data suppressed for respondent confidentiality.
2824	S4ASNPCT	S4 Q11A PERCENT OF ASIAN STUDENTS	This data suppressed for respondent confidentiality.
2826	S4HSPPCT	S4 Q11B PERCENT OF HISPANIC STUDENTS	This data recoded for respondent confidentiality.
2828	S4BLKPCT	S4 Q11C PERCENT OF BLACK STUDENTS	This data recoded for respondent confidentiality.
2830	S4WHTPCT	S4 Q11D PERCENT OF WHITE STUDENTS	This data suppressed for respondent confidentiality.
2832	S4INDPCT	S4 Q11E PERCENT OF AMERICAN INDIANS	This data suppressed for respondent confidentiality.
2834	S4OTHPCT	S4 Q11F PERCENT OF OTHER STUDENTS	This data recoded for respondent confidentiality.
2836	S4ELIBRK	S4 Q12A1 ELIGIBLE FOR FREE BREAKFAST	This data suppressed for respondent confidentiality.
2837	S4PARBRK	S4 Q12A2 PARTICIPATES IN BREAKFAST	This data suppressed for respondent confidentiality.
2838	S4ELILNC	S4 Q12B1 ELIGIBLE FOR FREE LUNCH	This data suppressed for respondent confidentiality.
2839	S4PARLNC	S4 Q12B2 PARTICIPATES IN FREE LUNCH	This data suppressed for respondent confidentiality.
2840	S4ELIRED	S4 Q12C1 ELIGIBLE IN REDUCED-PRICE LUNCH	This data suppressed for respondent confidentiality.
2841	S4PARRED	S4 Q12C2 PARTICIPATES IN RED-PRICE LUNCH	This data suppressed for respondent confidentiality.
2850	S4CHSRVD	S4 Q16 NUMBER SERVED BY TITLE 1	This data recoded for respondent confidentiality.
2862	S4CHLDNM	S4 Q18 # OF CHILDREN SITE ACCOMMODATES	This data recoded for respondent confidentiality.
2873	S4INSTCM	S4 Q20A # INSTRUCTIONAL COMPUTERS	This data recoded for respondent confidentiality.
2874	S4ADMNCM	S4 Q20B # INSTRUCT/ADMIN COMPUTERS	This data recoded for respondent confidentiality.
2875	S4TOTCM	S4 Q20C TOTAL NUMBER OF COMPUTERS	This data recoded for respondent confidentiality.

Table 7-15.—Recoded and suppressed data on the ECLS-K First-Grade Year Public-Use File (continued)

FieldID	FieldName	FieldLabel	Comment
2883	S4CMNITY	S4 Q23 COMMUNITY SCHOOL IS LOCATED IN	This data recoded for respondent confidentiality.
2932	S4NMBKPT	S4 Q31 NUMBER RETAINED GRADE-1ST	This data recoded for respondent confidentiality.
2937	S4MIGRSM	S4 Q32E SUMMER MIGRANT PROGRAMS	This data suppressed for respondent confidentiality.
2961	S4NMADMN	S4 Q37A2 NUMBER OF ADMINISTRATORS	This data recoded for respondent confidentiality.
2963	S4NMTECH	S4 Q37B2 NUMBER OF TEACHERS	This data recoded for respondent confidentiality.
2969	S4NMOMDD	S4 Q37E2 NUMBER OF PARENTS	This data recoded for respondent confidentiality.
2971	S4NMREPS	S4 Q37F2 NUMBER OF REPRESENTATIVES	This data recoded for respondent confidentiality.
2973	S4NUMOTH	S4 Q37G2 NUMBER OF OTHER GROUPS	This data recoded for respondent confidentiality.
2977	S4BILING	S4 Q40 BILINGUAL SERVICES PERCENT -1ST	This data suppressed for respondent confidentiality.
2978	S4ESLONL	S4 Q40 ESL SERVICES PERCENT -1ST	This data suppressed for respondent confidentiality.
2979	S4ESLBIL	S4 Q40 ESL AND BILINGUAL PERCENT -1ST	This data suppressed for respondent confidentiality.
2980	S4YRSESL	S4 Q41A YEARS FOR ESL SERVICES-1ST	This data recoded for respondent confidentiality.
2981	S4YRSBIL	S4 Q41B YEARS FOR BILINGUAL SERVICES-1ST	This data recoded for respondent confidentiality.
2982	S4SRVBTH	S4 Q41C YRS FOR ESL & BILINGUAL SER-1ST	This data recoded for respondent confidentiality.
3007	S4FTETOT	S4 Q49 TOTAL FTE SCHOOL STAFF	This data recoded for respondent confidentiality.
3008	S4ADMFTE	S4 Q50A FTE ADMINISTRATIVE STAFF	This data recoded for respondent confidentiality.
3009	S4TCHFTE	S4 Q50B FTE CLASSROOM TEACHERS	This data recoded for respondent confidentiality.
3010	S4GYMFTE	S4 Q50C FTE GYM/DRAMA/ART TEACHERS	This data recoded for respondent confidentiality.
3011	S4TT1FTE	S4 Q50D FTE TITLE 1 TEACHERS	This data recoded for respondent confidentiality.
3012	S4AIDFTE	S4 Q50E FTE TEACHER AIDES	This data recoded for respondent confidentiality.
3013	S4SPEFTE	S4 Q50F FTE SPECIAL ED TEACHERS	This data recoded for respondent confidentiality.
3014	S4SPAFTE	S4 Q50G FTE SPECIAL ED AIDES	This data recoded for respondent confidentiality.
3015	S4ESLFTE	S4 Q50H FTE ESL-BILINGUAL TEACHERS	This data recoded for respondent confidentiality.
3016	S4ESAFTE	S4 Q50I FTE ESL-BILINGUAL AIDES	This data recoded for respondent confidentiality.
3017	S4LIBFTE	S4 Q50J LIBRARY/MEDIA SPECIALISTS	This data recoded for respondent confidentiality.
3018	S4THRFTE	S4 Q50K SPEECH/PHYSICAL THERAPISTS	This data recoded for respondent confidentiality.
3019	S4NRSFTE	S4 Q50L FTE SCHOOL NURSES	This data recoded for respondent confidentiality.
3020	S4REAFTE	S4 Q50M FTE READING SPECIALIST TCHRS	This data recoded for respondent confidentiality.
3021	S4MSCFTE	S4 Q50N FTE MATH/SCIENCE SPECIALISTS	This data recoded for respondent confidentiality.
3022	S4FRLFTE	S4 Q50O FTE FOREIGN LNG SPECIALISTS	This data recoded for respondent confidentiality.
3025	S4TEAASN	S4 Q53 % ASIAN TEACHERS	This data suppressed for respondent confidentiality.
3027	S4ETHNIC	S4 Q53 % HISPANIC/LATINO TEACHERS	This data recoded for respondent confidentiality.
3029	S4TEABLK	S4 Q53 % BLACK TEACHERS	This data recoded for respondent confidentiality.
3033	S4TEAIND	S4 Q53 % AMERICAN INDIAN TEACHERS	This data suppressed for respondent confidentiality.
3090	S4BRTHYR	S4 Q61 YEAR PRINCIPAL WAS BORN	This data recoded for respondent confidentiality.
3091	S4ORIGIN	S4 Q62 PRINCIPAL IS HISPANIC/LATINO	This data suppressed for respondent confidentiality.
3092	S4RACE1	S4 Q63A PRINCIPAL IS AMERICAN INDIAN	This data suppressed for respondent confidentiality.
3093	S4RACE2	S4 Q63B PRINCIPAL IS ASIAN	This data suppressed for respondent confidentiality.
3094	S4RACE3	S4 Q63C PRINCIPAL IS BLACK	This data suppressed for respondent confidentiality.
3095	S4RACE4	S4 Q63D PRINCIPAL IS HAWAIIAN OR PAC IS	This data suppressed for respondent confidentiality.
3096	S4RACE5	S4 Q63E PRINCIPAL IS WHITE	This data suppressed for respondent confidentiality.
3097	S4YSTCH	S4 Q64A NUMBER OF YRS TEACHING	This data recoded for respondent confidentiality.
3098	S4TOTPRI	S4 Q64B NUMBER OF YRS AS PRINCIPAL	This data recoded for respondent confidentiality.
3099	S4PRINHR	S4 Q64C NUMBER YRS A PRINCIPAL HERE	This data recoded for respondent confidentiality.
3105	S4YRESL	S4 Q65F YEARS TEACHING ESL PROGRAM	This data recoded for respondent confidentiality.
3106	S4YRBILG	S4 Q65G YEARS TEACHING BILINGUAL ED	This data recoded for respondent confidentiality.
3119	S4EDLVL	S4 Q67 HIGHEST LEVEL OF EDUCATION	This data recoded for respondent confidentiality.

9. CREATING A LONGITUDINAL FILE

Longitudinal analyses with the Early Childhood Longitudinal Study-Kindergarten Class of 1998-99 (ECLS-K) can be conducted both "within school year" and "across school years." Examples of within-year analyses are those that look at children's growth in cognitive scores between fall and spring of kindergarten or between fall and spring of first grade. Such analyses do not require the combined use of kindergarten and first grade data. They can be conducted using just the kindergarten base year files or just the first grade files and are, therefore, not discussed in this chapter. Cross-year (K-1) analyses, on the other hand, are those that combine information from both the kindergarten year and the first grade year and are the focus of this chapter.

This chapter describes how to combine (or merge) the first grade and kindergarten files to create cross-year files for K-1 longitudinal analyses. The information contained in this chapter applies to users of the base year and first grade files. Users of the public-use files can consider using the public-use longitudinal file briefly described in chapter 1, which combines data from the base and first grade years. It contains longitudinal weights so that analysts can examine children's growth and development between kindergarten and first grade. Although it is somewhat streamlined, it contains most of the variables in the restricted-use files.

This chapter begins with a discussion of K-1 longitudinal analyses and the types of research questions that can be addressed with cross-year files. It then describes the merging procedures and the K-1 longitudinal weights.

9.1 Conducting Longitudinal Analyses

As described in chapter 1, one of the primary goals of the ECLS-K is to understand how children's early experiences influence their transition into kindergarten and their progression through the early elementary school years. A major strength of the ECLS-K design is that it captures important aspects of children's experiences as they occur. Thus, information about children's transition into kindergarten is measured in the fall of their kindergarten year and again in the spring. Capturing this information as it occurs means that the information is not distorted by faulty memory or by revisions to memory based on subsequent experiences. In addition, information from earlier points in time can be used as predictors of later events and experiences thereby strengthening the ability of researchers to make causal inferences.

In conducting K-1 longitudinal analyses with the ECLS-K data, it is important to keep in mind the sample design described in chapter 4. In particular, certain features of the design must be considered in conducting K-1 analyses. First, because the first grade data are only being released as child-based files, all K-1 analyses will, of necessity, be child-based. Second, the first grade data are not representative of all first grade schools or classrooms or teachers in the United States. Researchers conducting K-1 analyses should not attempt to use the data to describe the population of first grade classrooms, of teachers, or of schools containing first grades. However, information about the first grade schools can be used in the child-based analyses to examine, for example, the influence of the school environment on children's learning. Users may also examine the influence of the kindergarten year school characteristics on children's first grade experiences.

9.2 Examples of Research Questions

There are a variety of research questions that can be examined using the K-1 longitudinal files; the following are some examples:

1. How much do children's reading and math skills increase between the fall of kindergarten and the spring of first grade?

2. Do measures of school readiness at the beginning of kindergarten predict children's skill and knowledge levels at the end of first grade?

3. What family background characteristics (e.g., family poverty, parent education, maternal employment) affect children's later school outcomes?

4. Do children who easily adapted to a school setting in kindergarten do better in first grade than their peers who experienced more difficulty settling into school or are there no lingering effects of a slow adjustment to kindergarten?

5. To what extent do children experience gains or losses in their learning between spring of kindergarten and fall-first grade when school is out?

6. Do factors, such as participation in enrichment programs, help to minimize summer learning loss, after controlling for family background factors?

To study these and similar questions, researchers would combine information from two or more rounds of data collection, across the kindergarten and first grade years. For the first question, the researcher would need to examine differences between fall-kindergarten and spring-first grade assessment scores. To do this, one would combine fall-kindergarten data with spring-first grade data. Similarly, questions 2 and 3 (regarding the relationship between readiness at kindergarten entry—or maternal employment in that time frame—and first grade outcomes) would be examined by combining data from the same two timepoints. Note that for question 3 one would need to include data from the parent interview in the base year.

Researchers who want to examine the influence of children's kindergarten adjustment on their first grade performance, as in question 4, might use data from three rounds (i.e., fall-kindergarten, spring-kindergarten, and spring-first grade). For example, one could create variables from fall-kindergarten and spring-kindergarten to measure adjustment during kindergarten and then relate those variables to outcomes in the spring of first grade.

K-1 files can also be used to answer questions 5 and 6, focusing on summer learning gains or losses. For these questions, a researcher would combine data from spring-kindergarten with data from fall-first grade. Because the fall-first grade data are available for only a subsample of the ECLS-K children, these analyses would be restricted to the 5,424 children in the fall-first grade sample.

9.3 Merging Base Year Child-Level Data with the First Grade Child-Level Data

To create a K-1 data file, which combines data from the base year and first grade data collections, an analyst should use the ECLS-K Restricted-Use or Public-Use Base Year Electronic Code Book (ECB) (NCES 2000-079 or NCES 2001-029) and the ECLS-K First Grade Restricted-Use ECB (NCES 2001-102). The same procedures can be followed by users who wish to create a public-use K-1 longitudinal file themselves. Perform the following steps to merge the base year child-level variables needed for analysis with the first grade child-level variables needed:

1. Select the variables to be analyzed from the base year ECB child catalog and the variable CHILDID. This creates a "working taglist" (see section 8.4 in chapter 8 for more detail on how to create a working taglist).

2. Run the program generated after extraction to create a base year data set (DATA1).

3. Using the child catalog from the First-Grade Restricted-Use ECB, select the variables to be analyzed and the variable CHILDID.

4. Run the program generated after extraction to create a first grade data set (DATA2).

5. Sort DATA1 and DATA2 by CHILDID.

6. Merge DATA1 and DATA2 by CHILDID.

This merged file will contain 21,399 cases, some of which will not have K-1 longitudinal weights. For example, base year respondents who did not participate in either the fall or spring of first grade, movers who were not included in the first grade sample and children included through the first grade freshening process, will not have any K-1 longitudinal weights. To select cases with K-1 longitudinal data, a user can use a K-1 longitudinal weight appropriate to the analysis.

9.4 K-1 Longitudinal Weights

9.4.1 Type of K-1 Longitudinal Weights

K-1 longitudinal weights are used when analyzing data in the K-1 file created by merging base year and first grade data, whereas cross-sectional weights are used for analyses within one round of data collection. There are several sets of K-1 longitudinal weights computed for children with complete data from different combinations of rounds. All K-1 longitudinal weights are child-level weights. There are no K-1 longitudinal weights at the school or teacher level since school- and teacher-level weights are not computed for the first grade year.

The K-1 longitudinal weights, available on the first grade restricted-use ECB, are described in table 9-1. This table is designed to help users choose appropriate weights for their analysis. First, decide which two or more points in time are the focus of the analysis. The analysis could pertain to two points in time (any two of spring-kindergarten, fall-first grade, and spring-first grade), three points in time (any three of fall-kindergarten, spring-kindergarten, fall-first grade, and spring-first grade), or four points in time (all four rounds of data). For example, if the analysis uses spring-kindergarten and fall-first grade data, then the appropriate weights would be those beginning with C23 (denoting child-level data from round 2 AND round 3). Second, consider the source of the data, which also affects the choice of the weight. In table 9-1, details under "to be used in the analysis of …" provide guidance based on whether the data were collected through the child assessments, parent interviews, or teacher questionnaires A or B.

For the same example noted earlier, the two weights available are C23CW0 and C23PW0. If parent data from spring-kindergarten and fall-first grade are needed for the analysis, then C23PW0 should be used.

Base year longitudinal weights for the analysis of the base year data (within the kindergarten year) alone are described in the base year user's manuals. First grade longitudinal weights for the analysis of the first grade data (within the first grade year) alone are described in chapter 4 of this user's manual.

Table 9-1.— ECLS-K: K-1 longitudinal weights

K-1 longitudinal (panel) weights	
Weight	to be used for analysis of ...
C23CW0	child direct assessment data from BOTH spring-kindergarten and fall-first grade, alone or in conjunction with any combination of a limited set of child characteristics (e.g., age, sex, race-ethnicity).
C23PW0	parent interview data from BOTH spring-kindergarten and fall-first grade.
C123CW0	child direct assessment data from fall- AND spring-kindergarten AND fall-first grade, alone or in conjunction with any combination of a limited set of child characteristics (e.g., age, sex, race-ethnicity).
C123PW0	parent interview data from fall- AND spring-kindergarten AND fall-first grade.
C24CW0	child direct assessment data from BOTH spring-kindergarten and spring-first grade, alone or in conjunction with any combination of a limited set of child characteristics (e.g., age, sex, race-ethnicity).
C24PW0	parent interview data from BOTH spring-kindergarten and spring-first grade.
C124CW0	child direct assessment data from fall-kindergarten AND spring-kindergarten AND spring-first grade, alone or in conjunction with any combination of a limited set of child characteristics (e.g., age, sex, race-ethnicity).
C124PW0	parent interview data from fall-kindergarten AND spring-kindergarten AND spring-first grade.
C1_4CW0	child direct assessment data from all four rounds of data collection, alone or in conjunction with any combination of a limited set of child characteristics (e.g., age, sex, and race-ethnicity).
C1_4PW0	parent interview data from all four rounds of data collection.
Y2COMW0	child direct assessment data from fall-kindergarten AND spring-kindergarten AND spring-first grade, in conjunction with parent and/or teacher data from spring-first grade, AND one or more base year rounds of parent and/or teacher data.

K-1 longitudinal weights are used to produce estimates of differences between two or more rounds of data collection spanning across both kindergarten and first grade. Simple examples involving two rounds of data collection are the differences in children's mean assessment scores between spring kindergarten and spring first grade using C24CW0, and the difference in social rating scores as reported by parents using C24PW0 (social rating scores as reported by teachers and parents are not available for fall-first grade). K-1 longitudinal weights are also used to study the characteristics of children who were assessed in two or more rounds of data collection. For example, one can study the characteristics of kindergarten children who went to summer schools and the effect of summer schools on their assessment scores in fall-first grade for children who were assessed in both spring-kindergarten and fall-first grade. In this case, C23PW0 will be used to study the characteristics of the children as reported by their parents, and C23CW0 will be used to estimate the difference in assessment scores between spring-kindergarten and fall-first grade. As noted earlier, any longitudinal analysis that uses data from fall-first grade will be limited to a 27 percent subsample of children.

There may be combinations of data for which no weights were developed. For further advice on which weights to use when analyzing a complex combination of data, contact NCES at *ECLS@ed.gov.*

9.4.2 Weighting Procedures

In this section we discuss the statistical procedures used to produce the K-1 longitudinal weights, which are nearly identical to the procedures used for the cross-sectional and within first grade weights (see chapter 4). The differences are primarily in how eligible respondents are defined and in how adjustment cells are created. For example, in computing weight C23CW0, a respondent was defined as a child for whom both cross-sectional weights, C2CW0 and C3CW0, are nonzero. A child with a nonzero C23CW0 has both spring-kindergarten and fall-first grade scorable cognitive assessment data, or was excluded from the cognitive assessments because he or she is a child with disabilities or has a non-English and non-Spanish home language and was not able to pass the Oral Language Development Scale (OLDS).

The first stages of weighting were to:

- Adjust the base year school weight for base year school nonresponse, and
- Compute the base year child weight.

Base year child weights were then adjusted for:

- Subsampling of schools for the study (weights involving fall-first grade only),
- Subsampling of movers (all weights), and
- Child nonresponse (all weights).

Since the K-1 longitudinal weights do not concern children brought into the sample in spring-first grade through sample freshening, the adjustments having to do with freshening do not apply to longitudinal weights. The final stage of weighting was to rake the final adjusted weights to sample control totals. For a description of each stage of weighting, see section 4.3. The adjustment cells vary for each type of weight, depending on the characteristics of the schools, children or parents involved.

9.4.3 Characteristics of Longitudinal Weights

The statistical characteristics of the longitudinal weights are presented in table 9-2. For each weight, the number of cases with nonzero values is presented together with the mean weight, the standard deviation, the coefficient variation (i.e., the standard deviation as a percentage of the mean weight), the minimum value of the weight, the maximum value of the weight, the skewness, the kurtosis, and the sum of weights.

Table 9-2.—Characteristics of child-level K-1 longitudinal weights

Variable name	Number of cases	Mean	Standard deviation	CV (× 100)	Minimum	Maximum	Skewness	Kurtosis	Sum
C23CW0	5,216	739.84	587.55	79.42	68.23	7,182.37	3.98	21.56	3,858,997
C23PW0	4,861	793.83	515.75	64.97	84.26	5,853.21	2.97	13.04	3,858,805
C123CW0	4,729	815.99	646.25	79.20	76.08	7,696.79	3.89	21.55	3,858,824
C123PW0	4,295	898.37	597.89	66.55	95.35	6,421.30	3.05	14.20	3,858,492
C24CW0	16,371	234.81	200.69	85.47	1.78	3,272.40	4.22	31.65	3,844,009
C24PW0	14,938	257.25	198.94	77.34	1.93	2,580.41	3.30	19.64	3,842,784
C124CW0	15,001	256.28	228.52	89.17	1.54	3,877.43	3.71	24.60	3,844,472
C124PW0	13,413	286.40	214.80	75.00	2.06	3,275.79	3.84	26.53	3,841,463
C1_4CW0	4,542	847.78	639.83	75.47	77.56	7,528.68	3.49	18.68	3,850,619
C1_4PW0	4,012	959.07	617.93	64.43	108.75	6,780.92	2.86	13.48	3,847,785
Y2COMW0	13,983	274.83	241.55	87.89	2.03	3,803.82	4.26	29.97	3,842,961

The difference in the estimate of the population of students (sum of weights) between the different panels of students and types of weights is due to a combination of factors, among them: (1) the number of base year respondents who became ineligible (due to death, leaving the country, or being a nonsampled mover) after the base year, and (2) the adjustment of the weights for the children of unknown eligibility.

9.4.4 Variance Estimation

For each K-1 full sample weight listed in table 9-2, a set of replicate weights was calculated. Replicate weights are used in the jackknife replication method to estimate the standard errors of survey estimates. Any adjustments done to the full sample weights were repeated for the replicate weights.

For any full sample weight involving fall-first grade, there are 40 replicate weights. For all other weights, there are 90 replicate weights. The smaller number of replicates is due to the fact that only 30 percent of the full sample of schools were included in the fall-first grade subsample. See chapter 4 for a description of the fall-first grade sample and how the replicate weights were created. Each replicate weight variable name has the same weight prefix as for the full sample weight variable name. For example, the replicate weights for C23CW0 are C23CW1 through C23CW40; the replicate weights for C24CW0 are C24CW1 through C24CW90.

Stratum and first-stage unit identifiers used with the Taylor Series method are provided for each of the K-1 longitudinal weights in the file. They are described in table 9-3. For a description of the Taylor Series method, see section 4.4.2.

Specifications for computing standard errors are given in table 9-4. For each type of analysis described in table 9.4, users can choose between the replication method and the Taylor Series method for computing standard errors.

Table 9-3.—ECLS-K Taylor Series stratum and first-stage unit identifiers

Variable name	Description
C23CSTR	Sampling stratum–spring-kindergarten/fall-first grade longitudinal C-weights
C23CPSU	First-stage primary sampling unit within stratum–spring-kindergarten/fall-first grade longitudinal C-weights
C23PSTR	Sampling stratum–spring-kindergarten/fall-first grade longitudinal P-weights
C23PPSU	First-stage primary sampling unit within stratum–spring-kindergarten/fall-first grade longitudinal P-weights
C123CSTR	Sampling stratum–base year/fall-first grade longitudinal C-weights
C123CPSU	First-stage primary sampling unit within stratum–base year/fall-first grade longitudinal C-weights
C123PSTR	Sampling stratum–base year/fall-first grade longitudinal P-weights
C123PPSU	First-stage primary sampling unit within stratum–base year/fall-first grade longitudinal P-weights
C24CSTR	Sampling stratum–spring-kindergarten/spring-first grade longitudinal C-weights
C24CPSU	First-stage primary sampling unit within stratum–spring-kindergarten/spring-first grade longitudinal C-weights
C24PSTR	Sampling stratum–spring-kindergarten/spring-first grade longitudinal P-weights
C24PPSU	First-stage primary sampling unit within stratum–spring-kindergarten/spring-first grade longitudinal P-weights
C124CSTR	Sampling stratum–base year/spring-first grade longitudinal C-weights
C124CPSU	First-stage primary sampling unit within stratum–base year/spring-first grade longitudinal C-weights
C124PSTR	Sampling stratum–base year/spring-first grade longitudinal P-weights
C124PPSU	First-stage primary sampling unit within stratum–base year/spring-first grade longitudinal P-weights
C1_4CSTR	Sampling stratum–base year/first grade longitudinal C-weights
C1_4CPSU	First-stage primary sampling unit within stratum–base year/first grade longitudinal C-weights
C1_4PSTR	Sampling stratum–base year/first grade longitudinal P-weights
C1_4PPSU	First-stage primary sampling unit within stratum–base year/first grade longitudinal P-weights
Y2COMSTR	Sampling stratum–K-1 combined longitudinal weights
Y2COMPSU	First-stage primary sampling unit within stratum–K-1 combined longitudinal weights

For the replication method using WesVar, the full sample weight, the replicate weights, and the method of replication are required parameters. Variance estimation using the ECLS-K data should be done using the paired jackknife method (JK2). As an example, to compute the mean difference in reading scores between spring-kindergarten and spring-first grade and their standard errors, users need to specify C24CW0 as the full sample weight, C24CW1 to C24CW90 as the replicate weights, and JK2 as the method of replication.

For the Taylor Series method using either SUDAAN, SAS, or STATA, the full sample weight, the sample design, the nesting stratum, and primary sampling unit (PSU) variables are required. For the same example earlier, the full sample weight (C24CW0), the with replacement sample design (WR), the stratum variable (C24CSTR) and the PSU variable (C24CPSU) must be specified.

9.4.5 Design Effects

An important analytic device is to compare the statistical efficiency of survey estimates with what would have been obtained in a hypothetical and usually impractical simple random sample (SRS) of the same size. For a discussion of design effects and their use, see section 4.5. In this section, design effects are presented for selected illustrative estimates produced using kindergarten-first grade longitudinal weights. The tables that follow show estimates, standard errors, and design effects for selected means and proportions based on the ECLS-K child and parent data. For each survey item, the tables present the number of cases, the estimate, the standard error taking into account the actual sample design (Design SE), the standard error assuming SRS (SRS SE), the root design effect (DEFT), and the design effect (DEFF). Standard errors (Design SE) were produced using JK2.

Standard errors and design effects are presented in tables 9-5 to 9-9. In tables 9-5 and 9-6, data items are from the parent interview data. Full sample weights were used to compute the estimates, then the corresponding replicate weights were used to compute standard errors and design effects. In table 9-5, C24PW0 and the corresponding replicate weights were used. In table 9-6, C1_4PW0 and the corresponding replicate weights were used.

Table 9-4.—Specifications for computing standard errors

Type of analysis	Full sample weight	Computing standard errors: Replication method (WesVarPC): ID	Replicate weights	Jackknife method	Taylor Series method (SUDAAN & STATA): Sample design	Nesting variables	Approximating sampling errors: DEFT (Average root design effect)
Spring-kindergarten/fall-first grade longitudinal							1.890
	C23CW0	CHILDID	C23CW1 – C23CW40	JK2	WR[1]	C23CSTR C23CPSU	
	C23PW0	PARENTID	C23PW1 – C23PW40	JK2	WR	C23PSTR C23PPSU	
Kindergarten/fall-first grade longitudinal							not available[2]
	C123CW0	CHILDID	C123CW1 – C123CW40	JK2	WR	C123CSTR C123CPSU	
	C123PW0	PARENTID	C123PW1 – C123PW40	JK2	WR	C123PSTR C123PPSU	
Spring-kindergarten/spring-first grade longitudinal							1.759
	C24CW0	CHILDID	C24CW1 – C24CW90	JK2	WR	C24CSTR C24CPSU	
	C24PW0	PARENTID	C24PW1 – C24PW90	JK2	WR	C24PSTR C24PPSU	
Kindergarten/spring-first grade longitudinal							not available[2]
	C124CW0	CHILDID	C124CW1 – C124CW90	JK2	WR	C124CSTR C124CPSU	
	C124PW0	PARENTID	C124PW1 – C124PW90	JK2	WR	C124PSTR C124PPSU	
Kindergarten/first grade longitudinal							1.562
	C1_4CW0	CHILDID	C1_4CW1 – C1_4CW40	JK2	WR	C1_4CSTR C1_4CPSU	
	C1_4PW0	PARENTID	C1_4PW1 – C1_4PW40	JK2	WR	C1_4PSTR C1_4PPSU	
Combined longitudinal	Y2COMW0	CHILDID	Y2COMW1 –Y2COMW90	JK2	WR	Y2COMSTR Y2COMPSU	not available[2]

[1]WR = with replacement

[2]Design effects were not calculated using these weights.

Table 9-5.—ECLS-K spring-kindergarten/spring-first grade panel: standard errors and design effects for the full sample

Survey item	Variable name	Number of cases	Estimate	Design SE	SRS SE	DEFT	DEFF
Child characteristics (percent)							
Lived in single parent family	P4HFAMIL	14,936	23.59	0.610	0.348	1.755	3.079
Lived in two-parent family	P4HFAMIL	14,936	74.17	0.666	0.358	1.860	3.459
Mom worked 35 hours+/week	P4HMEMP	10,557	66.43	0.627	0.459	1.365	1.862
Primary case is center-based	P4PRIMNW	6,150	35.73	1.437	0.611	2.352	5.530
Primary case is home-based	P4PRIMNW	6,150	64.27	1.437	0.611	2.352	5.530
Parents had high school or less	WKPARED	14,938	36.95	0.940	0.395	2.381	5.668
Parents attended PTA	P4ATTENP	14,922	41.31	0.909	0.403	2.256	5.090
Did homework 3-4 times per week	P4HMWORK	14,924	39.78	0.950	0.401	2.372	5.627
Parents helped with homework 3-4 times/week	P4HLPHWK	14,441	37.67	0.765	0.403	1.897	3.599
Practiced reading, writing, numbers daily	P4RDWRNM	14,926	52.11	0.767	0.409	1.876	3.521
Visited library	P4LIBRAR	14,919	45.66	0.778	0.408	1.908	3.639
Used computer 1-2 times per week	P4COMPWK	10,083	45.31	0.689	0.496	1.390	1.932
Have family rule for TV	P4TVRULE	14,800	90.40	0.359	0.242	1.484	2.201
HH received foodstamp in last 12 months	P4FSTAMP	14,874	15.23	0.652	0.294	2.215	4.906
Child characteristics (mean)							
Age of child in months	R4_AGE	16,325	86.91	0.068	0.035	1.940	3.764
Child's household size	P4HTOTAL	14,936	4.56	0.024	0.012	2.072	4.293
Number of children <18 in child's HH	P4LESS18	14,936	2.55	0.021	0.010	2.205	4.861
Number of siblings in HH	P4NUMSIB	14,936	1.51	0.019	0.010	1.971	3.886
Number of hours watched TV after dinner	P4TVAFDH	14,779	0.77	0.013	0.007	1.983	3.932
Median						1.971	3.886
Mean						1.981	4.020
Standard deviation						0.320	1.219
Coefficient of variation						0.162	0.303
Minimum						1.365	1.862
Maximum						2.381	5.668

Table 9-6.—ECLS-K all-four round panel: standard errors and design effects for the full sample

Survey item	Variable name	Number of cases	Estimate	Design SE	SRS SE	DEFT	DEFF
Child characteristics (percent)							
Lived in single parent family	P4HFAMIL	4,011	22.66	0.851	0.661	1.287	1.656
Lived in two-parent family	P4HFAMIL	4,011	75.21	0.859	0.682	1.259	1.586
Mom worked 35 hours+/week	P4HMEMP	2,830	65.49	1.257	0.893	1.407	1.979
Primary case is center-based	P4PRIMNW	1,640	36.89	1.611	1.192	1.352	1.828
Primary case is home-based	P4PRIMNW	1,640	63.11	1.611	1.192	1.352	1.828
Parents had high school or less	WKPARED	4,012	36.25	1.095	0.759	1.443	2.082
Parents attended PTA	P4ATTENP	4,008	42.47	1.753	0.781	2.245	5.042
Did homework 3-4 times per week	P4HMWORK	4,008	38.78	1.299	0.770	1.688	2.850
Parents helped with homework 3-4 times/week	P4HLPHWK	3,889	36.55	1.102	0.772	1.427	2.037
Practiced reading, writing, numbers daily	P4RDWRNM	4,009	52.57	1.153	0.789	1.462	2.136
Visited library	P4LIBRAR	4,009	45.18	1.375	0.786	1.750	3.061
Used computer 1-2 times per week	P4COMPWK	2,754	46.15	1.251	0.950	1.317	1.734
Have family rule for TV	P4TVRULE	3,982	91.23	0.492	0.448	1.098	1.205
HH received foodstamp in last 12 months	P4FSTAMP	3,997	15.89	1.105	0.579	1.910	3.648
Child characteristics (mean)							
Age of child in months	R4_AGE	4,518	86.90	0.134	0.067	2.003	4.011
Child's household size	P4HTOTAL	4,011	4.63	0.039	0.022	1.762	3.106
Number of children <18 in child's HH	P4LESS18	4,011	2.60	0.031	0.018	1.691	2.861
Number of siblings in HH	P4NUMSIB	4,011	1.55	0.029	0.018	1.648	2.715
Number of hours watched TV after dinner	P4TVAFDH	3,979	0.78	0.020	0.013	1.581	2.500
Median						1.462	2.136
Mean						1.562	2.519
Standard deviation						0.288	0.959
Coefficient of variation						0.184	0.381
Minimum						1.098	1.205
Maximum						2.245	5.042

Differences in scores from the assessment data and in social rating scores as reported by the parents and teachers (spring-kindergarten/spring-first grade only) are presented in table 9-8. For the spring-kindergarten/fall-first grade panel, the differences in scores, their standard errors, and design effects were calculated using C23CW0 and the corresponding replicate weights. For the spring-kindergarten/spring-first grade panel, C24CW0 and the corresponding replicate weights were used for computing the differences in assessment scores and social rating scores reported by teachers, and their standard errors and design effects. C24PW0 and the corresponding replicate weights were used for the differences in social rating scores reported by parents, and their standard errors and design effects. The median design effect is 3.9 for the panel of spring-kindergarten/spring-first grade students and 2.1 for the panel of children in all four rounds of data collection (base year and first grade year).

Table 9-7 presents the median design effects for subgroups based on school type, child's gender and race-ethnicity, geographic region, level of urbanicity and the socioeconomic scale (SES quintiles) of the parents. For the spring-kindergarten/spring-first grade panel, the median design effects vary from 1.2 (Pacific Islanders) to 5.2 (children in the Midwest). For the panel involving all four rounds, the range of variability of the median design effects is very different from that for the spring-kindergarten/spring-first grade panel. The all-four-round panel has a much reduced sample size as it includes the fall-first grade subsample from the full base year sample. The median design effects range from 1.2 for children of other race-ethnicity to a high 8.6 for Pacific Islanders. The group of American Indian children also has a high median design effect of 6.4. The samples of Pacific Islanders and American Indians are highly clustered resulting in the higher design effects.

Table 9-9 presents the median design effects for the difference in scores for the same subgroups. The median design effects for the difference in scores between spring-kindergarten and fall-first grade are based only on three assessment scores; no social rating scores were collected in fall-first grade. For the panel spring-kindergarten/fall-first grade, the variation in the median design effects is large, especially for the group of American Indians, and unstable for certain race-ethnicity subgroups. For the group of American Indians, the smaller sample size, the highly clustered characteristic of the sample, and the large variation in the differences in scores, all contribute to the high design effects. Design effects of less than 1 are mainly due to the variability in the variance of estimates and are not viewed as having substantive implications.

Table 9-7.—ECLS-K panel: median design effects for subgroups

Subgroups	Spring-kindergarten/ spring-first grade[1] DEFT	DEFF	All four round panel[2] DEFT	DEFF
All students	1.971	3.886	1.462	2.136
Type of school				
Public	1.891	3.574	1.436	2.062
Private	1.931	3.729	1.697	2.879
Catholic private	1.710	2.923	1.348	1.816
Other private	1.746	3.048	1.606	2.580
Gender				
Male	1.684	2.837	1.358	1.845
Female	1.566	2.451	1.451	2.106
Race-ethnicity				
White	1.894	3.588	1.565	2.450
Black	1.471	2.163	1.274	1.624
Hispanic	1.457	2.122	1.184	1.403
Asian	1.586	2.514	1.440	2.074
Pacific Islander	1.085	1.178	2.938	8.632
American Indian	1.237	1.531	2.535	6.427
Other	1.260	1.587	1.092	1.193
Region				
Northeast	1.679	2.819	1.942	3.772
Midwest	2.270	5.151	1.762	3.106
South	2.033	4.135	1.665	2.772
West	1.645	2.707	1.514	2.291
Urbanicity				
Central city	1.866	3.482	1.462	2.138
Urban fringe and large town	1.924	3.703	1.622	2.632
Small town and rural area	2.056	4.226	1.798	3.233
SES quintiles				
First	1.444	2.084	1.218	1.484
Second	1.396	1.948	1.210	1.464
Third	1.462	2.136	1.248	1.558
Fourth	1.448	2.096	1.228	1.508
Fifth	1.492	2.226	1.269	1.611

[1]Each median is based on 19 items.

[2]Each median is based on 19 items.

Table 9-8.—ECLS-K spring-kindergarten/fall-first grade and spring-kindergarten/spring-first grade panels: standard errors and design effects of differences in scores for the full sample

Survey item	Variable name	Number of cases	Estimate	Design SE	SRS SE	DEFT	DEFF
Difference between spring-kindergarten and fall-first grade							
Reading score	C3RRSCAL-C2RRSCAL	4,915	5.31	0.173	0.079	2.181	4.756
Math score	C3RMSCAL-C2RMSCAL	5,126	5.40	0.106	0.074	1.434	2.057
General knowledge score	C3RGSCAL-C2RGSCAL	4,893	3.17	0.115	0.055	2.082	4.333
Median						2.082	4.333
Mean						1.899	3.715
Standard deviation						0.406	1.452
Coefficient of variation						0.214	0.391
Minimum						1.434	2.057
Maximum						2.181	4.756
Difference between spring-kindergarten and spring-first grade							
Reading score	C4RRSCAL-C2RRSCAL	15,539	22.71	0.148	0.073	2.021	4.085
Math score	C4RMSCAL-C2RMSCAL	16,126	15.74	0.113	0.049	2.302	5.300
General knowledge score	C4RGSCAL-C2RGSCAL	15,502	7.60	0.067	0.035	1.914	3.663
Approaches to learning-Parent	P4LEARN-P2LEARN	14,601	-0.03	0.006	0.004	1.534	2.352
Self-control-Parent	P4CONTRO-P2CONTRO	14,602	0.07	0.007	0.004	1.757	3.086
Social interaction-Parent	P4SOCIAL-P2SOCIAL	14,620	-0.04	0.007	0.005	1.504	2.261
Withdrawn-Parent	P4SADLON-P2SADLON	14,591	-0.01	0.005	0.004	1.335	1.783
Impulsive/overactive-Parent	P4IMPULS-P2IMPULS	14,436	-0.07	0.009	0.005	1.710	2.925
Approaches to learning-Teacher	T4LEARN-T2LEARN	14,287	-0.09	0.010	0.006	1.684	2.836
Self-control-Teacher	T4CONTRO-T2CONTRO	14,085	-0.01	0.010	0.005	1.834	3.363
Interpersonal-Teacher	T4INTERP-T2INTERP	13,994	-0.02	0.011	0.006	1.902	3.616
Externalizing problems-Teacher	T4EXTERN-T2EXTERN	14,138	0.00	0.008	0.005	1.573	2.474
Internalizing problems-Teacher	T4INTERN-T2INTERN	13,996	0.03	0.009	0.005	1.744	3.040
Median						1.744	3.040
Mean						1.755	3.137
Standard deviation						0.250	0.910
Coefficient of variation						0.143	0.290
Minimum						1.335	1.783
Maximum						2.302	5.300

Table 9-9.—ECLS-K spring-kindergarten/fall-first grade and spring-kindergarten/spring-first grade panels: median design effects of differences in scores for subgroups

Subgroups	Spring-kindergarten/ fall-first grade[1]		Spring-kindergarten/ spring-first grade[2]	
	DEFT	DEFF	DEFT	DEFF
All students	2.082	4.333	1.744	3.040
Type of school				
Public	1.967	3.870	1.607	2.581
Private	1.515	2.294	1.833	3.361
Catholic private	1.538	2.364	2.037	4.150
Other private	1.562	2.440	1.714	2.939
Gender				
Male	1.656	2.743	1.553	2.412
Female	1.740	3.027	1.473	2.171
Race-ethnicity				
White	1.830	3.349	1.591	2.531
Black	1.263	1.595	1.419	2.013
Hispanic	1.487	2.212	1.416	2.006
Asian	1.579	2.493	1.387	1.924
Pacific Islander	1.082	1.171	0.920	0.847
American Indian	7.032	49.442	1.013	1.026
Other	0.964	0.929	1.290	1.663
Region				
Northeast	1.949	3.800	1.584	2.508
Midwest	2.271	5.159	1.738	3.022
South	2.102	4.418	1.576	2.483
West	1.444	2.084	1.723	2.968
Urbanicity				
Central city	1.545	2.387	1.574	2.478
Urban fringe and large town	1.682	2.828	1.605	2.575
Small town and rural area	2.388	5.703	1.992	3.967
SES quintiles				
First	1.425	2.030	1.266	1.602
Second	1.425	2.032	1.378	1.899
Third	1.622	2.632	1.452	2.107
Fourth	1.313	1.724	1.366	1.865
Fifth	1.368	1.871	1.387	1.924

[1]Each median is based on 3 items.

[2]Each median is based on 13 items.

For the panels that include spring-first grade children, standard errors and design effects were not computed for items from the teacher and school administrator questionnaires since there are no teacher or school weights computed for the first-grade year. Although standard errors and design effects may also be calculated for the teacher and school administrator questionnaires at the child level, they are quite large compared to those typically found for the ECLS-K data. Design effects for teacher and school items are large because the intraclass correlation is 100 percent for children in the same school and very high for children in the same class; children attending the same school have the same school data, and children in the same class have the same teacher data.

NATIONAL CENTER FOR EDUCATION STATISTICS

Early Childhood Longitudinal Study
Kindergarten Class of 1998-99

USER'S MANUAL FOR THE ECLS-K THIRD GRADE PUBLIC-USE DATA FILE AND ELECTRONIC CODE BOOK

NCES 2004-001

Prepared by

Westat
Rockville, Maryland
Karen Tourangeau
Mike Brick
Thanh Le
Siu Wan
Margaret Weant
Christine Nord
Nancy Vaden-Kiernan
Mary Hagedorn
Elizabeth Bissett
Richard Dulaney
Jean Fowler

Educational Testing Service
Princeton, New Jersey
Judith Pollack
Donald Rock
Michael J. Weiss

University of Toledo
Toledo, Ohio
Sally Atkins-Burnett

The National Center for Education Statistics
U.S. Department of Education
Washington, DC
Elvira Germino Hausken
Jerry West

Education Statistics Services Institute
Washington, DC
Amy Rathbun
Jill Walston

U.S. Department of Education
Institute of Education Sciences

NCES 2004-001

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)

Chapter Page

List of Tables

Table

TABLE OF CONTENTS (continued)

List of Tables (continued)

TABLE OF CONTENTS (continued)

List of Tables (continued)

TABLE OF CONTENTS (continued)

List of Tables (continued)

TABLE OF CONTENTS (continued)

List of Tables (continued)

TABLE OF CONTENTS (continued)

List of Tables (continued)

TABLE OF CONTENTS (continued)

List of Tables (continued)

TABLE OF CONTENTS (continued)

List of Tables (continued)

TABLE OF CONTENTS (continued)

List of Exhibits

TABLE OF CONTENTS (continued)

List of Exhibits (continued)

GETTING STARTED

This document highlights key information you will need to work with the Early Childhood Longitudinal Study, Kindergarten Class of 1998–99 (ECLS-K) data and points you to the appropriate sections of the User's Manual so that you can get started quickly. To read more about any particular topic, go to the indicated section of the User's Manual. In this document, major differences between the third grade data collection and previous rounds are summarized; cautions and caveats about using the data are provided; and basic information about using the Electronic Code Book (ECB) is summarized.

You are working with the public-use file of the third grade data collection. In preparing the public-use file, 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 ensure confidentiality. The modifications that are implemented do not affect the overall data quality and most researchers should be able to find all that they need 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 7-13 in chapter 7 contains a list of the variables that have been modified. Section 7.8 contains additional information about the "masking" process.

Major Differences in the Third Grade Data Collection

- **Sample is not representative of third 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 third grade sample is not representative of third grade students, third grade teachers, or schools with third grades. Children who started their schooling in the U.S. in second or third grade are not represented in the sample. The data should not be used to make statements about third grade students, schools with third grades, or third grade teachers. See chapter 4, section 4.6 for more details on this point.

- **Children rate their perceptions of social skills and interest in school subjects**. In previous rounds of the ECLS-K, parents and teachers reported about children's social skills. For the first time in the ECLS-K, the children provided information about themselves by completing a short self-description questionnaire (SDQ). See sections 2.1 and 3.4 for additional information on the SDQ.

- **Social Rating Scale (SRS) is not collected from parents**. 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 grade, only teachers completed this scale. Sections 2.3.2 and 3.3 provide information about the SRS.

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

Cautions and Caveats

Users of previous rounds of the ECLS-K data have repeatedly asked certain questions. 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:

- **Not all sample children are in third grade**. The third grade data file includes children who were in third grade in spring 2002, and others who were either back or ahead a year or more. 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 four years (K-3) and some more (those who were repeating K in the base year). A very small number may have been in school less than four 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 and 4.4 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 third grade data collection. Section 7.6 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 5.6.1, tables 5-12 and 5-13 for more information on the response rates for movers and nonmovers.

- **Pay attention to missing data**. Users should always be sure to recode any missing data properly before conducting analyses. If 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 7.2 for a discussion of the different missing values codes and the circumstances when they are used.

- **There may be no perfect weight**. The third grade data file contains 3 sets of cross-sectional weights and 8 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 4.6.1 and 9.4.1.

- **Defining special populations**. The ECLS-K includes a number of analytic groups of interest that can be identified and studied separately. For example, the third grade file contains variables that identify children who have a disability diagnosed by a professional (P5DISABL), children receiving nonparental child care (P5CARNOW), and those who live in households with incomes below the poverty threshold (W3POVRTY). 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 (WKLANST), as well as other subgroups. These variables are not contained on the third grade cross-sectional data file, but will be available on the K–3 longitudinal data file to be released in late 2003. Users who desire to study a specific subpopulation should search the ECB using the "NARROW" feature of the ECB to list variables that might help them identify their population of interest. See section 8.3.1 for a description of this feature.

- **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 ECB contains variables that indicate the month, day, and year in which the direct assessment was administered. The ECB 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 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.

ECB Reference Guide

- **Electronic Code Book (ECB).** The ECB is designed to run under Windows 95®, Windows 98®, Windows 2000®, Windows XP®, or Windows NT® on a Pentium-class or higher PC. The PC should have a minimum of 20 megabytes (MB) of available disk space. The ECB offers the most convenient way to access the data because it enables users to search the names and labels of variables, to examine question wording and response categories for individual items, and to generate SAS, SPPS for Windows, or Stata programs for extracting selected variables (see section 8.1.2 for a description of the several features of the ECB). Section 8.2 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 child3p.dat in the CD-ROM root directory. For more information about the data file, see section 8.7 of the User's Manual. The ECB CD-ROM also contains Portable Document Format (PDF) files of the associated questionnaires and of the User's Manual.

- **Data File.** The third grade child catalog contains one record for each of 15,305 responding students in spring-third grade. Data collected from teachers and schools are stored in the child catalog. The file, named child3p.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 third grade data file contains a child identification variable (CHILDID) that uniquely identifies each record. The same ID is used across each round of the survey. Teachers on the child records are identified with ID variables T5_ID; schools are identified by the ID variables S5_ID. See sections 7.1 and 7.6 in the User's Manual for further information on these identification variables.

- **Instruments**. For the ECLS-K third grade data collection, data were collected using computer-assisted interviewing for parent interviews and child assessments. Self-administered questionnaires 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 you decide what variables to use in your analyses, you should always review the actual instruments. Seeing the specific wording of the questions and the context in which they are asked is useful in understanding the results of your analyses and can help minimize errors. Appendix A on the ECLS-K ECB CD-ROM contains, with some exceptions, the third 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, socioeconomic status. See section 7.4 and table 7-12 of the User's Manual for details on all the composites contained on the third grade public-use data file. It is strongly recommended that users give serious consideration to using the composite variables because these variables represent the best data the study has and some include sources not available on the data file.

- **Assessment Scales.** A key feature of the ELCS-K data is the assessments conducted on 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 third 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 four domains: language and literacy (reading and writing), science, social studies, 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 third grade teachers to tell 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. Children also rated their competence and popularity with peers and reported on 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, you will need to select the appropriate weights. Section 4.6.1 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 9.4.1 describes the K–3 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–3 longitudinal (panel) weights.

- **Creating a Longitudinal File**. It is possible to combine the third grade data with data from kindergarten and first grade. 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 2004.

1. INTRODUCTION

This manual provides guidance and documentation for users of the third 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 through fifth grade. 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 "third grade" is used throughout this document to refer to the data collections that took place in the 2001–02 school year, at which time most of the sampled children—but not all of them—were in third 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: Years 1998–2004

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
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.
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

The sample of children in the third 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 of children fielded in 2001–02 was not freshened with third graders who did not have a chance to be sampled in kindergarten or first grade (for example, because they were out of the country during their kindergarten and first grade year), this sample of children does not represent all third graders in 2001–02.

[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.

The vast majority of children in third grade in the 2001–02 school year are members of the cohort. However, third graders who repeated second or third 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. 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 final wave of data collection that is currently planned is scheduled for spring 2004 when most of the study children will be in the fifth grade.

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 through first grade. The ECLS-B focuses on the characteristics of children and their families that influence children's first experiences with the demands of formal school, as well as children's early health care and in- and out-of-home experiences. Together these cohorts will provide the depth and breadth of data required to more fully describe and understand children's health and 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 through fifth grade. 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 cognitive and academic growth as they move through elementary school.

1.1 Background

National policymakers and the public at large have increasingly recognized that the prosperity of the United States depends on the successful functioning of the American education system. There is also growing awareness that school reform efforts cannot focus solely on the secondary and postsecondary years but must pay attention to the elementary and preschool years as well. Increased policy interest in the early grades and the early childhood period is reflected in President Bush's No Child Left Behind Act (http://www.ed.gov.nclb) and in his Good Start, Grow Smart initiative (http://www.whitehouse.gov/infocus/earlychildhood).

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 affected by 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 these antecedent factors and enables stronger causal inferences to be made about their relationship to later academic progress.

The ECLS-K enables educational researchers and policy analysts to use a variety of perspectives on early childhood education, using techniques such as multilevel modeling to study how school and classroom factors affect the progress of individual children. The data collected will enable analysts to examine how children's status at school entry and performance in school are jointly determined by an interaction of child characteristics and school and family environments.

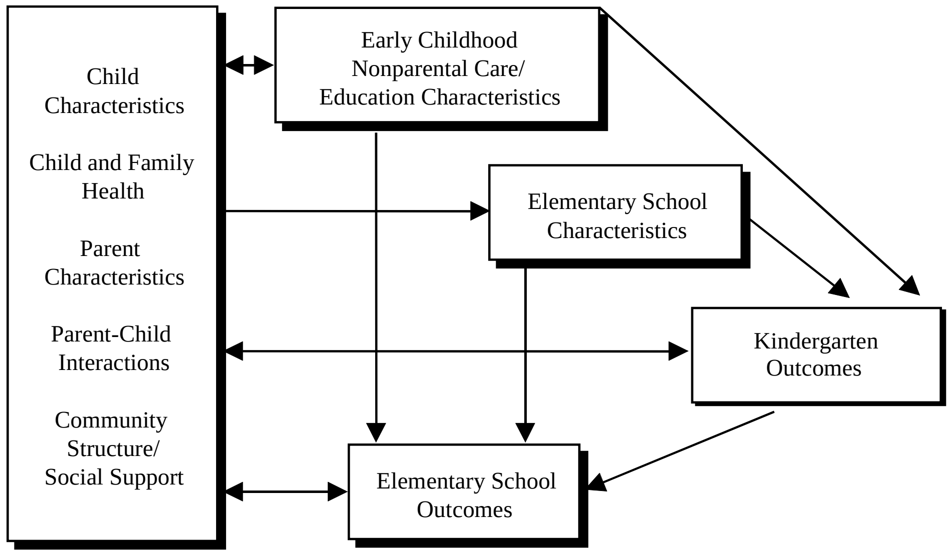
Data collected during the kindergarten year serve as baseline measures to examine how schooling shapes later individual development and achievement. The longitudinal nature of the study enables researchers to study children's cognitive, social, and emotional growth and to relate trajectories of change to variations in children's experiences in kindergarten and the early grades.

The spring-third grade data collection can be used to describe the diversity of the study children 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 and literacy practices 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 primary 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.

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., general knowledge, literacy, and quantitative) 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 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. 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. 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 a school fact sheet. 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 plan (IEP) on record. A student records abstract form is completed for each child in the study during each spring data collection.

  During the third grade data collection, school office staff were also asked to complete a school fact sheet. This form supplements the school administrator questionnaire with basic information about the school, including grade level, school type (public or private), length of school year, and attendance recordkeeping practices. This school fact sheet is only filled out once for each school in the study. Prior to the third grade data collection, the questions were part of the school administrator questionnaire.

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, and in some instances, entirely suppressed. 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 following data modifications account for the differences between the public-use and restricted-use data files:

- Outlier values are top- or bottom-coded;[6]
- 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) to reduce the risk;
- Some continuous variables are modified into categorical variables, and categories of certain categorical variables are collapsed;

[6] 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	1.97
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 outliner values are 0, 1, 4, 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	6.13
2	2,846	61.32
3	1,318	28.40
≥ 4	239	5.15

- 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.

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 the disabled, or children with specific non-English home languages or countries of birth, for example, will find that the restricted-use files contain a few more variables. 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 third grade data files.

- **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 not only can be used to analyze data collected in the third grade but it also provides 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. For more information on the third grade restricted-use data file, please see the *User’s Manual for the ECLS-K Third Grade Restricted-Use Data File and Electronic Code Book* (NCES 2003–003). 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 Files and Electronic Code Book* (NCES, forthcoming).

- **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. 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 not only can be used to analyze data collected in the first grade but also provides 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 Code Book* (NCES 2002–135) or the *User's Manual for the ECLS-K First Grade Restricted-Use Data Files and Electronic Code Book* (NCES 2002–128).

- **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 Files and Electronic Code Book* (NCES 2002–149).

- **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 Code Book: User's Manual* (NCES 2001–029) or the *ECLS-K Restricted-Use Base Year: Child File, Teacher File, and School File* (NCES 2000–097).

- 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).

- 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).

- The **student record abstract file** contains information from school records about children's school enrollment and attendance; Individualized Education Plan (IEP) and disability status; and home and school language. The student record 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).

- 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). 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 the third grade public-use data of the ECLS-K. 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 how the public-use third grade data file is structured; provides definitions of composite variables (chapter 7); describes how to install and use the Electronic Code Book (chapter 8); and describes how to use and merge the base year, first grade, and third grade files (chapter 9). The Electronic Code Book contains unweighted frequencies for all variables. Because this manual focuses on the third grade data collection, minimal information is provided about the base year or first 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 Code Book: User's Manual* (NCES 2001–029); the *ECLS-K Restricted-Use Base Year: Child File, Teacher File, and School File* (NCES 2000–097); the *User's Manual for the ECLS-K First Grade Public-Use Data Files and Electronic Code Book* (NCES 2002–135); or the *User's Manual for the ECLS-K First Grade Restricted-Use Data Files and Electronic Code Book* (NCES 2002–128). Additional information about the ECLS program can be found on the World Wide Web at http://nces.ed.gov/ecls.

2. DESCRIPTION OF DATA COLLECTION INSTRUMENTS

This chapter describes the survey instruments used during the third 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 third grade data collection. The instrumentation for the base year and first grade data collections are also shown. Similarities and differences between the third grade instruments and those used in the previous rounds are highlighted throughout this chapter.

The ECLS-K third grade data collection occurred in the spring of the 2001–02 school year. Data were collected using computer-assisted interviewing (CAI) for parent interviews and child assessments. Self-administered questionnaires were used to collect information from teachers (teacher questionnaires and special education teacher questionnaires) and school administrators or their designees (school administrator questionnaire, school fact sheet, and student records abstract). Field staff completed the school facilities check list. In addition, children completed a short self-description questionnaire on their own as part of the direct child assessments.

The third 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?”)* (Gresham and Elliott, 1990).
[2] Adapted with permission from *Self-Description Questionnaire I* (Marsh, 1990).

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

Instruments	1998–99 school year		1999–2000 school year		2001–02 school year
	Fall-kindergarten	Spring-kindergarten	Fall-first grade	Spring-first grade	Spring-third grade
Parent interview	X	X	X	X	X
Child assessments	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
Special education teacher questionnaire part A		X		X	X
Special education teacher questionnaire part B		X		X	X
Adaptive Behavior Scale		X		X	
Self-Description Questionnaire					X
School administrator questionnaire		X		X[3]	X
Student record abstract		X		X	X
School fact sheet					X
School facilities checklist		X		X	X
Salary and benefits questionnaire[4]		X			
Head Start verification[5]		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 *Early Childhood Longitudinal Study, Kindergarten Class of 1998–99 (ECLS-K) User's Manual for the ECLS-K First Grade Public-Use Data Files and Electronic Code Book* (NCES 2002–135; U.S. Department of Education, National Center for Education Statistics, 2002b) or the *Early Childhood Longitudinal Study, Kindergarten Class of 1998–99 (ECLS-K) User's Manual for the ECLS-K First Grade Restricted-Use Data Files and Electronic Code Book* (NCES 2002–128; U.S. Department of Education, National Center for Education Statistics, 2002c) 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-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.

[4] 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.

[5] 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 kindergarten, first grade, and third grade data collections, school years 1998–99, 1999–2000, and 2001–02.

2.1 Direct Child Assessments

One-on-one direct child assessments were administered using both hard-copy instruments and computer-assisted interviewing (CAI) in the spring of the 2001–02 school year. The children were assessed regardless of whether they were retained in second grade, promoted to third grade, or moved ahead to fourth. The assessments took about 90 minutes to administer. Exhibit 2-2 displays the major domains measured during the direct child assessments from all five rounds of data collection. As in the previous rounds, the third grade assessments included cognitive and physical components. In addition, the third grade assessment contained a socioemotional component completed by the children. The spring-third grade cognitive assessment scores include measures that can be compared to the base year assessments conducted in the fall of 1998 and the spring of 1999 and to the first grade assessments conducted in the fall of 1999 and the spring of 2000 to study children's gains in reading and mathematics. Chapter 3 contains a detailed description of the scores and information on their use and interpretation.

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

Direct child assessment domain	1998–99 school year		1999–2000 school year		2001–02 school year
	Fall-kindergarten	Spring-kindergarten	Fall-first grade	Spring-first grade	Spring-third grade
Language screener (Oral Language Development Scale [OLDS])[1]	X	/	/	/	
Reading (language and literacy)	X	X	X	X	X
Mathematical thinking	X	X	X	X	X
Socioemotional development					X
General knowledge (science and social studies)	X	X	X	X	
Science					X[2]
Psychomotor	X				
Height and weight	X	X	X	X	X

X Round that included the instrument.

/ OLDS was administered to language minority students who were new to the study in the spring or did not pass the cut score in the English version during the previous OLDS administration.

[1] The OLDS was given to children with a non-English language background to determine if the children understood English well enough to receive the direct child assessments in English. For further information on the OLDS, please refer to the *Early Childhood Longitudinal Study, Kindergarten Class of 1998–99 (ECLS-K) Base Year Public-Use Data Files and Electronic Code Book: User's Manual* (NCES 2001–029; U.S. Department of Education, National Center for Education Statistics, 2000) or the *Early Childhood Longitudinal Study, Kindergarten Class of 1998–99 (ECLS-K) Restricted-Use Base Year Child File, Teacher File, and School File* (NCES 2000–097; U.S. Department of Education, National Center for Education Statistics, 2001). The OLDS was not used in third grade because the vast majority of children passed it by spring-first grade.

[2] In spring-third grade, general knowledge assessment was replaced with a science assessment. Children received a science assessment 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 kindergarten, first grade, and third grade data collections, school years 1998–99, 1999–2000, and 2001–02.

The third 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 (see below). The assessor then administered the SDQ followed by the reading, math, and science assessments, and then by the physical measurements.

2.1.1 Socioemotional Development

To measure children's socioemotional development, the ECLS-K assessors administered the 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 competence and their 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 about reading grades, the difficulty of reading work, and their interest in and enjoyment of reading.
- **SDQ Mathematics** scale includes eight items about mathematics grades, the difficulty of mathematics work, and their interest in and enjoyment of mathematics.
- **SDQ School** scale includes seven items about how well they do in "all school subjects" and their enjoyment of "all school subjects."
- **SDQ Peer** scale includes six items about 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 about 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 seven items about internalizing problem behaviors such as feeling "sad a lot of the time," feeling lonely, feeling ashamed of mistakes, and worrying about school and friendships.

The items on the first four scales were adapted with permission from the *Self-Description Questionnaire I* (Marsh, 1990). The items in the two problem behavior scales were developed specifically for the ECLS-K.

Because children of this age have different levels of reading ability, 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. 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. 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 put the SDQ away after the child had completed it and 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 five 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.[3] 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 third grade direct cognitive assessments, as in previous years, included reading and mathematics domains. By third grade, however, children's knowledge of the world is more categorized into science and social studies domains. With limited time available for direct assessment, the third grade assessment included only the science domain. The third grade assessments also utilized a two-stage design. Easels were used to administer items in mathematics and science. The students also completed workbooks with open-ended mathematics questions. The reading passages and questions were in a

---

[3] For details on the two-stage assessment design, see the *Early Childhood Longitudinal Study, Kindergarten Class of 1998–99 (ECLS-K) Base Year Public-Use Data Files and Electronic Code Book: User's Manual* (NCES 2001–029; U.S. Department of Education, National Center for Education Statistics, 2000) or the *Early Childhood Longitudinal Study, Kindergarten Class of 1998–99 (ECLS-K) Restricted-Use Base Year Child File, Teacher File, and School File* (NCES 2000–097; U.S. Department of Education, National Center for Education Statistics, 2001) or the *Early Childhood Longitudinal Study, Kindergarten Class of 1998–99 (ECLS-K), Psychometric Report for Kindergarten Through First Grade* (NCES 2002–05; U.S. Department of Education, National Center for Education Statistics, 2002a).

booklet format to allow the student to refer back to the story when answering the questions. All questions were read by the assessor. In mathematics and science, all available response options were read to the child. However, the child read the response options in the reading assessment.

The ECLS-K third grade direct cognitive assessment battery was designed to assess children’s academic achievement in spring of third grade, and to provide a means of measuring growth since kindergarten entry. Child development and primary education experts consulted on the design and development of the assessment instruments. They recommended that the knowledge and skills assessed by the ECLS-K third 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 subjects of the elementary grades was made because of the central nature of these skills as antecedents of individuals’ later educational outcomes.

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 and first grade, but there were grade-appropriate changes in content and format. Test items were reviewed by elementary school curriculum specialists for appropriateness of content and difficulty, and for relevance to the test framework. In addition items were reviewed for sensitivity issues related to minority concerns. Items that passed these content, construct, and sensitivity screenings were field tested in spring 2000. The content validity of 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; Woodcock, McGrew, and Werder, 1994) that was also administered during the field test. Additional information about the development of the third grade cognitive assessment battery can be found in the *Early Childhood Longitudinal Study, Kindergarten Class of 1998–99 (ECLS-K), Psychometric Report for the Third Grade* (U.S. Department of Education, National Center for Education Statistics, forthcoming[a]).

2.1.2.1 Reading

The K-1 reading (language and literacy) assessment included questions designed to measure basic skills (print familiarity, letter recognition, beginning and ending sounds, rhyming sounds, “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.

The K-1 reading assessment contained five proficiency levels. These five levels reflect a progression of skills and knowledge. Children are thought to master a level if they pass the items within a level. If a child had mastered one of the higher proficiency levels, he or she was very likely to have 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. Three higher proficiency levels were added at the third grade level: literal inference, extrapolation, and evaluation.

Thus the third grade reading assessment contained five proficiency 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).

2.1.2.2 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 for the appropriate parts of the assessment.

The items in the K-1 mathematics assessment could also be grouped into five proficiency levels, though the math clusters were less homogeneous in content than the reading clusters. The clusters of math 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 grade mathematics assessment 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) are assessed in each of the strands. Some of the items draw 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.

Thus the items in the third grade mathematics assessment could be grouped into four proficiency levels. The clusters of third grade mathematics items included the following: (1) solving simple addition and subtraction problems; (2) solving simple multiplication and division problems and recognizing more complex number patterns; (3) demonstrating understanding of place value in integers to hundreds place; and (4) using knowledge of measurement and rate to solve word problems.

2.1.2.3 Science

The K-1 assessment battery differed from the third grade battery. The K-1 battery included a measure of general knowledge whereas the third 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 life, earth, space, and physical 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 grade battery 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-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 consisted of recording the children's height and weight in order to measure their physical growth and development. A Shorr Board (for measuring height) and a digital bathroom 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. Each height and weight was taken twice. For additional detail on the procedures used to collect height and weight, see the *Early Childhood Longitudinal Study, Kindergarten Class of 1998–99 (ECLS-K), Third Grade Methodology Report* (U.S. Department of Education, National Center for Education Statistics, forthcoming[b]).

2.2 Parent Interview

The third 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 percent) were conducted in person.

The parent interview for the spring-third grade data collection lasted on average 62 minutes and asked approximately 500 questions covering third grade school experiences, child care, parent characteristics, and child health. Exhibit 2-3 provides an overview of the topics covered in the third grade and in the previous rounds of data collection. As can be seen in the table, 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-third grade, among the questions added were ones on reading resources in the home (e.g., regular receipt of a newspaper or magazine or the availability of a dictionary or encyclopedia), the respondent's reading practices, and if there was a place set aside for the child to do homework. Topics that were dropped included the parent's report of the Social Rating Scale and attendance at religious services.

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

Parent questionnaire topic	1998–99 school year		1999–2000 school year		2001–02 school year
	Fall-kindergarten	Spring-kindergarten	Fall-first grade	Spring-first grade	Spring-third grade
Family structure	X	X	X	X	X
Demographics	X	X	X	X	X
Household roster	X	X	X	X	X
Marital status	X	X	X	X	X
Immigration status		X		X	X[1]
Primary language(s) spoken in home	X	/	/	/	/
Parent's involvement with child's school		X	X	X	X
Child care	X		X	X	X
Current arrangements with relatives	X		X	X	X
Current arrangements with nonrelatives	X		X	X	X
Current arrangements with centers	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		/	
Birth weight	X	/	/	/	
Physical functioning	X	/	/	X	X
Services for children with special needs	X	/	/	X	X
Social skills rating	X	X		X	
Home environment and cognitive stimulation	X	X	X	X	X
Frequency of literacy activities	X	X	X	X	X
Computer use		X	X	X	X
Television viewing		X	X	X	X
Homework					X
Family routines					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, and 2001–02—Continued

Parent questionnaire topic	1998–99 school year		1999–2000 school year		2001–02 school year
	Fall-kindergarten	Spring-kindergarte	Fall-first grade	Spring-first grade	Spring-third grade
Parental educational expectations for child	X		X	X	X
Neighborhood		X	X	X	X
Safety		X		X	X
Resources (e.g., community center, library)			X		
Parent education	X	/	/	X[2]	X[2]
Parent employment	X			X[2]	X[2]
Parent income		X		X	X
Welfare and other public assistance use	X	X		X	X
Parent/child interaction		X		X	X
Parent discipline		X			X
Parent health and emotional well-being		X			X
Relationships and social support	X	X			X
Marital satisfaction		X			X
Background data	X	X		X	
Mother's age at first birth	X				
Mother's age at child's birth				/	
WIC[3] 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
School involvement		X			X
Paternity		X		X	X
Child Support		X		X	X

X Round that included the topic.
/ Content areas 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] Supplemental Food Program for Women, Infants, and Children administered by the Food and Nutrition Service, U.S. Department of Agriculture.
SOURCE: U.S. Department of Education, National Center for Education Statistics, Early Childhood Longitudinal Study, Kindergarten Class of 1998–99 kindergarten, first grade, and third grade data collections, school years 1998–99, 1999–2000, and 2001–02.

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 (92 percent), the grade 3 respondent was the same as the respondent from the previous round. The child’s mother was the respondent in 87 percent of the cases and the child’s father in 9 percent.

2.3 Teacher Questionnaires

During the spring-third grade data collection, each child’s teacher received a self-administered questionnaire consisting of three distinct parts. The first section, part A, asked about the teacher’s classroom and the characteristics of the students, instructional activities and curricular focus, instructional practices in different subject areas (language arts, mathematics, science, and social studies), and student evaluation methods. The teacher was also asked about parent involvement. Only teachers of sampled children completed part A, unlike the base year when all kindergarten teachers in the school, regardless of whether they taught a sampled child, completed it. Part B asked questions on school and staff activities and the teacher’s views on teaching, the school environment, and overall school climate. Background questions about the teacher were also included in this section. Teachers were asked to complete one copy of part C for each of the sampled children in their classrooms; in this part, teachers were asked to respond to 39 questions about the child’s academic performance. The Academic Rating Scale (ARS) gathered data on each sampled child’s skills in areas of language and literacy, mathematical thinking, science, and social studies. Part C also included questions from the Social Rating Scale (SRS) that collected data on five areas of children’s social skills. The ARS and SRS are described in more detail in sections 2.3.1 and 2.3.2, respectively. The same teacher questionnaires were completed by the teacher of the sampled child regardless of the child’s grade level.

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-third 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, and 2001–02

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

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, and 2001–02—Continued

Content topic	1998–99 school year		1999–2000 school year		2001–02 school year
	Fall-kindergarten	Spring-kindergarten	Spring-first grade (First grade teacher)	Spring-first grade (Kindergarten teacher)	Spring-third grade (3rd grade teacher)
Teachers' evaluation and grading practices	X[2]	/	X[1]	X[2]	X[1]
Teachers' views on school readiness	X[2]	/	X[2]	X[2]	
Perceptions about school climate	X[2]	/	X[2]	X[2]	X[2]
Perception of personal influence on policies and classroom planning	X[2]	/	X[2]	X[2]	X[2]
Teacher demographic information	X[2]	/	X[2]	X[2]	X[2]
Teacher experience and education	X[2]	/	X[2]	X[2]	X[2]
Job satisfaction	X[2]	/	X[2]	X[2]	X[2]
Transition to school activities	X[2]	/	X[2]	X[2]	
Part C					
Indirect child cognitive evaluation by teacher (ARS)	X	X	X	X	X
Language and literacy, mathematics, general knowledge (science and social studies)	X	X	X	X	X
Social skills (SRS)	X	X	X	X	X
Additional information on sampled child		X	X	X	X
Participation in special services and programs		X	X	X	X

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, and 2001–02—Continued

Content topic	1998–99 school year		1999–2000 school year		2001–02 school year
	Fall-kindergarten	Spring-kindergarten	Spring-first grade (First grade teacher)	Spring-first grade (Kindergarten teacher)	Spring-third grade (3rd grade teacher)
Overall academic skills and physical activity levels		X	X	X	X
Reading group participation		X	X	X	X
Parental involvement		X	X	X	X

X Round that included the construct.
/ Content areas asked only of new teacher participants.
[1] Topic is in teacher questionnaire part A.
[2] Topic is in teacher questionnaire part B.
SOURCE: U.S. Department of Education, National Center for Education Statistics, Early Childhood Longitudinal Study, Kindergarten Class of 1998–99 kindergarten, first grade, and third grade data collections, school years 1998–99, 1999–2000, and 2001–02.

2.3.1 Academic Rating Scale

The kindergarten and first grade ARS contained three scales: language and literacy, mathematics, and general knowledge. There are four scales of the third grade ARS: language and literacy, mathematical thinking, science, and social studies. 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 third grade based on the teacher’s past observation and experience with the child. In the third 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. Thus, each sampled child’s primary or homeroom classroom teacher was asked to forward the questionnaire 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 both first and third grade, but the third grade item sets 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.

Similarly, in mathematics the item about demonstrating understanding of place value has a similar stem in both grades, but the third grade item sets 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.

Below is a description of the content of the third grade ARS:

- The **Language and Literacy** section of the ARS consists of eight 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.
- In the **Mathematical Thinking** section, teachers rate each child on nine items that tap the following skills: number concepts (place value, fractions, and estimation), data analysis, measurement, operations (division), geometry, application of mathematical strategies, and creating and extending patterns.
- 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, earth and space, and physical science.
- The **Social Studies** section of the ARS consists of six items. Teachers are asked to rate each child's knowledge and understanding of cultural differences, economics, geography (map skills and the interaction between humans and the environment), history, and government.

See chapter 3, section 3.2 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. These items are intended to measure approaches to learning, self-control, and interpersonal skills. The items were rated on a scale of one (never) to four (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. In third grade, examination of the responses suggested a different perception of 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 the file includes a peer relations score that combines responses on both the interpersonal items and self-control items that relate to peers, as well as these scales reported separately to facilitate comparison with earlier rounds of data collection. 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 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**. In the third grade, the teacher ratings indicated that self-control and interpersonal skills are so strongly related that they form a single scale that represents the child's skill 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 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?")* (Gresham and Elliott, 1990).

2.4 Special Education Teacher Questionnaires

In the spring-third 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 these services. The supervisor then listed special education staff working with each child (e.g., speech pathologists, reading instructors, and audiologists). Questionnaires were given to these special education teachers and related services providers. 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.

The spring-third grade special education teacher questionnaires were very similar to the ones used in previous rounds. The only differences were that questions on transition to school were not asked and a few new questions were added. Exhibit 2-5 provides a summary of the content areas addressed in the special education teacher questionnaires in spring-third 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, and 2001–02

Content topic	1998–99 school year spring-kindergarten	1999–2000 school year spring-first grade	2001–02 school year spring-third grade
Part A (Teacher Level)			
Teacher's sex	X	X	X
Teacher's age	X	X	X
Teacher's race/ethnicity	X	X	X
Teaching experience	X	X	X
Educational background	X	X	X
Special education teacher background	X	X	X
Location of service provision	X	X	X
Student load per week	X	X	X
Part B (Child Level)			
Disability category	X	X	X
IEP goals for the school year	X	X	X
Extent of services	X	X	X
Types of services provided for the year	X	X	X
Primary placement	X	X	X
Teaching practices, methods, and materials	X	X	X
Assistive technologies used by child	X	X	X
General education goals, expectations, and assessments	X	X	X
Collaboration/communication with child's general education teacher	X	X	X
Frequency of communicating with child's parents	X	X	X

See notes at end of exhibit.

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

Content topic	1998–99 school year spring-kindergarten	1999–2000 school year spring-first grade	2001–02 school year spring-third grade
Receipt of formal evaluations in the past year	X	X	X
When child first had an individualized education plan (IEP)			X
Likelihood child will have an IEP next school year			X
Percentage of IEP goals that have been met this school year			X

X Round that included the topic.
NOTE: Data were 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 kindergarten, first grade, and third grade data collections, school years 1998–99, 1999–2000, and 2001–02.

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 or children 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 2002. This self-administered questionnaire was intended to gather information about the school, student body, teachers, school policies, and administrator characteristics. The questionnaire was divided into nine sections. The first seven sections of the school administrator questionnaire 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. Exhibit 2-6 summarizes the content areas addressed in this questionnaire in spring-third grade and previous rounds.

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

Content topic	1998–99 school year	1999–2000 school year		2001–02 school year
		Spring-first grade		
	Spring-kindergarten	Returning schools	New schools	Spring-third grade
School characteristics	X	/	X	
School type	X		X	
Admission requirements	X			
School size	X	X	X	X
Average daily attendance				X
Student characteristics	X	X	X	X
Race/ethnicity of students	X	X	X	X
Children eligible for special services	X	X	X	X
Types of kindergarten programs	X			
School facilities and resources	X	/	X	X
Computer equipment	X	X	X	X
Community characteristics and school safety	X	X	X	X
Teaching and other school staff characteristics	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
School policies and programs	X	/	X	X
Assessments, testing, and retention	X	X	X	X
School-family-community connections	X	/	X	X
Programs and activities for families	X		X	X
Parent involvement and participation	X	X	X	
Programs for special populations	X	X	X	X
ESL[1] and bilingual education	X	X	X	X
Special education	X	/	X	X
Gifted and talented	X		X	X
Principal characteristics	X	X	X	X
Sex, race/ethnicity, age of principal	X	X	X	X
Experience and education	X	X	X	X
School governance and climate	X	X	X	X
Goals and objectives for teachers	X	X	X	X
School functioning and decisionmaking	X	X	X	X

/ Fewer details on the topic were collected than for new schools.
[1] English as a second language.
SOURCE: U.S. Department of Education, National Center for Education Statistics, Early Childhood Longitudinal Study, Kindergarten Class of 1998–99 kindergarten, first grade, and third grade data collections, school years 1998–99, 1999–2000, and 2001–02.

For nonresponding and late-responding schools, interviewers were trained to visit the school and encourage the school administrators to complete the questionnaire. If necessary, the interviewers were to sit down with the administrators to help them fill out the questionnaire. However, if the school administrators were still reluctant to complete the full questionnaire, the interviewers were instructed to obtain key information. This key information covered such topics as the school environment, particularly the safety of the school; school policies and practices; school programs for special populations; staffing and teacher characteristics; and principal characteristics.

2.6 School Fact Sheet

The school fact sheet collects basic information about the school including the grades taught in the school, school sector and focus, the length of the school year, and whether the school keeps student attendance records. Some of this information had been included in the school administrator questionnaire and student record abstract in previous rounds. A separate school fact sheet was developed for the spring-third grade round for ease of administration.

2.7 School Facilities Checklist

ECLS-K supervisors completed the facilities checklist during their visits to the school in the spring of third grade. The facilities checklist collects information about the (1) number of portable classrooms on school grounds, (2) presence of security measures, (3) presence of environmental factors that may affect the learning environment, and (4) overall learning climate of the school.

2.8 Student Records Abstract Form

School staff completed the student records abstract form for each sampled child in the spring of kindergarten, first grade, and third grade. This instrument was used to obtain information about the child's attendance record, presence of and details on a child's IEP, and the type of language or English proficiency screening that the school used. A copy of each child's report card was also obtained. The spring-third grade version 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 5.4 for more detail on the collection of these forms.

REFERENCES

Gresham, F.M. and Elliott, S.N. (1990). *Social Skills Rating Scale: Elementary Scale A ("How Often?").* Circle Pines, MN: American Guidance Services, Inc.

Marsh, H.W. (1990). *Self-Description Questionnaire I.* Campbelltown, New South Wales, Australia: University of Western Sydney, Macarthur.

U.S. Department of Education, National Center for Education Statistics. (2000). *Early Childhood Longitudinal Study, Kindergarten Class of 1998–99 (ECLS-K) Base Year Public-Use Data Files and Electronic Code Book: User's Manual* (NCES 2001–029). Washington, DC: U.S. Department of Education.

U.S. Department of Education, National Center for Education Statistics. (2001). *Early Childhood Longitudinal Study, Kindergarten Class of 1998–99 (ECLS-K) Restricted-Use Base Year Child File, Teacher File, and School File* (NCES 2000–097). Washington, DC: U.S. Department of Education.

U.S. Department of Education, National Center for Education Statistics (2002a). *Early Childhood Longitudinal Study–Kindergarten Class of 1998–99 (ECLS-K), Psychometric Report for Kindergarten Through First Grade* (NCES 2002–05). Washington, DC: U.S. Department of Education.

U.S. Department of Education, National Center for Education Statistics. (2002b). *Early Childhood Longitudinal Study, Kindergarten Class of 1998–99 (ECLS-K) User's Manual for the ECLS-K First Grade Public-Use Data Files and Electronic Code Book* (NCES 2002–135). Washington, DC: U.S. Department of Education.

U.S. Department of Education, National Center for Education Statistics. (2002c). *Early Childhood Longitudinal Study, Kindergarten Class of 1998–99 (ECLS-K) User's Manual for the ECLS-K First Grade Restricted-Use Data Files and Electronic Code Book* (NCES 2002–128). Washington, DC: U.S. Department of Education.

U.S. Department of Education, National Center for Education Statistics (forthcoming [a]). *Early Childhood Longitudinal Study, Kindergarten Class of 1998–99 (ECLS-K), Psychometric Report for the Third Grade*. Washington, DC: U.S. Department of Education.

U.S. Department of Education, National Center for Education Statistics (forthcoming [b]). *Early Childhood Longitudinal Study, Kindergarten Class of 1998–99 (ECLS-K), Third Grade Methodology Report.* Washington, DC: U.S. Department of Education.

Woodcock, R.W., McGrew, K.S., and Werder, J.K. (1994). *Woodcock-McGrew-Werder Mini-Battery of Achievement.* Itasca, IL: Riverside Publishing.

This page intentionally left blank.

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 third 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 third grade direct cognitive assessment contained items in reading, mathematics, and science. In each subject area, children received a 15- to 17-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 third grade direct cognitive assessment built on the framework established in the kindergarten and first grade rounds of data collection, but differed in several important respects:

- **No 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. Once they achieved a score sufficient for assessment in English, the OLDS was not administered in subsequent rounds of data collection. At kindergarten entry, about 15 percent of the ECLS-K participants were found to need screening for English proficiency. By spring of first grade, less than 6 percent of the sample was screened, and nearly two-thirds of them achieved the score required to go on to the rest of the assessment. The number of sampled children who might still lack English proficiency two years later, in third grade, was assumed to be so small that the language screening assessment was unnecessary. Therefore, the OLDS was not administered in the third grade data collection.

[1] This user's manual is applicable to the data gathered during the 2001–02 school year; information contained in this manual about data gathered during the 1998–99 school year (base year of the study) and 1999–01 school year (first grade) is provided primarily for background and comparison purposes.
[2] See chapter 2, section 2.1, 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 third grade could be expected to have advanced beyond the levels covered by the kindergarten/first grade (K-1) assessments, a new set of assessment instruments was developed for the third grade. Some of the K-1 assessment items were retained in the third grade forms to support development of a longitudinal score scale.

- **Science assessment:** The K-1 general knowledge assessment included basic natural science concepts as well as concepts in social studies. For third grade, a science assessment replaced the general knowledge assessment. There was no longitudinal scale for measuring gains in science through third grade, because the third grade science assessment was not comparable to the K-1 general knowledge assessment.

- **Assessment format:** The format of the third 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. A workbook of one to seven questions that required computations or written responses was added to the third grade mathematics assessment. 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.

- **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 grade assessment and are described in detail below.

One of the critical goals of the ECLS-K was to measure children's growth in cognitive achievement across the early elementary school years. Due to budgetary constraints, data were not collected in 2000–01, when most of the sampled children were in second grade. The absence of second grade data presented a challenge for establishing longitudinal scales to link the first grade to third grade scores. Very few children answered the most difficult items in the spring-first grade data collection correctly. Third grade field-test assessment results indicated that these same items would be too easy for the vast majority of third graders. The ability levels of first graders overlapped with those of third graders only in the tails of the distributions. Without any second grade data, it would have been difficult to place the items reliably along the difficulty scale, making it difficult to accurately estimate cognitive gains from first to third grade. In order to bridge this gap, reading and math assessments were administered to a sample of approximately 900 second graders in 43 schools. While the bridge sample was a convenience sample and was not designed to be nationally representative, efforts were made to include a diverse sample of children and schools. About 77 percent of the bridge sample children were White and 23 percent minority; 30 public schools and 13 religious or other private schools participated; and attention was given to recruiting schools spanning a wide range of socioeconomic (SES) levels. However, because the bridge sample participants did not constitute a nationally representative sample of second graders, and were not part of the ECLS-K longitudinal sample, no scores were reported for this group. The purpose of the bridge sample was to obtain data on the performance of the assessment items, rather than to track the progress of the children themselves, in order that reliable gain scores could be estimated for the first-to-third graders in the ECLS-K sample. The longitudinal scores necessary for measuring gain over time were estimated by pooling the four rounds of kindergarten/first-grade data with the data from the ECLS-K third graders and the second grade bridge sample. Details of the scoring procedures will be described in the *Early Childhood Longitudinal Study, Kindergarten Class of 1998–99* (*ECLS-K) Third Grade Methodology Report* (U.S. Department of Education, National Center for Education Statistics forthcoming).

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 3-9 show the types of scores, variable names, descriptions, and summary statistics for the direct cognitive assessment. The name and description for each variable in the tables begin with a "C," indicating that it is a child variable, and a data collection round number, either 1 (fall-kindergarten), 2 (spring-kindergarten), 3 (fall-first grade), 4 (spring-first grade), or 5 (spring-third grade). Weighted means use weight C1_5SC0, the round 1-2-3-4-5 panel weight, for the four kindergarten and first grade rounds, and the round 5 cross-sectional weight, C5CW0, to represent population estimates for third grade. Kindergarten and first grade IRT scores, T-scores, and proficiency probability scores in this data base differed slightly from the corresponding scores in the previously released data files because they were re-estimated along with the bridge and third grade scores. In addition, all kindergarten and first grade score statistics presented here differ from previous estimates because the panel weight used restricted estimates to children who participated in all five rounds of data collection.

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. 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 K-1 reading routing test emphasized pre-reading skills, while the routing test in third grade contained easy and difficult decoding words, selecting the best word to complete a sentence, and a series of questions based on a reading passage. 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, scores on the third grade routing test, which contained more difficult items, are *not* comparable with the kindergarten or first grade number-right scores. The third grade routing test number-right scores should *not* be compared with the kindergarten or first grade routing test number-right scores.

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 test.

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

Variable name	Description	Range of values	Weighted mean	Standard deviation
C1R2RNOR	C1 RC2 Reading Routing #Right - K-1 Assmt	0 - 20	5.9	4.0
C2R2RNOR	C2 RC2 Reading Routing #Right - K-1 Assmt	0 - 20	10.0	4.1
C3R2RNOR	C3 RC2 Reading Routing #Right - K-1 Assmt	0 - 20	11.7	4.2
C4R2RNOR	C4 RC2 Reading Routing #Right - K-1 Assmt	0 - 20	16.3	3.7
C1R2MNOR	C1 RC2 Math Routing #Right - K-1 Assmt	0 - 16	4.5	3.0
C2R2MNOR	C2 RC2 Math Routing #Right - K-1 Assmt	0 - 16	7.3	3.4
C3R2MNOR	C3 RC2 Math Routing #Right - K-1 Assmt	0 - 16	8.9	3.4
C4R2MNOR	C4 RC2 Math Routing #Right - K-1 Assmt	0 - 16	11.8	3.0

NOTE: Table estimates based on C1_5SC0 weight. Table estimates may differ from those reported in earlier user's manuals and the *ECLS-K Psychometric Report for Kindergarten Through First Grade* (NCES 2002–05) because of sample attrition. See chapter 7, section 7.3 for variable naming conventions.
SOURCE: U.S. Department of Education, National Center for Education Statistics, Early Childhood Longitudinal Study, Kindergarten Class of 1998–99 third grade data collection, school year 2001–02.

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

Variable name	Description	Range of values	Weighted mean	Standard deviation
C5R2RNR3	C5 RC2 Reading Routing #Right - Gr3 Assmt	0 - 15	9.9	2.8
C5R2MNR3	C5 RC2 Math Routing #Right - Gr3 Assmt	0 - 17	8.8	4.4
C5SROUNR	C5 Science Routing Test - Number Right	0 - 15	8.2	3.4

NOTE: Table estimates based on C5CW0 weight. See chapter 7, section 7.3 for variable naming conventions.
SOURCE: U.S. Department of Education, National Center for Education Statistics, Early Childhood Longitudinal Study, Kindergarten Class of 1998–99 third grade data collection, school year 2001–02.

3.1.2 Item Response Theory (IRT) 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 bridge sample and third grade data to stabilize the longitudinal estimates. In addition, the maximum values of the reading and mathematics scale scores have been extended to include the more difficult items administered in the third grade assessments. As a result, the re-estimated K-1 IRT scores in this database differ from the IRT scores in the kindergarten and first grade files previously released.

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 154 questions in all of the first- and second-stage reading forms, the 123 questions in all of the mathematics forms, and the 62 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 IRT scale scores at fall-kindergarten from the IRT scale scores at spring-first grade, spring-first grade from spring-third grade, and so forth. (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.) Gain scores are not available for science because the science assessment is new in third grade and is not comparable to the general knowledge assessment administered in the earlier rounds.

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

Table 3-3. Direct cognitive assessment: Item Response Theory scale scores: School year 2001–02

Variable name	Description	Range of values	Weighted mean	Standard deviation
C1R2RSCL	C1 RC2 Reading IRT Scale Score	0 – 154	27.4	10.3
C2R2RSCL	C2 RC2 Reading IRT Scale Score	0 – 154	38.6	13.4
C3R2RSCL	C3 RC2 Reading IRT Scale Score	0 – 154	44.8	16.4
C4R2RSCL	C4 RC2 Reading IRT Scale Score	0 – 154	66.9	20.9
C5R2RSCL	C5 RC2 Reading IRT Scale Score	0 – 154	106.1	20.7
C1R2MSCL	C1 RC2 Math IRT Scale Score	0 – 123	21.2	8.9
C2R2MSCL	C2 RC2 Math IRT Scale Score	0 – 123	31.1	11.6
C3R2MSCL	C3 RC2 Math IRT Scale Score	0 – 123	38.0	13.3
C4R2MSCL	C4 RC2 Math IRT Scale Score	0 – 123	54.5	16.2
C5R2MSCL	C5 RC2 Math IRT Scale Score	0 – 123	83.3	18.3
C5SSCALE	C5 Science IRT Scale Score	0 – 62	33.5	10.0

NOTE: Table estimates for C1-C4 variables are based on C1_5SC0 weight; estimates for C5 variables are based on C5CW0 weight. Table estimates may differ from those reported in earlier user's manuals and the *ECLS-K Psychometric Report for Kindergarten Through First Grade* (NCES 2002-05) because of re-estimation of scores on a longitudinal scale that includes third grade, and because of sample attrition. See chapter 7, section 7.3 for variable naming conventions.
SOURCE: U.S. Department of Education, National Center for Education Statistics, Early Childhood Longitudinal Study, Kindergarten Class of 1998–99 third grade data collection, school year 2001–02.

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 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 (C1R2RTSC) 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 third grade (C5R2RTSC) 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.

See table 3-4 for variable names, descriptions, and ranges for the standardized T-scores. Weighted means and standard deviations for the kindergarten and first grade scores in this table deviate slightly from the mean 50.0, standard deviation 10.0 metric because of sample attrition.

Table 3-4. Direct cognitive assessment: standardized scores: School year 2001–02

Variable name	Description	Range of values	Weighted mean	Standard deviation
C1R2RTSC	C1 RC2 Reading T-Score	0 - 96	50.6	10.1
C2R2RTSC	C2 RC2 Reading T-Score	0 - 96	50.6	9.9
C3R2RTSC	C3 RC2 Reading T-Score	0 - 96	50.4	9.8
C4R2RTSC	C4 RC2 Reading T-Score	0 - 96	50.2	9.9
C5R2RTSC	C5 RC2 Reading T-Score	0 - 96	50.0	10.0
C1R2MTSC	C1 RC2 Math T-Score	0 - 96	50.2	10.1
C2R2MTSC	C2 RC2 Math T-Score	0 - 96	50.3	10.1
C3R2MTSC	C3 RC2 Math T-Score	0 - 96	50.4	9.8
C4R2MTSC	C4 RC2 Math T-Score	0 - 96	50.5	9.8
C5R2MTSC	C5 RC2 Math T-Score	0 - 96	50.0	10.0
C5STSCOR	C5 Science T-Score	0 - 96	50.0	10.0

NOTE: Table estimates for C1-C4 variables are based on C1_5SC0 weight; estimates for C5 variables are based on C5CW0 weight. Table estimates may differ from those reported in earlier user's manuals and the *Early Childhood Longitudinal Study, Kindergarten Class of 1998–99* (*ECLS-K) Psychometric Report for Kindergarten Through First Grade* (NCES 2002-05; U.S. Department of Education, National Center for Education Statistics, 2002) because of re-estimation of scores on a longitudinal scale that includes third grade, and because of sample attrition. See chapter 7, section 7.3 for variable naming conventions.
SOURCE: U.S. Department of Education, National Center for Education Statistics, Early Childhood Longitudinal Study, Kindergarten Class of 1998–99 third grade data collection, school year 2001–02.

3.1.3 Item Cluster Scores

Several item cluster scores were 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.

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 grade reading forms because nearly all children had mastered them by the end of first grade.

A set of four relatively difficult decoding items was reported for the third grade assessment. These were words that were unlikely to be in most children's everyday vocabulary but could be sounded out phonetically. 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; however, the grade three decoding score was an entirely new task, so comparisons with scores in the earlier rounds are not meaningful.

See table 3-5 for variable names, descriptions, ranges, weighted means, and standard deviations for the reading cluster scores: print familiarity and decoding score.

Table 3-5. Direct cognitive assessment: Reading cluster scores: School year 2001–02

Variable name	Description	Range of values	Weighted mean	Standard deviation
C1R2RPRN	C1 RC2 Print Familiarity	0 – 3	1.8	1.1
C2R2RPRN	C2 RC2 Print Familiarity	0 – 3	2.4	0.9
C3R2RPRN	C3 RC2 Print Familiarity	0 – 3	2.6	0.8
C4R2RPRN	C4 RC2 Print Familiarity	0 – 3	2.8	0.6
C5R2RDEC	C5 RC2 Decoding Score Gr 3	0 – 4	1.1	1.2

NOTE: Table estimates for C1-C4 variables are based on C1_5SC0 weight; estimates for C5 variables are based on C5CW0 weight. Table estimates may differ from those reported on earlier user files because of sample attrition. See chapter 7, section 7.3 for variable naming conventions.
SOURCE: U.S. Department of Education, National Center for Education Statistics, Early Childhood Longitudinal Study, Kindergarten Class of 1998–99 third grade data collection, school year 2001–02.

3.1.3.2 Science

The 15 routing test items of the third grade science assessment tapped a range of basic concepts, with five questions each in life science, physical science, and earth science:

- Life Science: a sample of concepts related to anatomy/health, animal characteristics/ behavior, and 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.

Number-right scores for these item clusters are reported. The items were not selected to have comparable levels of difficulty within each set. For example, the mean of 3.0 for the life science cluster compared with 2.6 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 3-6 for variable names, descriptions, ranges, weighted means, and standard deviations for the science cluster scores.

Table 3-6. Direct cognitive assessment: Science cluster scores: School year 2001–02

Variable name	Description	Range of values	Weighted mean	Standard deviation
C5LIFESC	C5 Life Science Gr3	0 – 5	3.0	1.4
C5PHYSSC	C5 Physical Science Gr3	0 – 5	2.7	1.4
C5EARTSC	C5 Earth Science Gr3	0 – 5	2.6	1.3

NOTE: Table estimates based on C5CW0 weight. See chapter 7, section 7.3 for variable naming conventions.
SOURCE: U.S. Department of Education, National Center for Education Statistics, Early Childhood Longitudinal Study, Kindergarten Class of 1998–99 third grade data collection, school year 2001–02.

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 several points along the score scale 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 8 reading and 7 mathematics proficiency levels (exhibits 3-1 and 3-2) were identified in the reading and mathematics assessments for kindergarten through third grade. No proficiency scores were computed for the science assessment because the questions did not follow a hierarchical pattern.

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 third 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 five rounds of data collection, less than 7 percent of reading response patterns, and less than 5 percent of math 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.

Exhibit 3-1. Reading proficiency levels, kindergarten through third grade: School years 1998–99, 1999–2000, and 2001–02

Reading proficiency level	Description
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

SOURCE: U.S. Department of Education, National Center for Education Statistics, Early Childhood Longitudinal Study, Kindergarten Class of 1998–99 kindergarten, first, and third grade data collections, school years 1998–99, 1999–2000, and 2001–02.

Exhibit 3-2. Mathematics proficiency levels, kindergarten through third grade: School years 1998–99, 1999–2000, and 2001–02

Mathematics proficiency level	Description
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

SOURCE: U.S. Department of Education, National Center for Education Statistics, Early Childhood Longitudinal Study, Kindergarten Class of 1998–99 kindergarten, first, and third grade data collections, school years 1998–99, 1999–2000, and 2001–02.

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.

3.1.4.1 Highest Proficiency Level Mastered

Mastery of a proficiency level was defined as answering correctly at least 3 of the 4 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. 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 grade, 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. 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 and used to assign an imputed right or wrong answer. These imputed responses were then aggregated in the same manner as actual responses to determine mastery at each of the missing levels. More than 80 percent of the “highest level” scores in both reading and 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, by imputation, or by a combination of 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 3-7 for variable names, descriptions, and weighted percentages for the highest proficiency level-mastered scores.

Table 3-7. Third grade direct cognitive assessment: highest proficiency level mastered, in percent: School year 2001–02

Variable name	Description	Below Level 1	Level 1	Level 2	Level 3	Level 4	Level 5	Level 6	Level 7	Level 8
C1R2RPF	C1 RC2 Reading Highest Prof Lvl Mastered	35	36	14	13	1	1	0	0	0
C2R2RPF	C2 RC2 Reading Highest Prof Lvl Mastered	7	20	23	37	9	3	1	0	0
C3R2RPF	C3 RC2 Reading Highest Prof Lvl Mastered	4	12	19	40	15	6	2	1	0
C4R2RPF	C4 RC2 Reading Highest Prof Lvl Mastered	1	3	5	14	34	30	11	3	1
C5R2RPF	C5 RC2 Reading Highest Prof Lvl Mastered	0	0	0	2	5	21	26	25	21
C1R2MPF	C1 RC2 Math Highest Prof Lvl Mastered	8	36	36	17	3	0	0	0	†
C2R2MPF	C2 RC2 Math Highest Prof Lvl Mastered	2	14	30	37	15	2	0	0	†
C3R2MPF	C3 RC2 Math Highest Prof Lvl Mastered	2	7	20	39	26	5	0	0	†
C4R2MPF	C4 RC2 Math Highest Prof Lvl Mastered	0	2	6	21	48	20	3	0	†
C5R2MPF	C5 RC2 Math Highest Prof Lvl Mastered	0	0	0	5	20	31	28	15	†

†Not applicable. Eight proficiency levels were defined for reading, 7 for mathematics. See chapter 7, section 7.3 for variable naming conventions.
NOTE: Table estimates for C1-C4 variables based on C1_5SC0 weight; estimates for C5 variables based on C5CW0 weight. Detail may not sum to totals because of rounding.
SOURCE: U.S. Department of Education, National Center for Education Statistics, Early Childhood Longitudinal Study, Kindergarten Class of 1998–99 third grade data collection, school year 2001–02.

3.1.4.2 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 level 1 probability at time 1 from the level 1 probability at time 2 indicates whether a student advanced in mastery of the particular set of level 1 skills during this time interval. Thus, students' school experiences can be related to improvements in specific skills.

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

Table 3-8. Third grade direct cognitive assessment: proficiency probability scores—reading: School year 2001–02

Variable name	Description	Range of values	Weighted mean	Standard deviation
C1R2RPB1	C1 RC2 Prob1 - Letter Recognition	0 - 1	0.70	0.36
C1R2RPB2	C1 RC2 Prob2 - Beginning Sounds	0 - 1	0.31	0.33
C1R2RPB3	C1 RC2 Prob3 - Ending Sounds	0 - 1	0.17	0.27
C1R2RPB4	C1 RC2 Prob4 - Sight Words	0 - 1	0.03	0.14
C1R2RPB5	C1 RC2 Prob5 - Word in Context	0 - 1	0.01	0.10
C1R2RPB6	C1 RC2 Prob6 - Literal Inference	0 - 1	0.00	0.04
C1R2RPB7	C1 RC2 Prob7 - Extrapolation	0 - 1	0.00	0.01
C1R2RPB8	C1 RC2 Prob8 - Evaluation	0 - 1	0.00	0.01
C2R2RPB1	C2 RC2 Prob1 - Letter Recognition	0 - 1	0.94	0.17
C2R2RPB2	C2 RC2 Prob2 - Beginning Sounds	0 - 1	0.70	0.32
C2R2RPB3	C2 RC2 Prob3 - Ending Sounds	0 - 1	0.51	0.34
C2R2RPB4	C2 RC2 Prob4 - Sight Words	0 - 1	0.15	0.27
C2R2RPB5	C2 RC2 Prob5 - Word in Context	0 - 1	0.04	0.17
C2R2RPB6	C2 RC2 Prob6 - Literal Inference	0 - 1	0.01	0.08
C2R2RPB7	C2 RC2 Prob7 - Extrapolation	0 - 1	0.00	0.02
C2R2RPB8	C2 RC2 Prob8 - Evaluation	0 - 1	0.00	0.01
C3R2RPB1	C3 RC2 Prob1 - Letter Recognition	0 - 1	0.97	0.13
C3R2RPB2	C3 RC2 Prob2 - Beginning Sounds	0 - 1	0.82	0.27
C3R2RPB3	C3 RC2 Prob3 - Ending Sounds	0 - 1	0.66	0.32
C3R2RPB4	C3 RC2 Prob4 - Sight Words	0 - 1	0.27	0.34
C3R2RPB5	C3 RC2 Prob5 - Word in Context	0 - 1	0.09	0.25
C3R2RPB6	C3 RC2 Prob6 - Literal Inference	0 - 1	0.03	0.14
C3R2RPB7	C3 RC2 Prob7 - Extrapolation	0 - 1	0.01	0.06
C3R2RPB8	C3 RC2 Prob8 - Evaluation	0 - 1	0.00	0.03
C4R2RPB1	C4 RC2 Prob1 - Letter Recognition	0 - 1	1.00	0.04
C4R2RPB2	C4 RC2 Prob2 - Beginning Sounds	0 - 1	0.96	0.13
C4R2RPB3	C4 RC2 Prob3 - Ending Sounds	0 - 1	0.92	0.19
C4R2RPB4	C4 RC2 Prob4 - Sight Words	0 - 1	0.74	0.34
C4R2RPB5	C4 RC2 Prob5 - Word in Context	0 - 1	0.42	0.41
C4R2RPB6	C4 RC2 Prob6 - Literal Inference	0 - 1	0.15	0.29
C4R2RPB7	C4 RC2 Prob7 - Extrapolation	0 - 1	0.04	0.13
C4R2RPB8	C4 RC2 Prob8 - Evaluation	0 - 1	0.02	0.07
C5R2RPB1	C5 RC2 Prob1 - Letter Recognition	0 - 1	1.00	0.00
C5R2RPB2	C5 RC2 Prob2 - Beginning Sounds	0 - 1	1.00	0.00
C5R2RPB3	C5 RC2 Prob3 - Ending Sounds	0 - 1	1.00	0.01
C5R2RPB4	C5 RC2 Prob4 - Sight Words	0 - 1	0.98	0.08
C5R2RPB5	C5 RC2 Prob5 - Word in Context	0 - 1	0.93	0.20
C5R2RPB6	C5 RC2 Prob6 - Literal Inference	0 - 1	0.74	0.35
C5R2RPB7	C5 RC2 Prob7 - Extrapolation	0 - 1	0.42	0.37
C5R2RPB8	C5 RC2 Prob8 - Evaluation	0 - 1	0.26	0.27

NOTE: Table estimates for C1-C4 variables are based on C1_5SC0 weight; estimates for C5 variables are based on C5CW0 weight. Table estimates may differ from those reported in earlier user's manuals and the *ECLS-K Psychometric Report for Kindergarten Through First Grade* (NCES 2002-05) because of re-estimation of scores on a longitudinal scale that includes third grade, and because of sample attrition. See chapter 7, section 7.3 for variable naming conventions.
SOURCE: U.S. Department of Education, National Center for Education Statistics, Early Childhood Longitudinal Study, Kindergarten Class of 1998–99 third grade data collection, school year 2001–02.

Table 3-9. Third grade direct cognitive assessment: proficiency probability scores—mathematics: School year 2001–02

Variable name	Description	Range of values	Weighted mean	Standard deviation
C1R2MPB1	C1 RC2 Prob1 – Count, Number, Shape	0 - 1	0.91	0.19
C1R2MPB2	C1 RC2 Prob2 – Relative Size	0 - 1	0.54	0.35
C1R2MPB3	C1 RC2 Prob3 – Ordinality, Sequence	0 - 1	0.20	0.31
C1R2MPB4	C1 RC2 Prob4 – Add/Subtract	0 - 1	0.04	0.12
C1R2MPB5	C1 RC2 Prob5 – Multiply/Divide	0 - 1	0.00	0.04
C1R2MPB6	C1 RC2 Prob6 – Place Value	0 - 1	0.00	0.00
C1R2MPB7	C1 RC2 Prob7 – Rate & Measurement	0 - 1	0.00	0.00
C2R2MPB1	C2 RC2 Prob1 - Count, Number, Shape	0 - 1	0.99	0.06
C2R2MPB2	C2 RC2 Prob2 - Relative Size	0 - 1	0.83	0.25
C2R2MPB3	C2 RC2 Prob3 - Ordinality, Sequence	0 - 1	0.53	0.39
C2R2MPB4	C2 RC2 Prob4 - Add/Subtract	0 - 1	0.17	0.26
C2R2MPB5	C2 RC2 Prob5 - Multiply/Divide	0 - 1	0.02	0.08
C2R2MPB6	C2 RC2 Prob6 - Place Value	0 - 1	0.00	0.01
C2R2MPB7	C2 RC2 Prob7 - Rate & Measurement	0 - 1	0.00	0.00
C3R2MPB1	C3 RC2 Prob1 - Count, Number, Shape	0 - 1	0.99	0.04
C3R2MPB2	C3 RC2 Prob2 - Relative Size	0 - 1	0.92	0.18
C3R2MPB3	C3 RC2 Prob3 - Ordinality, Sequence	0 - 1	0.73	0.35
C3R2MPB4	C3 RC2 Prob4 - Add/Subtract	0 - 1	0.32	0.33
C3R2MPB5	C3 RC2 Prob5 - Multiply/Divide	0 - 1	0.05	0.14
C3R2MPB6	C3 RC2 Prob6 - Place Value	0 - 1	0.00	0.03
C3R2MPB7	C3 RC2 Prob7 - Rate & Measurement	0 - 1	0.00	0.00
C4R2MPB1	C4 RC2 Prob1 - Count, Number, Shape	0 - 1	1.00	0.01
C4R2MPB2	C4 RC2 Prob2 - Relative Size	0 - 1	0.98	0.08
C4R2MPB3	C4 RC2 Prob3 - Ordinality, Sequence	0 - 1	0.94	0.19
C4R2MPB4	C4 RC2 Prob4 - Add/Subtract	0 - 1	0.70	0.32
C4R2MPB5	C4 RC2 Prob5 - Multiply/Divide	0 - 1	0.23	0.30
C4R2MPB6	C4 RC2 Prob6 - Place Value	0 - 1	0.03	0.11
C4R2MPB7	C4 RC2 Prob7 - Rate & Measurement	0 - 1	0.00	0.02
C5R2MPB1	C5 RC2 Prob1 - Count, Number, Shape	0 - 1	1.00	0.00
C5R2MPB2	C5 RC2 Prob2 - Relative Size	0 - 1	1.00	0.00
C5R2MPB3	C5 RC2 Prob3 - Ordinality, Sequence	0 - 1	1.00	0.01
C5R2MPB4	C5 RC2 Prob4 - Add/Subtract	0 - 1	0.96	0.11
C5R2MPB5	C5 RC2 Prob5 - Multiply/Divide	0 - 1	0.75	0.32
C5R2MPB6	C5 RC2 Prob6 - Place Value	0 - 1	0.39	0.39
C5R2MPB7	C5 RC2 Prob7 - Rate & Measurement	0 - 1	0.14	0.28

NOTE: Table estimates for C1-C4 variables are based on C1_5SC0 weight; estimates for C5 variables are based on C5CW0 weight. Table estimates may differ from those reported in earlier user's manuals and the *ECLS-K Psychometric Report for Kindergarten Through First Grade* (NCES 2002-05) because of re-estimation of scores on a longitudinal scale that includes third grade, and because of sample attrition. See chapter 7, section 7.3 for variable naming conventions.
SOURCE: U.S. Department of Education, National Center for Education Statistics, Early Childhood Longitudinal Study, Kindergarten Class of 1998–99 third grade data collection, school year 2001–02.

Some examples of interpretation and use of the proficiency probability scores whose means appear in table 3-8 are the following:

- At entry to kindergarten about 70 percent (mean probability = .70) of children were proficient at letter recognition (C1R2RPB1).
- 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 (C2R2RPB4) compared to 74 percent a year later (C4R2RPB4).
- There were only small gains in letter recognition after spring-kindergarten, because most children, 94 percent, knew their letters by this time (C2R2RPB1).
- Children’s skills in making inferences based on cues directly stated in text (literal inference) increased dramatically between first and third grade, from 15 percent (C4R2RPB6) to 74 percent (C5R2RPB6).
- In spring-third grade, most children had not yet demonstrated understanding of the author’s craft or making connections between a problem in the narrative and similar life problems. Only 26 percent mastered the evaluation level (C5R2RPB8).

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.

- **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 were used as longitudinal measures of overall growth. 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 lessons from 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 third grade may be working on learning to make evaluative judgments based on reading material, which would show up as large gains in reading level 8. Less skilled readers may show their largest gains between first and third grade at levels 5 or 6, comprehension of words in context and literal inference. 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. Highest proficient 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 third-grade reading, math, and science assessments are based on 15, 17, and 15 items respectively (20, 16, and 12 items for the K-1 reading, math 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., Decoding Score Gr 3) and science (e.g., Life Science Gr 3) are based on a count of the number correct for a particular 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 five 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 set of dichotomous pass/fail scores, reported as a single highest mastery level. 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 20 percent of these scores, the item data was supplemented with IRT-based estimates to avoid complications associated with missing data that was not missing at random. The highest proficiency level mastered should be treated as an ordinal variable.

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 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 *Early Childhood Longitudinal Study, Kindergarten Class of 1998–99* (*ECLS-K) Psychometric Report for Kindergarten Through First Grade* (NCES 2002-05; U.S. Department of Education, National Center for Education Statistics, 2002) describes this analysis in 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. 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 8 reading and 7 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 .23, 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, .75, is equivalent to a population mastery rate of 75 percent. While most children were making their largest gains at level 5, a small number of children were advancing their skills in solving word problems based on rate and measurement, level 7. The mastery rate for level 7 advanced from near zero at the end of first grade to about 14 percent at the end of third grade. 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 outcome measures to be explained by background and process variables.

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 gains 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 (2002).

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. 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-first grade to spring-third 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.

3.1.7 Reliability

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. These reliabilities, ranging from .88 to .96, apply to all of the scores derived from the theta estimate, namely, the IRT scale scores, T-scores, and proficiency probabilities (see table 3-10). Alpha coefficients for the routing test number correct ranged from .75 to .86 for the third grade assessment forms (see table 3-11). The third grade reading alpha is somewhat lower than in earlier rounds, at least in part due to the third grade form having fewer items (15) than the 20 items in the K-1 routing test. Conversely, the alpha coefficient for the math routing test was slightly higher in third grade. The increase in the number of mathematics routing items, from 16 in the K-1 form to 17 in third grade, probably accounts for a small part of this difference. Split-half reliabilities were computed for the item cluster scores in reading and science (see table 3-12). These reliabilities were higher for the reading clusters (.60 to .67) than for the science scores (.46 to .59). 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.

The score indicating the highest proficiency level mastered is based on a combination of raw item scores and IRT-based imputations. For a majority of students, highest level mastered could be determined on the basis of actual item responses alone, while for others imputations were required. Thus, the reported score represents a hybrid of different methodologies, as well as a collapsing of separate measurements into one.

Standard measures of reliability are not suitable for assessing the reliability of the highest proficiency level score, for several reasons. Split-half reliabilities at each individual proficiency level could be calculated for children with complete sets of item responses, but would not apply to levels or students whose item response data was supplemented with IRT-based estimates. Similarly, the reliability of the IRT theta would be relevant only to the small percentage of proficiency level scores for which IRT-based estimates alone provided the pass/fail determinations. Furthermore, the score denoting the highest level mastered reduces the series of pass/fail level scores to a single composite, so any reliability estimates obtained for individual levels would not necessarily represent the reliability of the single reported score.

As a result, the statistics reported in table 3-13 are not traditional reliability measures. Another approach was employed, based on the idea that reliability statistics are estimates of consistency of measurement under different circumstances, such as different sets of assessment items, or in this case, different methodologies. The two scoring methods, actual item responses and imputations, were compared, and statistics on the agreement between the methods presented in table 3-13 serve as reliability estimates for the highest level mastered scores. First, highest proficiency level mastered was determined on the basis of item responses alone for all of the children (about 80 percent in each round of data collection) who could be categorized without requiring imputations. Then, for the same group of children, highest level was obtained using only imputations, and ignoring actual item responses. The percent of agreement of the scores resulting from independent application of the two methods can be considered a proxy for the reliability of the hybrid scoring procedure. The numbers in table 3-13 are the percentages of exact agreement between these two extremes, and the percentages of scores that were within one level of each other. When discrepancies were examined grade-by-grade and level-by-level, there did not appear to be a pattern of either method consistently overestimating or underestimating highest proficiency level when compared with the alternate method. The vast majority of the highest proficiency level scores in the data file were determined on the basis of item responses alone, or item responses supplemented by IRT estimates. 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 3-10 through 3-13 present the reliability statistics for all of the assessment scores.

Table 3-10. Reliability statistics of Item Response Theory (IRT)-based scores, by round of data collection and domain: School years 1998–99, 1999–2000, and 2001–02

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

† Not applicable. There was no science assessment prior to third grade.
NOTE: The IRT-based scores consist of the IRT scale scores, T-scores, and proficiency probabilities. See sections 3.1.2 and 3.1.4 for a discussion of these scores. 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.
SOURCE: U.S. Department of Education, National Center for Education Statistics, Early Childhood Longitudinal Study, Kindergarten Class of 1998–99 kindergarten, first grade, and third grade data collections, school years 1998–99, 1999–2000, and 2001–02.

Table 3-11. Reliability statistics of routing test number correct (alpha coefficient), by round of data collection and domain: School years 1998–99, 1999–2000, and 2001–02

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

† Not applicable. There was no science assessment prior to third grade.
NOTE: 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. See section 3.1 for a discussion of the routing tests and section 3.1.1 for a discussion of number-right scores.
SOURCE: U.S. Department of Education, National Center for Education Statistics, Early Childhood Longitudinal Study, Kindergarten Class of 1998–99 kindergarten, first grade, and third grade data collections, school years 1998–99, 1999–2000, and 2001–02.

Table 3-12. Split-half reliability statistics for item-cluster-based scores, by round of data collection and cluster: School years 1998–99, 1999–2000, and 2001–02

Type	Fall-kindergarten	Spring-kindergarten	Fall-first grade	Spring-first grade	Spring-third grade
Print Familiarity	.70	.68	.68	.60	†
Decoding Score	†	†	†	†	.67
Life Science	†	†	†	†	.59
Physical Science	†	†	†	†	.49
Earth Science	†	†	†	†	.46

† Not applicable. Cluster not collected.
NOTE: 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.
SOURCE: U.S. Department of Education, National Center for Education Statistics, Early Childhood Longitudinal Study, Kindergarten Class of 1998–99 kindergarten, first grade, and third grade data collections, school years 1998–99, 1999–2000, and 2001–02.

Table 3-13. Percent agreement of highest proficiency level-mastered score, by round of data collection: School years 1998–99, 1999–2000, and 2001–02.

Category	Fall-kindergarten	Spring-kindergarten	Fall-first grade	Spring-first grade	Spring-third grade
Reading					
Exact Agreement	68	57	57	59	54
Exact + Off by 1	97	95	95	95	94
Mathematics					
Exact Agreement	58	56	56	61	61
Exact + Off by 1	97	96	96	98	98

NOTE: 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.
SOURCE: U.S. Department of Education, National Center for Education Statistics, Early Childhood Longitudinal Study, Kindergarten Class of 1998–99 kindergarten, first grade, and third grade data collections, school years 1998-99, 1999-2000, and 2001-02.

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 grade (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 third-grade assessment forms. Some compromises in matching target percentages were necessary to satisfy constraints related to other issues, including linking to K-1 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 third-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; Woodcock, McGrew, and Werder, 1994) in the spring 2000 field test of third 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 3-14.

Table 3-14. Validity coefficients for reading and mathematics field test item pools: School year 2001–02

Category	Reading	Mathematics
Reliability of MBA (computed both ways)	.84 and .86	.81 and .82
Square root of reliability	.92 and .93	.90 and .91
Correlation of MBA x ECLS grade 3 field assessment item pool	.83	.84

SOURCE: U.S. Department of Education, National Center for Education Statistics, Early Childhood Longitudinal Study, Kindergarten Class of 1998–99 third grade data collection, school year 2001–02.

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 four domains: language and literacy (reading and writing), science, social studies, and mathematical thinking. Teachers rated the child's skills, knowledge, and behaviors on a scale from "Not Yet" to "Proficient" (see table 3-15). 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 third grade, the classroom teacher most knowledgeable of the child's academic achievement in the four domains may not be the primary or homeroom teacher. The primary teacher was asked to forward the rating form to the teacher most knowledgeable of the particular domain to complete the ratings. 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 3-15. Academic Rating Scale response scale: School year 2001–02

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

SOURCE: U.S. Department of Education, National Center for Education Statistics, Early Childhood Longitudinal Study, Kindergarten Class of 1998–99 third grade data collection, school year 2001–02.

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 measures only the products of children's achievement. 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 math problems, or investigate a scientific phenomenon.

Due to time constraints, the direct cognitive battery does not include a scale of children's knowledge in social studies. On the ARS teachers reported the children's knowledge and understanding of civics, geography, history, culture, and economics.

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 3-16).

Table 3-16. Person separation reliability statistics[1] for the third-grade Rasch-based score, by category: School year 2001–02

Category	Grade 3
ARS Language and Literacy	.95
ARS Mathematical Thinking	.94
ARS Science	.95
ARS Social Studies	.93

[1]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 third grade data collection, school year 2001–02.

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 third grade (T5) ARS scores are shown in table 3-17. 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 third 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 and first grade. 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 and mathematics scores, and fewer than 5 percent of science and social studies scores, failed to meet the completeness criterion.

Table 3-17. Third grade Academic Rating Scale: variable names, descriptions, ranges, weighted means, and standard deviations: School year 2001–02

Variable name	Description	Range of values	Weighted mean	Standard deviation
T5ARSLIT	T5 Literacy ARS Score	1 - 5	3.27	0.89
T5ARSMAT	T5 Math ARS Score	1 - 5	3.08	0.75
T5ARSSCI	T5 Science ARS Score	1 - 5	3.16	0.94
T5ARSSOC	T5 Social Studies ARS Score	1 - 5	3.01	0.86

NOTE: Table estimates based on C5CW0 weight. See chapter 7, section 7.3 for variable naming conventions.
SOURCE: U.S. Department of Education, National Center for Education Statistics, Early Childhood Longitudinal Study, Kindergarten Class of 1998–99 third grade data collection, school year 2001–02.

Tables 3-18 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 third grade items.

Table 3-18. Spring-third grade Academic Rating Scale language and literacy item difficulties: School year 2001–02

Item difficulty	Item number and abbreviated content
2.69	Q3. Conveys ideas clearly when speaking
2.72	Q6. Reads third grade books (fiction) independently with comprehension
2.74	Q5. Reads fluently
2.87	Q4. Uses various strategies to gain information
3.03	Q7. Reads and comprehends expository text
3.19	Q10. Makes mechanical corrections when reviewing a rough draft
3.22	Q8. Composes multi-paragraph stories/reports,
3.28	Q9. Rereads and reflects on writing, making changes to clarify or elaborate

NOTE: Items are arranged in order of difficulty. 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.
SOURCE: U.S. Department of Education, National Center for Education Statistics, Early Childhood Longitudinal Study, Kindergarten Class of 1998–99 third grade data collection, school year 2001–02.

Table 3-19. Spring-third grade Academic Rating Scale mathematical thinking item difficulties: School year 2001–02

Item difficulty	Item number and abbreviated content
2.45	Q7. Shows understanding of place value with whole numbers
2.53	Q3. Creates and extends patterns
2.71	Q5. Recognizes properties of shapes and relationships among shapes
2.73	Q9. Surveys, collects, and organizes data into simple graphs
2.78	Q4. Uses a variety of strategies to solve math problems
2.81	Q6. Uses measuring tools accurately
2.83	Q8. Makes reasonable estimates of quantities and checks answers
3.16	Q10. Models, reads, writes, and compares fractions
3.69	Q11. Divides a 3 digit number by a 1 digit number

NOTE: Items are arranged in order of difficulty. 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.
SOURCE: U.S. Department of Education, National Center for Education Statistics, Early Childhood Longitudinal Study, Kindergarten Class of 1998–99 third grade data collection, school year 2001–02.

Table 3-20. Spring-third grade Academic Rating Scale science item difficulties: School year 2001–02

Item difficulty	Item number and abbreviated content
2.77	Q5. Classifies and compares living and non-living things in different ways
2.84	Q3. Makes logical predictions when conducting scientific investigations
2.87	Q8. Demonstrates understanding of life science concepts
2.93	Q9. Demonstrates understanding of earth and space science concepts
2.97	Q6. Forms explanations and conclusions based on observation and investigation
3.01	Q7. Demonstrates understanding of physical science concepts
3.07	Q4. Communicates scientific information

NOTE: Items are arranged in order of difficulty. 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.
SOURCE: U.S. Department of Education, National Center for Education Statistics, Early Childhood Longitudinal Study, Kindergarten Class of 1998–99 third grade data collection, school year 2001–02.

Table 3-21. Spring-third grade Academic Rating Scale social studies item difficulties: School year 2001–02

Item difficulty	Item number and abbreviated content
2.56	Q7. Knows how to use maps and globes to locate and derive information
2.67	Q3. Identifies similarities and differences in group habits and living patterns
2.84	Q6. Recognizes the reciprocal influence of environment on people
2.88	Q5. Demonstrates understanding of the ways in which the past influences the present
3.19	Q4. Shows understanding of the purpose and structure of government functions
3.33	Q8. Demonstrates understanding of the U. S. economic system

NOTE: Items are arranged in order of difficulty. 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.
SOURCE: U.S. Department of Education, National Center for Education Statistics, Early Childhood Longitudinal Study, Kindergarten Class of 1998–99 third grade data collection, school year 2001–02.

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 kindergarten and first grade teacher questionnaires, 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 should not 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 3-25 provide standard errors (SE) for each of the Rasch scores for third 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-third grade Academic Rating Scale language and literacy standard errors: School year 2001–02

Score	Measure	Standard error (SE)	Score	Measure	Standard Error (SE)	Score	Measure	Standard error (SE)
8	1.00E	.45	19	2.45	.15	30	3.56	.15
9	1.32	.26	20	2.54	.15	31	3.66	.15
10	1.53	.20	21	2.64	.15	32	3.75	.15
11	1.68	.18	22	2.74	.16	33	3.85	.15
12	1.80	.16	23	2.84	.16	34	3.95	.16
13	1.90	.16	24	2.94	.16	35	4.06	.16
14	2.00	.15	25	3.05	.16	36	4.17	.17
15	2.09	.15	26	3.15	.16	37	4.30	.18
16	2.18	.15	27	3.26	.16	38	4.46	.21
17	2.27	.15	28	3.36	.16	39	4.68	.27
18	2.36	.15	29	3.46	.16	40	5.00E	.45

NOTE: The “Score” column is the sum of the raw score ratings. “Measure” is the Rasch-based score. The column labeled “Standard error (SE)” is the corresponding standard error of measurement for those scores. E=estimated extreme score.
SOURCE: U.S. Department of Education, National Center for Education Statistics, Early Childhood Longitudinal Study, Kindergarten Class of 1998–99 third grade data collection, school year 2001–02.

Table 3-23. Spring-third grade Academic Rating Scale mathematical thinking standard errors: School year 2001–02

Score	Measure	Standard error (SE)	Score	Measure	Standard error (SE)	Score	Measure	Standard error (SE)
9	1.00E	.49	22	2.46	.14	35	3.41	.14
10	1.34	.28	23	2.53	.14	36	3.49	.15
11	1.55	.21	24	2.60	.14	37	3.57	.15
12	1.69	.18	25	2.67	.14	38	3.66	.15
13	1.79	.16	26	2.74	.14	39	3.75	.16
14	1.88	.15	27	2.82	.14	40	3.85	.17
15	1.96	.14	28	2.89	.14	41	3.97	.18
16	2.04	.14	29	2.96	.14	42	4.11	.20
17	2.11	.14	30	3.03	.14	43	4.29	.24
18	2.18	.13	31	3.11	.14	44	4.58	.32
19	2.25	.13	32	3.18	.14	45	5.00E	.52
20	2.32	.13	33	3.26	.14			
21	2.39	.13	34	3.33	.14			

NOTE: The “Score” column is the sum of the raw score ratings. “Measure” is the Rasch-based score. The column labeled “Standard error (SE)” is the corresponding standard error of measurement for those scores. E=estimated extreme score.
SOURCE: U.S. Department of Education, National Center for Education Statistics, Early Childhood Longitudinal Study, Kindergarten Class of 1998–99 third grade data collection, school year 2001–02.

Table 3-24. Spring-third grade Academic Rating Scale science standard errors: School year 2001–02

Score	Measure	Standard error (SE)	Score	Measure	Standard error (SE)	Score	Measure	Standard error (SE)
7	1.00E	.40	17	2.36	.15	27	3.63	.19
8	1.29	.23	18	2.47	.15	28	3.80	.20
9	1.48	.16	19	2.58	.16	29	3.97	.19
10	1.61	.15	20	2.71	.17	30	4.12	.17
11	1.73	.15	21	2.85	.18	31	4.25	.16
12	1.83	.15	22	2.99	.17	32	4.37	.17
13	1.94	.15	23	3.12	.16	33	4.51	.18
14	2.05	.15	24	3.24	.16	34	4.71	.24
15	2.15	.15	25	3.36	.16	35	4.99E	.40
16	2.26	.15	26	3.48	.17			

NOTE: The “Score” column is the sum of the raw score ratings. “Measure” is the Rasch-based score. The column labeled “Standard error (SE)” is the corresponding standard error of measurement for those scores. E=estimated extreme score.
SOURCE: U.S. Department of Education, National Center for Education Statistics, Early Childhood Longitudinal Study, Kindergarten Class of 1998–99 third grade data collection, school year 2001–02.

Table 3-25. Spring-third grade Academic Rating Scale social studies standard errors: School year 2001–02

Score	Measure	Standard error (SE)	Score	Measure	Standard error (SE)	Score	Measure	Standard error (SE)
6	1.00E	.45	15	2.47	.17	24	3.71	.19
7	1.33	.27	16	2.59	.17	25	3.86	.19
8	1.56	.21	17	2.72	.17	26	4.02	.20
9	1.73	.19	18	2.84	.17	27	4.19	.21
10	1.87	.18	19	2.97	.18	28	4.39	.23
11	2.00	.17	20	3.11	.18	29	4.65	.28
12	2.12	.17	21	3.25	.19	30	5.00E	.46
13	2.23	.17	22	3.40	.19			
14	2.35	.17	23	3.55	.19			

NOTE: The “Score” column is the sum of the raw score ratings. “Measure” is the Rasch-based score. The column labeled “Standard error (SE)” is the corresponding standard error of measurement for those scores. E=estimated extreme score.
SOURCE: U.S. Department of Education, National Center for Education Statistics, Early Childhood Longitudinal Study, Kindergarten Class of 1998–99 third grade data collection, school year 2001–02.

Classroom teachers were asked to forward the ARS to the child’s teacher most knowledgeable about the child’s performance in each of the subject areas to complete. The majority of teachers rated more than one student on the ARS. The number of students rated by each teacher ranged from one to more than 20. 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 third-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 earlier rounds of data collection, SRS questions had been asked of both teachers and parents.) Teachers used a frequency scale (see table 3-26) 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 3-26. Social Rating Scale response scale: School year 2001–02

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 third grade data collection, school year 2001–02.

Five teacher SRS scales were developed based on responses to the scale. The scale scores on all SRS scales are 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 24 of the 26 third grade SRS items were the same as items in the K-1 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 grade scales due to a high number of maximum scores on the field assessment 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 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.

Variable names for the teacher scores, descriptions, ranges, weighted means, and standard deviations for these scales are shown in table 3-27. About 90 percent of the children whose teachers provided social ratings data were in third grade during the round 5 data collection, and about 9 percent were in first or second grade. Numbers in the table are for third graders, with scores for children who at round 5 were still in first or second grade shown in parentheses. The number of children who had advanced to fourth or fifth grade by round 5 was too small to be analyzed separately.

Table 3-27. Teacher Social Rating Scale scores: variable names, descriptions, ranges, weighted means, and standard deviations: School year 2001–02

Variable name	Description	Range of values	Weighted mean	Standard deviation
T5LEARN	T5 Approaches to Learning	1 - 4	3.1 (2.7)	0.7 (0.7)
T5CONTRO	T5 Self-Control	1 - 4	3.2 (3.0)	0.6 (0.7)
T5INTERP	T5 Interpersonal	1 - 4	3.1 (2.8)	0.6 (0.7)
T5EXTERN	T5 Externalizing Problem Behaviors	1 - 4	1.7 (1.9)	0.6 (0.7)
T5INTERN	T5 Internalizing Problem Behaviors	1 - 4	1.6 (1.8)	0.5 (0.6)
T5SCINT	T5 Combo of Self-Control & Interpersonal (Peer Relations)	1 - 4	3.2 (2.9)	0.6 (0.6)

NOTE: Table estimates based on C5CW0 weight. Numbers outside of parentheses represent children in third grade at the time of assessment. Numbers within parentheses represent first and/or second graders at the time of assessment. See chapter 7, section 7.3 for variable naming conventions.
SOURCE: U.S. Department of Education, National Center for Education Statistics, Early Childhood Longitudinal Study, Kindergarten Class of 1998–99 third grade data collection, school year 2001–02.

The split-half reliabilities for the teacher SRS scales are high (see table 3-28). Reliabilities are nearly identical for third graders in round 5 and for children who were not yet in third grade, so the table contains only reliabilities for the whole sample.

Table 3-28. Split-half reliability for the teacher Social Rating Scale scores: School year 2001–02

Variable name	Description	Split-half reliability
T5LEARN	T5 Approaches to Learning	.91
T5CONTRO	T5 Self-control	.79
T5INTERP	T5 Interpersonal	.89
T5EXTERN	T5 Externalizing Problem Behaviors	.89
T5INTERN	T5 Internalizing Problem Behaviors	.76
T5SCINT	T5 Combo of Self-Control & Interpersonal (Peer Relations)	.92

NOTE: See chapter 7, section 7.3 for variable naming conventions.
SOURCE: U.S. Department of Education, National Center for Education Statistics, Early Childhood Longitudinal Study, Kindergarten Class of 1998–99 third grade data collection, school year 2001–02.

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. The factor intercorrelations with the internalizing problem behaviors are the lowest. The absolute values of correlations among the teacher SRS factors range from 0.32 to 0.81, with nearly identical patterns for third graders and for children who were still in first or second grade.

3.4 Self-Description Questionnaire

For the first time in the ECLS-K, third 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 distractability, while "Internalizing Problems" scale included items on sadness, loneliness, and anxiety. For further description of the Self-Description Questionnaire (SDQ) see chapter 2. 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 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).

Table 3-29. Self-Description Questionnaire scale reliabilities (alpha coefficient): School year 2001–02

Variable name	Description	Number of items	Alpha coefficient
C5SDQRDC	C5 SDQ Prcvd Interest/Competence - Reading	8	.87
C5SDQMTC	C5 SDQ Prcvd Interest/Competence - Math	8	.90
C5SDQSBC	C5 SDQ Prcvd Interest/Competence - All Sbj	6	.79
C5SDQPRC	C5 SDQ Prcvd Interest/Competence - Peer Rl	6	.79
C5SDQEXT	C5 SDQ Externalizing Problems	6	.77
C5SDQINT	C5 SDQ Internalizing Problems	8	.81

NOTE: See chapter 7, section 7.3 for variable naming conventions.
SOURCE: U.S. Department of Education, National Center for Education Statistics, Early Childhood Longitudinal Study, Kindergarten Class of 1998–99 third grade data collection, school year 2001–02.

Table 3-30. Self-Description Questionnaire scale range, mean, and standard deviation (weighted): School year 2001–02

Variable name	Description	Range of Values	Weighted mean	Standard Deviation
C5SDQRDC	C5 SDQ Prcvd Interest/Competence - Reading	1 - 4	3.26	0.66
C5SDQMTC	C5 SDQ Prcvd Interest/Competence - Math	1 - 4	3.16	0.79
C5SDQSBC	C5 SDQ Prcvd Interest/Competence - All Sbj	1 - 4	2.92	0.66
C5SDQPRC	C5 SDQ Prcvd Interest/Competence - Peer Rl	1 - 4	3.03	0.65
C5SDQEXT	C5 SDQ Externalizing Problems	1 - 4	2.02	0.71
C5SDQINT	C5 SDQ Internalizing Problems	1 - 4	2.22	0.74

NOTE: Table estimates based on C5CW0 weight. See chapter 7, section 7.3 for variable naming conventions.
SOURCE: U.S. Department of Education, National Center for Education Statistics, Early Childhood Longitudinal Study, Kindergarten Class of 1998–99 third grade data collection, school year 2001–02.

REFERENCES

Linacre, J., and Wright, B. (2000). *WINSTEPS.* Chicago, IL: MESA Press.

Rock, D.A., and Pollack, J.M. (2002). *A Model-based Approach to Measuring Cognitive Growth in Pre-reading and Reading Skills During the Kindergarten* (ETS Research Report RR-02-18). Princeton, NJ: Educational Testing Service.

U.S. Department of Education, National Center for Education Statistics (2002). *Early Childhood Longitudinal Study, Kindergarten Class of 1998–99* (*ECLS-K) Psychometric Report for Kindergarten Through First Grade* (NCES 2002-005) by Donald A. Rock and Judith M. Pollack, Educational Testing Service, Elvira Germino Hausken, project officer. Washington, DC: U.S. Department of Education.

U.S. Department of Education, National Center for Education Statistics (forthcoming). *Early Childhood Longitudinal Study, Kindergarten Class of 1998–99* (*ECLS-K) Third Grade Methodology Report*. Washington, DC: U.S. Department of Education.

Woodcock, R.W., McGrew, K.S., and Werder, J.K. (1994). *Woodcock-McGrew-Werder Mini-Battery of Achievement*, Itasca, IL: Riverside Publishing.

4. SAMPLE DESIGN AND IMPLEMENTATION

The Early Childhood Longitudinal Study, Kindergarten Class of 1998–99 (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, the effort for 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. As in the first grade data collection, only a subsample of students who had transferred from their kindergarten school was followed. In third grade, however, the subsampling rate applied to transferred children was slightly higher: children whose home language is non-English (also known as children belonging to the language minority group) and who moved for the first time in 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.

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 five-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

			Census region			
SR status	MSA status	Total	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 kindergarten data collection, school year 1998–99.

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. A procedure was implemented to identify kindergarten programs that would be operational at the time of ECLS-K's base year data collection, but that were not included in the frame just described. These were newly opened schools that were not listed in the CCD and the PSS, and 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. Schools that were discovered to be ineligible during recruitment have been omitted from the tabulation.

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
Religious affiliation			
Catholic	120	†	120
Other religious	149	†	149
Nonreligious, private	94	†	94
School type			
Regular	1,162	893	269
Ungraded	4	1	3
No grade beyond K	49	19	30
Unknown	62	1	61

† Not applicable.
SOURCE: U.S. Department of Education, National Center for Education Statistics, Early Childhood Longitudinal Study, Kindergarten Class of 1998–99 kindergarten data collection, school year 1998–99.

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 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.

4.2 Fall-First Grade Subsample

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

Base year schools in the 62 fall-first grade sampled PSUs were stratified by frame source (original public, original private, freshened, etc.) 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.

[1] See the *Early Childhood Longitudinal Study, Kindergarten Class of 1998–99* (*ECLS-K) Base Year Public-Use Data Files and Electronic Code Book: User's Manual* (NCES 2001–029; U.S. Department of Education, National Center for Education Statistics, 2000) for a detailed description of the base year sample.

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

Characteristic	Total	Sector: Public	Sector: Private
Total	22,666	17,777	4,889
Region			
Northeast	4,262	3,045	1,217
Midwest	5,628	4,292	1,336
South	7,461	6,179	1,282
West	5,315	4,261	1,054
Type of locale			
Large city	4,550	3,365	1,185
Midsize city	4,728	3,569	1,159
Urban fringe of large city	6,470	4,945	1,525
Urban fringe of midsize city	1,644	1,434	210
Large town	714	577	137
Small town	1,905	1,485	420
Rural	2,655	2,402	253
Religious affiliation			
Catholic	2,510	†	2,510
Other religious	1,445	†	1,445
Nonreligious, private	934	†	934
School type			
Regular	21,436	17,390	4,046
Ungraded	56	24	32
No grade beyond K	663	338	325
Unknown	511	25	486
Composite child race			
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
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.

SOURCE: U.S. Department of Education, National Center for Education Statistics, Early Childhood Longitudinal Study, Kindergarten Class of 1998–99 kindergarten data collection, school year 1998–99.

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
Religious affiliation			
Catholic	29	†	29
Other religious	33	†	33
Nonreligious, private	21	†	21
School type			
Regular	292	222	70
Ungraded	1	1	0
No grade beyond K	18	5	13

† Not applicable.
SOURCE: U.S. Department of Education, National Center for Education Statistics, Early Childhood Longitudinal Study, Kindergarten Class of 1998–99 first grade data collection, school year 1999–2000.

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. "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). Base year nonrepondents would be adjusted for during weighting. Because of the additional burden of school recruiting, the cost of collecting data for a child who transferred from the school in which he or she was originally sampled greatly exceeds that for a child who stayed enrolled. To contain these costs, 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.

Prior to subsampling with equal probability, children were stratified into groups of nonmovers, movers with information identifying their new schools, and movers without such identifying information. 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 for fall-first grade. Region, locale, religious 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. 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.

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

Table 4-5. Number (unweighted) of children subsampled 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
Religious affiliation			
Catholic	535	†	535
Other religious	254	†	254
Nonreligious, private	415	†	415
School type			
Regular	5,374	4,338	1,036
Ungraded	24	24	0
No grade beyond K	138	84	54
Unknown	114	0	114
Composite child race			
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 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	9	521
High school graduate	1,252	128	1,124
Vocational/technical	335	50	285
Some college	1,419	300	1,119
College graduate	1,038	358	680
Masters	398	157	241
Ph.D./professional	255	130	125
Unknown	423	72	351

† Not applicable.
NOTE: School characteristics (i.e., region, locale, religious 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 first grade data collection, school year 1999–2000.

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, new from Catholic dioceses, new from local governments, etc.). Finally within each frame source, schools were stratified by response status, and 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 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 children in the spring-first grade sample, excluding freshened students. Region, locale, religious affiliation, and school type describe the school at which the child attended kindergarten.

Table 4-6. Number (unweighted) of 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
Religious affiliation			
Catholic	2,091	†	2,091
Other religious	1,139	†	1,139
Nonreligious, private	606	†	606
School type			
Regular	17,277	13,971	3,306
Ungraded	40	24	16
No grade beyond K	420	235	185
Unknown	347	18	329
Composite child race			
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

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	Sector: Private
Highest parent level of education			
Less than high school	1,529	1,491	38
High school graduate	3,779	3,356	423
Vocational/technical	1,078	926	152
Some college	4,211	3,313	898
College graduate	3,348	2,194	1,154
Masters	1,191	719	472
Ph.D./professional	749	395	354
Unknown	2,199	1,854	345

† Not applicable.
NOTE: School characteristics (i.e., region, locale, religious 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 first grade data collection, school year 1999–2000.

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 alphabetic 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 was linked to the base year sampled student (i.e., was assigned that student’s probability of selection) 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 fall 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 *Early Childhood Longitudinal Study, Kindergarten Class of 1998–99 (ECLS-K) User's Manual for the ECLS-K First Grade Public-Use Data Files and Electronic Code Book* (NCES 2002-135; U.S. Department of Education, National Center for Education Statistics, 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. 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. This only affects children who had not moved out of the original sample schools before third grade. If they had moved before third grade, than 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 for analytic purposes. 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 children in the spring-third grade sample, excluding freshened students. Region, locale, religious affiliations, and school type describe the school at which the child attended kindergarten.

4.5 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 spring-first grade and again in spring-third grade.

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

Characteristic	Total	Sector: Public	Sector: 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
Religious affiliation			
Catholic	1,924	†	1,924
Other religious	1,036	†	1,036
Nonreligious, private	544	†	544
School type			
Regular	15,930	12,901	3,029
Ungraded	34	23	11
No grade beyond K	391	222	169
Unknown	315	20	295
Composite child race			
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.

NOTE: School characteristics (i.e., region, locale, religious 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 third grade data collection, school year 2001–02.

The first procedure was 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. 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 and last procedure, applied in first and third 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.

Table 4-8 shows the sample size for each round of data collection of the ECLS-K, and the response status of the children in each round. Fall-first grade is not included in this table, as it pertains only to a subsample of the ECLS-K children. Tables 4-9 and 4-10 show the same children separately by the original sample school type (public/private).

Table 4-8. Number (unweighted) of children in the ECLS-K sample, by response status and data collection round: School years 1998–99, 1999–2000, and 2001–02

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
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

† 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.
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.
SOURCE: U.S. Department of Education, National Center for Education Statistics, Early Childhood Longitudinal Study, Kindergarten Class of 1998–99 kindergarten, first, and third grade data collections, school years 1998–99, 1999–2000, and 2001–02.

Table 4-9. Number (unweighted) of public school children in the ECLS-K sample, by response status and data collection round: School years 1998–99, 1999–2000, and 2001–02

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
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

† 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.
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.
SOURCE: U.S. Department of Education, National Center for Education Statistics, Early Childhood Longitudinal Study, Kindergarten Class of 1998–99 kindergarten, first, and third grade data collections, school years 1998–1999, 1999-2000, and 2001–02.

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

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
Spring-first grade	4,573[2]	11	21	686	192	3,663
Spring-third grade	4,573	23	39	988	288	3,235

† 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.

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.

SOURCE: U.S. Department of Education, National Center for Education Statistics, Early Childhood Longitudinal Study, Kindergarten Class of 1998–99 kindergarten, first, and third grade data collections, school years 1998–99, 1999–2000, and 2001–02.

The number of children who participated in the base year and first grade and third grade data collections is 13,698 (10,900 in original public schools and 2,798 in original private schools). This represents 64 percent of the base year respondents or 60 percent of children sampled for the base year.

4.6 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. As in the first grade year, only child-level weights were computed for third grade. 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 third grade sample was not freshened with third 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 grade data are representative of the population cohort rather than all third graders in 2001–02. The estimated number of children from the ECLS-K is approximately 96 percent of all third graders. While the vast majority of children in third grade in the 2001–02 school year are members of the cohort, third graders who repeated second or third 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. 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.

Several sets of weights were computed for third 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; parent instruments; several types of teacher instruments; 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 sampling weights be computed for use in analyzing the ECLS-K data. Several combinations of kindergarten through third grade longitudinal weights were also computed. Details on these longitudinal weights are available in chapter 9. This section describes the different types of third grade cross-sectional weights, how they were calculated, how they should be used, and their statistical characteristics.

4.6.1 Types of Sample Weights

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

- C5CW0 is nonzero if the child has completed assessment data or the child was excluded from direct assessment due to a disability.
- C5PW0 is nonzero if the child has completed parent interview.
- C5CPTW0 is nonzero if the child has completed assessment data and parent interview data and teacher data from questionnaire part B.

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 third 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 teacher questionnaires.

Exhibit 4-1. ECLS-K third grade cross-sectional weights: School year 2001-02

Weight	to be used for analysis of ...
C5CW0	third 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) any third grade teacher questionnaire A, B or C data, and (c) data from the school administrator questionnaire or school fact sheet.
C5PW0	third grade parent interview data alone or in combination with (a) third grade child assessment data, (b) third grade teacher questionnaire A, B, or C data, and (c) data from the school administrator questionnaire or school fact sheet. *Exception*: If data from the parent AND child assessment AND teacher questionnaire A or B (not C) are used then C5CPTW0 should be used.
C5CPTW0	third grade direct child assessment data combined with third grade parent interview data AND third grade teacher data alone or in conjunction with data from the school administrator or school fact sheet 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) third grade data collection, school year 2001–02.

Weight C5CW0 is used to estimate child-level characteristics or assessment scores for third grade. Examples of such estimates are the percent of third grade children who are male, the percent of children who are API, the percent of children who are 9 years old when they begin third grade, and the mean reading score of children in third grade. 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 fact sheet. 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.

When analyzing child assessment data in conjunction with teacher data collected in third grade, weight C5CW0 should be used. An example of the use of C5CW0 is in the analysis of the relationship between children's approaches to learning as rated by their teachers, the teacher's type of teaching certification, and the children's cognitive skills and knowledge. Some data may be missing because some teachers did not complete the questionnaire, but these are the most appropriate weights for

[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, this is no longer applicable, since there were no children not assessed due to language ability.

this type of analysis. However, different weights should be used for analysis of child data in conjunction with both parent and teacher data (C5CPTW0).

C5PW0 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 are read to every day. 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, when exploring the relationship between home literacy behaviors and children's reading skills.

C5CPTW0 is used when child direct assessment *and* teacher *and* parent data 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. These weights should be not be used for estimates using only direct child assessment data or only parent interview data.

Careful consideration should be given to which set of weights is appropriate for the desired analysis. Using the wrong weights will result in more biased or inefficient estimates. For example, if C5CPTW0 were used in an analysis of child and teacher/classroom data only, then the resulting estimates will be inefficient compared to estimates using C5CW0. The lower parent response causes C5CPTW0 to result in lower sample size with positive weights. There may be combinations of data from a different source for which no weights were developed, but most analyses are possible from the weights provided.

The distribution of schools by number of sampled students with nonzero third grade weights and the mean number of sampled students with nonzero weights per school are useful in analysis using hierarchical linear modeling. These are given in table 4-11. In third grade, 70 percent of schools in the sample have five or fewer ECLS-K students with nonzero third grade weights; 94 percent of these schools with small numbers of children are schools where students transferred to (not in tables). For this reason, schools are classified in table 4-11 on the basis of the number of students who had never transferred schools.

Table 4-11. Distribution of originally sampled schools by number of children with nonzero weights and by type of third grade sample weights: School year 2001–02

Sample	Number of cases					Mean cases per school
	1 – 5	6 – 10	11 – 15	16 – 20	21 – 27	
Spring-third grade						
C5CW0	93	224	344	219	18	12
C5PW0	129	301	309	151	9	11
C5CPTW0	180	278	249	131	7	10

SOURCE: U.S. Department of Education, National Center for Education Statistics, Early Childhood Longitudinal Study, Kindergarten Class of 1998–99 (ECLS-K) third grade data collections, school year 2001–02.

4.6.2 Weighting Procedures

The third grade sample included all base year respondents as defined earlier and a supplemental sample of first graders brought in through a sample freshening procedure implemented in spring-first grade. Only a subsample of children who moved from the schools they were attending when they were sampled originally was followed into their new schools. However, children who belong to the language minority group and who had not moved out of the original sample schools at anytime during the first grade year were all followed into their new third grade schools if they moved from the original sample school during their third grade year. 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 for the first grade year. They are described again 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.

The third and last stage was to rake the weights adjusted in the second stage to sample-based control totals.

The computation of the initial child weights is described in section 4.6.3. The subsequent weight adjustments are described in section 4.6.4. Section 4.6.5 describes the different types of weights computed for spring-third 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.

4.6.3 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 4.6.3.1 for the school-level weights, and in section 4.6.3.2 for the child-level weights.

4.6.3.1 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 but 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 type (public, Catholic private, non-Catholic private, or nonsectarian 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.

4.6.3.2 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.

4.6.3.2.1 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 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 type, 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 child weights in the base year. For a description of the computation of child weights in the base year, see chapter 4, section 4.3.4 of the *Early Childhood Longitudinal Study, Kindergarten Class of 1998–99* (*ECLS-K) Base Year Public-Use Data Files and Electronic Code Book: User's Manual* (NCES 2001–029: U.S. Department of Education, National Center for Education Statistics, 2000).

4.6.3.2.2 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 linked back to the child sampled in first grade and further adjusted for nonresponse due to not obtaining the data from the sample of freshened children.

First the school base year weight adjusted for school nonresponse (as described in section 4.6.3.1) 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 type and census region.

[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.

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

Next, 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 type, locale, region, school enrollment class, and child characteristics such as age group, sex, and race/ethnicity.

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 moved into a new school. For this two-step adjustment, each child was classified as a (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 the step before) 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 step 1 of all first graders (group a or d) over the sum of the weights adjusted in step 1 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 type and census region.

The weights thus created for children sampled in kindergarten were then linked to the children that they 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.

4.6.4 Computation of Spring-Third Grade Child Weights

The initial child weights described in section 4.6.3 were adjusted for movers between the base year and third grade and nonresponse in third grade, and raked to sampled-based control totals to obtain the final spring-third grade child weights.

4.6.4.1 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. Once a child was identified as a mover, he stayed a mover (unless he moved back to the original sample school). The spring-first grade follow flags were maintained for all children in the spring-third grade sample except for children whose home language was not English. For these language minority children, their spring-first grade flags were switched to 1 if they were not already equal to 1 and if they had not already been subsampled out because they moved in spring-first grade. Thus, children who moved out of their original sample school were followed in the random 50 percent of schools where the follow flag was set to 1, and language minority children were followed at 100 percent if they had not moved previously. The adjustment factor for subsampling movers 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; 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, whether they were movers in spring-first grade, whether they were language minority children, the school type of their original sample school, and the region where their original sample school was located. Seven children with large weights had their weights trimmed by half. However, the weights were not redistributed because the total sum of weights was re-established in the raking procedure that came later.

4.6.4.2 Adjustment for Nonresponse

After the adjustment for subsampling movers, the child weights were adjusted for nonresponse. As in spring-first 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. 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 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 (C5PW0 only), and the school type including whether the child was homeschooled (C5CPTW0 only).

4.6.4.3 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 4.6.3. 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-third grade. The sum of weights thus calculated is the estimated number of third graders in spring 2002. 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-third 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.

Before raking the C5CPTW0 weights, 11 responding movers had their nonresponse-adjusted weights trimmed and the excess weight redistributed among the remaining responding movers so that the sum of weights before trimming was equal to the sum of weights after trimming.

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 type, region, urbanicity, sex, age, race/ethnicity, SES, language minority status, whether sampled in kindergarten or first grade, and if sampled in kindergarten, mover status.

4.6.5 Types of Weights and Their Use

The different types of cross-sectional weights are described in section 4.6.1 and their use was summarized in exhibit 4-1. They were all created as described in sections 4.6.4.2 and 4.6.4.3, but the definition of which children were eligible respondents varied. The adjustment for movers was done once, then the resulting weights were adjusted for nonresponse separately for C5CW0, C5PW0 and C5CPTW0.

4.6.5.1 Weights To Be Used With Direct Child Assessment Data (C5CW0)

In spring-third grade, responding children for this type of weight were eligible children who had spring-third grade scorable direct child cognitive assessment data, or children with disabilities who according to specifications in their IEP 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-third grade, responding children were classified using rules similar to those used in spring-first grade.

Table 4-12 shows the number of children who were not assessed due to the following special situations: children with disabilities, children who moved out of their original sample schools and were not flagged to be followed, children who moved and were flagged to be followed but could not be located or moved into a school in a nonsampled county, and children who moved outside of the country or who were deceased. Of these, only children with disabilities had weights.

Table 4-12. Number of children who were not assessed in spring-third grade, by special situations: School year 2001–02

Special situation	Number of children: Unweighted	Number of children: Weighted
Spring-first grade		
Children with disabilities[1]	75	22,334
Moved from original sample schools		
Subsampled, not to be followed	4,117	784,957
Nonlocatable or moved to nonsampled PSU	1,481	282,443
To be followed but were ineligible in spring	72	13,599

[1] These children's individualized education programs (IEPs) specifically prohibited assessments.
SOURCE: U.S. Department of Education, National Center for Education Statistics, Early Childhood Longitudinal Study, Kindergarten Class of 1998–99 third grade data collection, school year 2001–02.

4.6.5.2 Weights To Be Used With Parent Data (C5PW0)

The weight C5PW0 is to be used with parent interview data. In spring-third 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 third grade children, not to the parents of third grade children.

4.6.5.3 Weights To Be Used With a Combination of Child Direct Assessment Data and Parent Interview Data and Teacher Data (C5CPTW0)

The weight C5CPTW0 is to be used for analysis involving child, parent, and teacher data. A respondent for this type of weight was defined as a child who had scorable cognitive assessment data for spring-third grade (or children with disabilities), whose parent completed the FSQ section of the parent interview for spring-third grade, and whose teacher completed part B of the teacher questionnaire.

4.6.6 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-third 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 C5CW0 are C5CW1 through C5CW90. 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 4.7.

4.6.7 Characteristics of Sample Weights

The statistical characteristics of the sample weights are presented in table 4-13. 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 types of weight is due a combination of factors, among them: (1) the number of first graders who became ineligible in third 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.

Table 4-13. Characteristics of the first grade child-level weights: School year 1999–2000

Sample	Number of cases	Mean	Standard deviation	CV (× 100)	Minimum	Maximum	Skewness	Kurtosis	Sum
C5CW0	14,470	272.18	242.53	89.10	1.54	3,376.78	3.21	18.45	3,938,513
C5PW0	13,489	291.92	241.71	82.80	1.63	3,654.05	3.23	18.83	3,937,758
C5CPTW0	10,395	378.75	435.34	114.94	2.58	5,209.19	3.38	15.25	3,937,125

SOURCE: U.S. Department of Education, National Center for Education Statistics, Early Childhood Longitudinal Study, Kindergarten Class of 1998–99 third grade data collection, school year 2001–02.

4.7 Variance Estimation

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

4.7.1 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:

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

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 4.8 were produced using WesVar and JK2.

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.

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 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 type (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 also numbered sequentially from 1 to 63, and so on until the list was exhausted, thus forming the desired 63 variance strata.

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

A child sampled in first grade through the freshening process was assigned to the same replicate as the originally sampled child to whom the child was linked. When the child sampled in first grade was assigned a full sample weight (see section 4.6.3.2), 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, SUDAAN and AM.

4.7.2 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. 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, 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 or AM. They are described in table 4-14.

Table 4-14. ECLS-K Taylor Series stratum and first-stage unit identifiers

Variable name	Description
C5TCWSTR	Sampling stratum—spring-third grade C-weights
C5TCWPSU	First-stage sampling unit within stratum—spring-third grade C-weights
C5TPWSTR	Sampling stratum—spring-third grade P-weights
C5TPWPSU	First-stage sampling unit within stratum—spring-third grade P-weights
C5CPTSTR	Sampling stratum—spring-third grade CPT-weights
C5CPTPSU	First-stage sampling unit within stratum—spring-third grade CPT-weights

SOURCE: U.S. Department of Education, National Center for Education Statistics, Early Childhood Longitudinal Study, Kindergarten Class of 1998–99 third grade data collection, school year 2001–02.

4.7.3 Specifications for Computing Standard Errors

Specifications for computing standard errors are given in table 4-15. For each type of analysis described in the table, users can choose the replication method or the Taylor Series method for computing standard errors.

Table 4-15. Specifications for computing standard errors, spring-third grade: School year 2001–02

		Computing standard errors					Approximating sampling errors
		Replication method (WesVar, SUDAAN or AM)			Taylor Series method (SUDAAN, STATA, SAS or AM)		
Type of analysis	Full sample weight	ID	Replicate weights	Jackknife method	Sample design	Nesting variables	DEFT (Average root design effect)
Spring-third grade cross-sectional	C5CW0	CHILDID	C5CW1 – C5CW90	JK2	WR[1]	C5TCWSTR C5TCWPSU	1.841
	C5PW0	PARENTID	C5PW1 – C5PW90	JK2	WR[1]	C5TPWSTR C5TPWPSU	
	C5CPTW0	CHILDID	C5CPTW1 – C5CPTW90	JK2	WR[1]	C5CPTSTR C5CPTPSU	

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

SOURCE: U.S. Department of Education, National Center for Education Statistics, Early Childhood Longitudinal Study, Kindergarten Class of 1998–99 third grade data collection, school year 2001–02.

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-third grade child-level estimates (e.g., mean reading scores) and their standard errors, users need to specify CHILDID in the ID box of the WesVar data file screen, C5CW0 as the full sample weight, C5CW1 to C5CW90 as the replicate weights, and JK2 as the method of replication.

For the Taylor Series method using SUDAAN, STATA, SAS 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 (C5CW0), the stratum variable (C5TCWSTR), and the PSU variable (C5TCWPSU) must be specified. The "with replacement" sample design option, WR, must also be specified if using SUDAAN.

The last column in table 4-15 gives the average root design effect that can be used to approximate the standard errors for each type of analysis. For a discussion of the use of design effects, see section 4.8.1.

4.8 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.

4.8.1 Use of Design Effects

Methods of computing standard errors for the ECLS-K are replication and Taylor Series linearization. If a statistical analysis software package such as SPSS (Statistical Program for the Social Sciences) is used, the standard errors should be corrected using *DEFT*, since these programs calculate standard errors, assuming the data were collected with a simple random sample. The standard error 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.

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

4.8.2 Average 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. One way to produce design effects for analysts' use is to calculate them for a number of variables and average them. The averaging can be done overall and for selected subgroups. The tables that follow show estimates, standard errors, and design effects for selected means and proportions based on the ECLS-K third grade child, parent, teacher, and school data. For each survey item, the tables present 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 estimate, the variable name as it appears in the ECLS-K first grade Electronic Code Book 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.

Standard errors and design effects for the child-level items are presented in tables 4-16 and 4-17. The survey items were selected so that there was a mix of items from the direct child assessment, the parent interview, and the teacher child-level questionnaire They include the different scores from the direct child assessment, 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 4-16 presents design effects for the third grade sample, with a median design effect of 3.3. Table 4-17 presents the median design effects for subgroups based on school type, child’s sex and race/ethnicity, geographic region, level of urbanicity, and the socioeconomic scale (SES quintiles) of the parents. The median design effect varies from 1.3 (Pacific Islanders) to 3.6 (children in small towns and rural areas). The variation in the design effects is largely a function of the sample size as well as the homogeneity of the children within schools.

In spring-third grade, as in spring-first grade, design effects are not computed for items from the teacher and school administrator’s questionnaires since there are no teacher or school weights computed for the first grade year nor for the third grade year. Although standard errors 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.

Table 4-16. ECLS-K standard errors and design effects by selected child and parent variables, for the full sample—child and parent data: School year 2001–02

Survey item	Variable name	Number of cases	Estimate	Design SE[1]	SRS SE[2]	DEFT[3]	DEFF[4]
Child scores (mean)							
Reading scale score	C5R2RSCL	14,280	106.05	0.437	0.173	2.521	6.354
Math scale score	C5R2MSCL	14,374	83.25	0.433	0.152	2.843	8.083
Science scale score	C5SSCALE	14,351	33.46	0.222	0.083	2.661	7.079
Self-described : Externalizing problems	C5SDQEXT	14,379	2.02	0.012	0.006	2.063	4.254
Self-described : Internalizing problems	C5SDQINT	14,379	2.23	0.013	0.006	2.187	4.782
Self-described : Competence in math	C5SDQMTC	14,379	3.16	0.011	0.006	1.718	2.951
Self-described : Competence in peer relation	C5SDQPRC	14,378	3.03	0.007	0.005	1.323	1.750
Self-described : Competence in reading	C5SDQRDC	14,379	3.26	0.009	0.005	1.690	2.855
Self-described : Competence in all subject	C5SDQSBC	14,379	2.92	0.008	0.005	1.497	2.240
Approaches to learning-Teacher	T5LEARN	11,701	2.99	0.010	0.007	1.526	2.330
Self-control-Teacher	T5CONTRO	11,592	3.17	0.010	0.006	1.651	2.725
Interpersonal-Teacher	T5INTERP	11,558	3.05	0.009	0.006	1.407	1.979
Externalizing problems-Teacher	T5EXTERN	11,676	1.73	0.010	0.006	1.760	3.097
Internalizing problems-Teacher	T5INTERN	11,577	1.67	0.008	0.005	1.590	2.529
Child and parent characteristics from parent interview (percent)							
Lived in single parent family	P5HFAMIL	13,489	27.80	0.627	0.386	1.626	2.644
Lived in two-parent family	P5HFAMIL	13,489	69.70	0.711	0.396	1.797	3.229
Mom worked 35 hours+/week	P5HMEMP	9,790	67.08	0.782	0.475	1.646	2.709
Primary care is center-based	P5PRIMNW	4,765	36.41	1.315	0.697	1.887	3.561
Primary care is home-based	P5PRIMNW	4,765	63.59	1.315	0.697	1.887	3.561
Parents had high school or less	W3PARED	13,489	33.42	0.882	0.406	2.172	4.716
Household income category below median	W3INCCAT	13,489	42.88	0.995	0.426	2.335	5.450
Parent attended PTA	P5ATTENP	13,470	43.03	0.887	0.426	2.080	4.327
Practiced reading, writing, numbers every day	P5RDWRNM	13,364	48.90	0.609	0.433	1.407	1.981
Visited library	P5LIBRAR	13,362	54.01	0.807	0.431	1.871	3.500
Used computer 1-2 times per week	P5HOMECM P5COMPWK	10,671	41.57	0.625	0.477	1.311	1.718
Had internet access	P5HOMECM P5INTACC	10,417	85.61	0.466	0.344	1.356	1.840
Used computer 1-2 times per week for homework	P5HOMECM P5CMPEDU	10,394	52.58	0.593	0.490	1.210	1.465
Had family rule for TV	P5TVHOME P5TVRULE	13,285	91.58	0.323	0.241	1.342	1.802
Did homework 3-4 times per week	P5OFTDHW	13,344	40.55	0.675	0.425	1.588	2.522
Have someone help with reading homework	P5HELPR	13,161	98.50	0.164	0.106	1.546	2.389
Completely true that child and self have close time	P5WARMCL	13,021	68.48	0.572	0.407	1.404	1.971
Took away privilege when child angry	P5HITPRV	13,016	62.17	0.996	0.425	2.342	5.486
Self-reported in very good health	P5HEALTH	12,995	86.81	0.560	0.297	1.885	3.555
Household received food stamp in last 12 months	P5FSTAMP	13,256	15.15	0.729	0.311	2.341	5.479
Child characteristics from teacher questionnaire C (percent)							
Enrolled in third grade	T5GRADE	11,721	88.33	0.619	0.297	2.086	4.351
Average in language skills	T5RTLANG	11,581	70.78	0.773	0.423	1.828	3.343
Average in science/social studies	T5RTSCI	11,500	79.68	0.891	0.375	2.375	5.639
Average in math skills	T5RTMTH	11,544	77.21	0.802	0.390	2.054	4.219

See notes at end of table.

Table 4-16. ECLS-K standard errors and design effects by selected child and parent variables, for the full sample—child and parent data: School year 2001–02—Continued

Survey item	Variable name	Number of cases	Estimate	Design SE[1]	SRS SE[2]	DEFT[3]	DEFF[4]
Child characteristics (mean)							
Age of child in months	R5AGE	14,393	111.22	0.093	0.040	2.346	5.506
Child's BMI	CBMI	13,866	18.69	0.044	0.033	1.332	1.775
Child's household size	P5HTOTAL	13,489	4.56	0.024	0.012	1.960	3.843
Number of children <18 in child's HH	P5LESS18	13,489	2.52	0.020	0.010	2.009	4.036
Number of siblings in HH	P5NUMSIB	13,489	1.55	0.018	0.010	1.850	3.421
Number of hours watched TV after dinner	P5TVAFDH	13,263	0.89	0.012	0.007	1.700	2.891
Median						1.813	3.343
Mean						1.841	3.559
Standard deviation						0.397	1.567
Coefficient of variation						0.216	0.440
Minimum						1.210	1.465
Maximum						2.843	8.083

[1]Design SE is the standard error under the ECLS-K sample design. For an explanation of this statistic, see section 4.8.
[2] SRS SE is the standard error assuming simple random sample. For an explanation of this statistic, see section 4.8.
[3] DEFT is the root design effect. For an explanation of this statistic, see section 4.8.
[4] DEFF is the design effect. For an explanation of this statistic, see section 4.8.
SOURCE: U.S. Department of Education, National Center for Education Statistics, Early Childhood Longitudinal Study, Kindergarten Class of 1998–99 third grade data collection, school year 2001–02.

Table 4-17. ECLS-K median design effects for subgroups—child, parent, and teacher questionnaire part C data: School year 2001–02

Subgroups	Spring-third grade	
	DEFT[1]	DEFF[2]
All students	1.813	3.343
School type		
Public	1.669	2.841
Private	1.788	3.203
Catholic private	1.796	3.262
Other private	1.585	2.576
Sex		
Male	1.606	2.641
Female	1.512	2.326
Race/ethnicity		
White	1.717	2.945
Black	1.514	2.305
Hispanic	1.392	1.962
Asian	1.481	2.178
Pacific Islander	1.134	1.277
American Indian	1.387	1.905
Other	1.369	1.870
Region		
Northeast	1.675	2.812
Midwest	1.789	3.208
South	1.765	3.290
West	1.663	2.765
Urbanicity		
Central city	1.608	2.601
Urban fringe and large town	1.636	2.658
Small town and rural area	1.903	3.558
SES quintiles		
First	1.462	2.128
Second	1.430	2.049
Third	1.429	2.037
Fourth	1.376	1.868
Fifth	1.453	2.112

[1]DEFT is the root design effect. For an explanation of this statistic, see section 4.8.
[2]DEFF is the design effect. For an explanation of this statistic, see section 4.8.
NOTE: Each median is based on 44 items.
SOURCE: U.S. Department of Education, National Center for Education Statistics, Early Childhood Longitudinal Study, Kindergarten Class of 1998–99 third grade data collection, school year 2001–02.

REFERENCES

Durbin, J. (1967). Design of Multi-stage Surveys for the Estimation of Sampling Errors. *Journal of the Royal Statistical Society*, Series C (Applied Statistics), Vol. 16, 152-164.

Kish, L. (1965). *Survey Sampling*. New York: John L. Wiley & Sons.

U.S. Department of Education, National Center for Educational Statistics (1995–96). *Common Core of Data, Public School Universe Survey*. Washington, DC: U.S. Department of Education.

U.S. Department of Education, National Center for Education Statistics (2000). *Early Childhood Longitudinal Study, Kindergarten Class of 1998–99 (ECLS-K) Base Year Public-Use Data Files and Electronic Code Book: User's Manual* (NCES 2001–029) Washington, DC: U.S. Department of Education.

U.S. Department of Education, National Center for Education Statistics (2002). *Early Childhood Longitudinal Study, Kindergarten Class of 1998–99 (ECLS-K) User's Manual for the ECLS-K First Grade Public-Use Data Files and Electronic Code Book* (NCES 2002-135). Washington, DC: U.S. Department of Education.

U.S. Department of Education, National Center for Education Statistics (1998). *Private School Universe Survey, 1995–96*. Washington, DC: U.S. Department of Education.

U.S. Department of the Interior, Bureau of Indian Affairs (1995-96). *Office of Indian Education Programs Education Directory*. Washington, DC: U.S. Department of the Interior.

5. DATA COLLECTION METHODS AND RESPONSE RATES

The following sections discuss the data collection procedures and response rates in the third 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 field staff training (section 5.2), preassessment school contacts (section 5.3), spring-third grade data collection (section 5.4), and quality control procedures (section 5.5). Spring-third grade completion rates are presented and discussed in section 5.6.

5.1 Overview of Data Collection Methods

The ECLS-K third grade data collection was conducted in the fall and spring of the 2001–02 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 locating information about their new schools. Spring data collection included the direct child assessments, parent interviews, teacher and school questionnaires, student record abstract, and facilities checklist. The activities, begun in fall data collection, to locate children and gain cooperation of the schools into which they transferred continued in spring data collection. The content and timeline of the third grade data collections are shown in exhibit 5-1.

Computer-assisted personal interviewing (CAPI) was the mode of data collection for the child assessments, and telephone and in-person computer-assisted interviewing (CAI) was the mode of data collection for 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.

5.2 Field Staff Training

Several in-person training sessions were conducted to prepare staff for the third grade data collection. In the fall of 2001, supervisors were trained to contact original schools and recruit transfer

Exhibit 5-1. Timeline of third grade data collection: School year 2001–02

Third grade		
2001	2002	
Fall	Winter	Spring

Fall 2001:

Advance school contact

Tracing sampled households

Spring 2002:

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 third grade data collection, school year 2001–02.

schools. In the spring of 2002, three trainings were held: one for staff who only conducted parent interviews, one for field supervisors, and one for assessors. Field supervisors managed all the data collection activities within their assignment, supervising the assessors and interviewers and conducting child assessments and parent interviews. Assessors conducted the child assessments and parent interviews. Interviewers only conducted parent interviews.

5.2.1 Advance Contact and Recruitment Training

In September 2001, field supervisors were trained for 3 days to contact original sampled schools and transfer schools to set up the data collection in the spring. A total of 50 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 moved from the schools they attended in the first 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 first 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-Third Grade Training

Field supervisors, interviewers, and assessors were trained for the spring-third grade data collection in three sessions in February and March 2002. Prior to the March in-person training session, supervisors and assessors completed 8 hours of home study training on the study design, field procedures, and computer keyboard skills.

5.2.2.1 Parent Interviewer-Only Training

Supervisors and staff assigned to complete only parent interviews during the spring data collection attended a 2-day training in February 2002. Trainers presented the content of the parent interview and discussed protocols for interviewing. The interviewers practiced using the CAI system on laptops during interactive lectures and role plays. Supervisors had an additional day of training to practice using the FMS to organize and track production and to discuss management techniques for overseeing their teams of interviewers. Eight (8) supervisors and 66 interviewers completed training.

5.2.2.2 Field Supervisor Training

Field supervisor training preceded the assessor training and lasted for 3 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 moved from their first 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. Seventy-seven (77) field supervisors completed training.

5.2.2.3 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 to conduct the direct child assessments. This included following standardized procedures for administration of all assessment items as well as giving children neutral praise with the sampled children. 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 5 days; field supervisors were also trained to perform all assessor activities. Two hundred sixty-six (266) assessors and 77 field supervisors completed training.

5.2.2.4 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, all field staff 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. Each level form of an assessment domain was reviewed in detail during an interactive lecture. Time was then given to each trainee to review and practice administering it individually. 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. Just over a quarter of the trainees (26 percent or 84 trainees) did not pass at least one element of the certification exercises on the first attempt. Nineteen (19) percent did not pass the reading exercises, with only a small percentage not passing the math (2 percent) and science (5 percent) exercises. Most likely, this was due to the complexity of the reading scoring rubrics and the unfamiliarity of the exercises themselves (reading exercises were distributed first, with math 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 *Early Childhood Longitudinal Study, Kindergarten Class of 1989–1999, Third Grade Methodology Report* (U.S. Department of Education, National Center for Education Statistics, forthcoming [c]) for additional detail.

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 the yellow reading level—see chapter 3) of the assessment to third grade 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. Trainees who scored 85 percent or above were certified as 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. Number and percent of trainees, by scores on training certification form: School year 2001–02

Trainees	Number	Percent
Total	343	100
Score on certification form		
85 percent or above	337	98.2
70–84 percent	6	1.8
Below 70 percent	0	0

SOURCE: U.S. Department of Education, National Center for Education Statistics, Early Childhood Longitudinal Study, Kindergarten Class of 1998–99 third grade data collection, school year 2001–02.

The majority of the trainees (98.2 percent) scored at or above 85 percent on the certification form, with only 1.8 percent scoring 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 qualified to administer the child assessments after they conducted a second assessment on a third grade child who was not part of the ECLS-K sample.

5.3 Fall Preassessment School Contact

Beginning in September 2001, all participating ECLS-K schools, i.e., schools that participated in fall or spring of kindergarten or first grade, were contacted by telephone to prepare for the spring data collection. When children were identified as transferring to another school, the child's new school (and district, if necessary) was recruited.

5.3.1 Advance Mailings

In September 2001, an advance package was mailed via Federal Express to all participating ECLS-K schools asking them to prepare for the preassessment contact 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 third 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 preassessment contact was made by telephone between September and November 2001. The 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 withdrew from their spring-first grade school. Schools were determined to be ineligible for third grade data collection if no sampled children were currently enrolled. Original sampled schools became ineligible if second 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 visit for original sampled schools, identified ECLS-K sampled children who were no longer enrolled at the school, collected locating information for those children, identified each enrolled child's regular and special education teacher, reviewed parental consent status, obtained information on special accommodations[1] during assessment for the enrolled sampled children, and answered any questions the school coordinator may have had.

[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.

5.3.2.1 Identifying ECLS-K Sampled Children Who Withdrew From the School

Field supervisors asked the school coordinators to identify ECLS-K children who 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 transferred, only a subset were followed to their new school (see section 4.4.1 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.

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

5.3.2.3 Reviewing Parent Consent

Although parental consent was obtained in the base year (and, in some schools, in the first grade year), field supervisors reviewed the parental consent with the school coordinator to determine if the base year or first grade consent was acceptable for third grade. If the schools required consent to be re-obtained or changed the type of consent that was required (e.g., from implicit to explicit), parent letters and consent forms were mailed either 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.

5.3.2.4 Contacting Families of Home-Schooled Children

As part of the advance school contact, children who were home schooled in previous rounds were identified. The status of home-schooled children who were identified in round 1 through 4 was verified with their parents and updated as necessary. In addition, some home-schooled children were identified by the schools during the preassessment contact. Their status was also verified with their parents during data collection. Parents of these children were contacted in September through November 2001 to determine if the child was still home schooled 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.

5.3.3 Preparing for Spring-Third Grade Data Collection

In order to ensure that as many of the sampled children as possible were contacted in the spring, locating efforts were undertaken in the winter of 2001. Staff in Westat’s Telephone Research Center (TRC) traced children who could not be located during the preassessment school contact phase. 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 5.4.4 for more details about children who transferred schools in third grade.

A mailing to the post office requesting change of address information for sampled households was also conducted in the winter of 2001.

5.4 Spring-Third 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-third grade data collection. Eligibility for the study was not dependent on the child’s current grade, that is, children were eligible whether they were promoted to third grade or were retained in second grade.

Table 5-2. Results of the Telephone Research Center's locating efforts, spring–third grade: School year 2001–02

Result	Number	Percent
Total cases worked	781	100.0
Located and entered into database	307	39.3
Unlocatable	426	54.5
Out of scope	27	3.5
Final refusal	16	2.0
Partially located	3	0.4
Unable to locate due to language barrier	2	0.3

NOTE: "Unlocatable" means that the children and their households could not be found using the available tracing and locating strategies; "out of scope" means that a child was ineligible to participate; "final refusal" means that the child's family indicated that they did not want to participate; "partially located" means that the tracing and locating effort yielded some information about the child, but not enough to definitively locate the child; "unable to locate due to language barriers" means that the household language was not English and no staff were available who were bilingual in that language.
SOURCE: U.S. Department of Education, National Center for Education Statistics, Early Childhood Longitudinal Study, Kindergarten Class of 1998–99 third grade data collection, school year 2001–02.

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 three 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, distributed and collected all school and teacher questionnaires, and completed school facilities checklists. The majority of field staff members in third grade were continuing from previous rounds of data collection; a few new staff were hired in areas where no experienced ECLS-K staff lived.

5.4.1 Preassessment School Contact

Based on the information collected in the fall of 2001, packets of hard-copy teacher and school administrator questionnaires were assembled and mailed to schools in February 2002, along with letters confirming the scheduled visits to the school. Teachers and school administrators were asked to either complete the questionnaires for pickup on assessment day, or to return the questionnaires in a Federal Express mailer that was provided in the packet. Letters were also mailed to parents reminding them of the spring-third grade data collection activities.

Field supervisors conducted most preassessment activities by telephone starting in March 2002. The preassessment activities for these schools were similar to those conducted in previous rounds of data collection and included confirming the assessment date and the receipt of the hard-copy questionnaires and arranging for space to conduct the assessments.

5.4.2 Timeline of the Direct Child Assessments

The direct child assessments were conducted from March through June 2002, the same time of year as in prior spring data collections. Conducting the child assessments began in March with 91 percent of the assessments completed between April and May and a small percentage (9 percent) 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.

5.4.2.1 Conducting the Direct Child Assessments

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 children that were 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. When scheduling schools in the fall, an attempt was made to conduct the direct child assessments at about the same point in time from the beginning of school year and at the end of the year to increase the chances that exposure to instruction was about the same for all children. The third grade direct child assessments averaged 94 minutes.

Table 5-3 displays the number of completed child assessments for each round of data collection, including spring-third grade. All of the assessments in spring-third grade were completed in English. Most (74.6 percent) of these assessments were completed in original schools, although the number of assessments in transfer schools has grown at each data collection point. In spring-third grade, a quarter of the sample was assessed in transfer schools.

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

Characteristic	Fall-kindergarten		Spring-kindergarten		Fall-first grade		Spring-first grade		Spring-third grade	
	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
Assessed										
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
In Spanish	1,008	5.3	724	3.6	176	3.3	286	1.7	†(a)	†(a)
In other language	410	2.1	229	1.1	33	0.6	37	0.2	†(a)	†(a)
With accommodation[1]	515	2.7	579	2.9	195	3.7	761	4.6	814	5.6
Excluded	88	0.5	70	0.4	28	0.5	47	0.3	74	0.5
Partial complete	107	0.6	43	0.2	17	0.3	31	0.2	49	0.3
Child assessments completed	19,147	100.0	19,987	100.0	5,297	100.0	16,622	100.0	14,502	100.0
Original sampled school	19,147	100.0	19,463	97.4	4,867	91.9	14,830	89.2	10,820	74.6
Transfer school	†(b)	†(b)	524	2.6	430	8.1	1,792	10.8	3,682	25.4

†(a) Not applicable. The assessment was conducted only in English in third grade.
†(b) Not applicable. There were no transfer schools in fall-kindergarten.
[1]The term *accommodation* in this table is the field operational definition of accommodation, which includes the wearing of glasses and hearing aids. These types of aids were systematically tracked to ensure that every child had the same chance at a successful assessment. With this information, assessors could prompt a child, for example, 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 5.4.4.) 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 kindergarten, first, and third grade data collections, school years 1998–99, 1999–2000, and 2001–02.

5.4.2.2 Accommodations and Exclusions

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

Table 5-4. Number of children excluded from and requiring an accommodation in the spring-third grade assessments: School year 2001–02

Category	Number of children
Exclusions	
Excluded for disability	74
Accommodations[1]	
Alternative setting accommodation	33
Scheduling/timing accommodation	65
Health care aide present	6
Assistive device	4

[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 third grade data collection, school year 2001–02.

5.4.3 Conducting the Parent Interview

Parent interview procedures mirrored those of the base year and first grade. The parent interview was administered, primarily by telephone interview using CAI, between March and July 2002.

Sixteen percent of the parent interviews were completed in March, 54 percent were completed in April and May, and 30 percent were completed in June or later. The parent interview averaged 62 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-third grade. In third grade, only 2.4 percent of all completed parent interviews were conducted in person; 7.2 percent of all completed parent interviews were conducted in a language other than English with 95.8 percent of completed non-English interviews conducted in Spanish. A special effort to build parent response rates was conducted between July 5 and 31, 2002 and yielded an additional 7.3 percentage points in the response rate. Almost 8 percent (7.7 percent) of the parent interviews were not completed because of locating problems.

5.4.4 Conducting Data Collection on Children Who Withdrew 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 first 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-first 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-first 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 third grade; 9,889 children were identified as having transferred from the school in which they were enrolled during the spring of first grade. Of the 9,889 children who moved in spring-third grade, 5,668 were in scope, i.e., children selected to be followed, and followed (57.3 percent of total movers). The remaining 4,221 mover children were out of scope and were not followed; no child assessments or parent interview was conducted for these children.

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, and 2001–02

Data collection modes and language	Fall-kindergarten		Spring-kindergarten		Fall-first grade		Spring-first grade		Spring-third grade	
	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
In person	618	3.4	619	3.3	211	4.2	456	2.9	319	2.4
By phone	17,379	96.6	18,288	96.7	4,862	95.8	15,120	97.1	13,185	97.6
Parent interviews completed	17,997	100.0	18,907	100.0	5,073	100.0	15,576	100.0	13,504	100.0
In English	17,379	96.6	17,482	92.5	4,717	93.0	14,319	91.9	12,416	91.9
In Spanish	618	3.4	1,321	7.0	351	6.9	1,071	6.9	932	6.9
In other language	0	0	81	0.4	0	0	75	0.5	41	0.3
Partial complete	0	0	23	0.1	5	0.1	111	0.7	115	0.9

NOTE: This table reflects 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 kindergarten, first, and third grade data collections, school years 1998–2002.

Table 5-6. Number and percent of spring-third grade children who moved from their spring-first grade school, by scope and completion category: School year 2001–02

Scope and completion category	Spring-third grade	
	Number of children	Percent
Total movers	9,889	100.0
Out-of-scope[1]	4,221	42.7
Did not follow[2]	4,102	97.2
Moved to outside of U.S.[2]	117	2.7
Deceased[2]	2	<1
In-scope and followed[1]	5,668	57.3
Completed assessment[3]	3,682	65.0
Unlocatable[3]	607	10.7
Nonsampled primary sample unit[3]	871	15.4
Assessment refused[3]	323	5.7
Not assessed/absent[3]	185	3.3

[1] Percent based on total movers.
[2] Percent based on out-of-scope children.
[3] Percent based on in-scope children.
NOTE: Detail may not sum to totals due to rounding.
SOURCE: U.S. Department of Education, National Center for Education Statistics, Early Childhood Longitudinal Study, Kindergarten Class of 1998–99 third grade data collection, school year 2001–02.

Different data collection strategies were followed for children who moved, depending on where they moved to and the status of their new school. Data collection was attempted for children who moved and were flagged as “follow” in spring-third 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.

- 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. Parent interviews were attempted for all children.

- 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. Parent interviews were attempted for all children.

- 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. Parent interviews were attempted for all children.
  - If the school was outside the PSU, no child, school, or teacher data were collected. The parent interview was still attempted.
- For children who were not enrolled in school in the spring (including children who were home schooled), 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. Parent interviews were attempted for all children.

Of the children who were identified as movers in third grade and who were selected to be followed, 15.4 percent moved into a school outside the PSU, and 10.7 percent could not be located. Assessments were completed for 65 percent of the movers who were followed in the spring-third grade data collection, and parent interviews were completed for 68 percent of these children.

5.4.5 Teacher and School Data Collection

Data were collected from school administrators, regular classroom teachers, and special education teachers from March through June 2002.

The school and teacher questionnaires were mailed to the school coordinators in February 2002. Using the child-teacher linkage information collected in the fall, a packet of questionnaires was assembled for each regular and special education teacher. The regular teacher packet included a cover letter, a sheet explaining the study and its goals, and teacher questionnaire part A, teacher questionnaire part B, and teacher questionnaire part C for each student who had been linked to the teacher in the fall. The special education teacher packet contained a cover letter and summary sheet, special education teacher questionnaire part A, and special education teacher questionnaire part B for each sampled student linked to the teacher. Packets were bundled together by school and mailed to the school coordinator for distribution. If the school and/or teacher and school administrator were not identified in the fall advance contact, then the supervisor gathered the relevant information during the preassessment call in the spring and mailed the packets at that time.

Teachers were asked to complete individual ratings for the sampled children in their classrooms, and they were paid $7 for each child rating (teacher questionnaire part C) they completed. In addition, school staff were asked to complete a student record abstract after the school year closed and were paid $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 by telephone and visits to the schools to ensure that questionnaires were not missing critical information and that completed questionnaires were mailed to Westat. To improve response rates, in early September 2002 a package was mailed to all schools with outstanding school fact sheets or student record abstracts with a request to complete and return questionnaires. School staff were prompted by telephone for the return of the questionnaires and abstracts through October 2002. The hard-copy followup increased child-level responses rates for the school fact sheet by 10 percent and the student record abstract by 12 percent.

5.5 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 practiced the direct child assessments with third 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.

5.5.1 Child Assessments Observations

Field supervisors conducted on-site observations of the child assessments and completed the child observation form. Our quality control plan called for conducting two observations for each of the 266 assessors who completed training. The first observation was to be within 2 weeks of the start of the assessments, and the second observation within 3 weeks of the first observation. These procedures were followed for the majority of assessors (over 80 percent), but some assessors were observed only once due to the school year ending or to the travel distance involved.

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.

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. Table 5-7 presents the result of the observations.

Table 5-7. Results of the child assessments observations, spring-third grade: School year 2001-02

Observations[1]	Number	Percent
Total	468	100
Score on certification form		
85 percent or above	468	100
70–84 percent	0	0
Below 70 percent	0	0

[1] A total of 258 assessors were initially observed. (Eight trainees chose not to continue with data collection.) Of these, 200 received a second observation visit and 10 received a third observation visit. No assessor scored below 85 percent on any of the observation visits.

SOURCE: U.S. Department of Education, National Center for Education Statistics, Early Childhood Longitudinal Study, Kindergarten Class of 1998–99 third grade data collection, school year 2001–02.

5.5.2 Validation of Parent Interviews

Approximately 10 percent of the completed parent interviews were called back by a field supervisor (i.e., validated). The first parent interview completed by an assessor was always validated. 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.

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
- Eight to ten questions from the parent interview that were re-asked of the parent.

5.5.3 Validations 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-third grade data collection.

Field managers conducted the school validations by telephone. The first school that each team completed 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-third grade, a total of 155 school visits were validated with no negative reports of the assessment team or study made by school staff.

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.

5.5.4 Assessor Interrater Reliability

As part of the child assessments observation described in section 5.5.1, 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. These items had open-ended responses that called for interpretation on the part of the assessor to determine whether a child's response was correct. 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 is very high throughout all the levels. It is highest for math (97 percent or better on all levels) and lowest for reading, with the reading yellow level (the medium reading level) showing

Table 5-8. Interrater reliability on child assessment validation item, by subject area and level: School year 2001–02

Subject and level	Number of observations	Number of validation items	Percent agreement: assessors and observers[1]
Reading	462	13	90
Routing	†	0	†
Low (Red)	118	3	93
Middle (Yellow)	262	5	87
High (Blue)	82	5	94
Mathematics	464	7	99
Routing	464	1	100
Low (Red)	135	1	99
Middle (Yellow)	186	2	99
High (Blue)	143	3	97
Science	453	12	93
Routing	453	3	95
Low (Red)	140	1	100
Middle (Yellow)	228	4	89
High (Blue)	85	4	93

† Not applicable.

[1] Percent agreement was calculated as follows: number of validation items observed in which observer agreed with the assessor/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 third grade data collection, school year 2001–02.

the lowest percent agreement (87 percent). The reading yellow level path received a relatively large number of observations (262) 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 yellow level (the medium science level) also had a relatively higher opportunity for disagreement (228 observations and 4 validation codes) and it, too, exhibited a somewhat lower interrater reliability (89 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. More details on the instruments can be found in the *Early Childhood Longitudinal Study, Kindergarten Class of 1998–99 (ECLS-K) Psychometric Report for the Third Grade* (U.S Department of Education, National Center for Education Statistics, forthcoming [b]).

5.6 Spring-Third Grade Completion Rates

In the sections that follow, spring-third grade completion rates are presented for three groups of students: (1) students sampled in kindergarten, (2) students sampled in first grade through the freshening procedure, and (3) both groups combined. Completion rates were computed with the same procedures used for spring-first grade to allow for comparisons of completion rates between the two 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 assessments or parent interview) and children sampled in spring-first grade as part of sample freshening as described in section 4.3.2.

5.6.1 Students Sampled in Kindergarten

Tables 5-9 to 5-11 present weighted and unweighted child-level completion rates for spring-third grade data collection, broken out by school characteristics. These rates pertain to children who were sampled as part of the kindergarten cohort in the base year. (Rates for students sampled in first grade through the student sample freshening procedure can be found in table 5-19.) 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.

Relative to spring-first grade, the overall completion rates for the child assessments (80.8 percent) and the parent interview (77.8 percent) both decreased 7 percentage points in spring-third grade. The decrease is almost completely due to the increase in the number of children who moved outside of the sampled PSUs or moved within the sampled PSUs but could not be located (the numbers slightly more than doubled in both cases). These children are included in the category labeled “Unknown” for each of the different school characteristics (tables 5-9 and 5-11). The category includes children who were unlocatable as their whereabouts were unknown, and those children who had moved into a nonsampled county. If no information concerning the child’s school was available, they were included in the “Unknown” category. The completion rates for the child assessments are quite high and uniform across school characteristics (ranging from 97.1 percent in schools with 750 or more enrolled to 100 percent in

Table 5-9. Number of completed child-level cases and child-level completion rates among 1998–99 kindergartners in spring-third grade, by survey type and school characteristics: School year 2001–02

School characteristic[1]	Child assessments			Parent interview		
		Completion rate			Completion rate	
	Completes[2]	Weighted	Unweighted	Completes[3]	Weighted	Unweighted
All school types	14,349	80.8	86.1	13,392	77.8	80.3
School type						
Public	11,671	98.1	98.5	10,268	86.3	86.6
Private	2,629	98.7	99.2	2,452	92.6	92.5
Catholic	1,662	99.0	99.3	1,546	92.9	92.4
Other private	967	98.1	99.0	906	92.3	92.7
Unknown school type	49	2.6	2.3	672	36.2	31.0
Type of locale						
Large city	2,431	97.8	98.3	2,038	80.9	82.4
Mid-size city	2,474	98.9	99.3	2,216	88.6	88.9
Urban fringe of large city	4,169	97.4	98.0	3,723	86.9	87.5
Urban fringe of mid-size city	942	97.7	98.0	832	86.2	86.6
Large town	375	100.0	100.0	355	93.4	94.7
Small town	1,032	99.4	99.2	967	90.0	93.0
Rural—outside MSA	1,498	98.8	99.4	1,350	89.7	89.6
Rural—inside MSA	1,201	98.2	98.7	1,075	88.0	88.3
Unknown	227	8.5	9.7	836	39.5	35.6
School size (total enrollment)						
1 to 299	3,078	98.2	98.8	2,809	89.8	90.2
300 to 499	4,562	98.5	98.9	4,118	88.5	89.2
500 to 749	4,043	98.5	98.7	3,569	87.3	87.2
750 or more	2,505	97.1	97.9	2,116	81.4	82.7
Unknown	161	6.2	7.0	780	38.3	34.1

See notes at end of table.

Table 5-9. Number of completed child-level cases and child-level completion rates among 1998–99 kindergartners in spring-third grade, by survey type and school characteristics: School year 2001–02—Continued

School characteristic[1]	Child assessment			Parent interview		
		Completion rate			Completion rate	
	Completes[2]	Weighted	Unweighted	Completes[3]	Weighted	Unweighted
Percent non-White enrolled						
0 – 10	4,580	99.3	99.4	4,269	92.3	92.7
11 – 49	4,594	97.5	98.2	4,179	88.3	89.4
50 – 89	2,564	98.0	98.4	2,197	83.8	84.3
90 – 100	2,412	97.8	98.2	1,932	78.5	78.6
Unknown	199	7.3	8.6	815	38.9	35.1
Region						
Northeast	2,667	98.6	98.8	2,411	88.8	89.3
Midwest	3,677	98.1	98.5	3,405	90.8	91.2
South	4,674	97.6	98.3	4,038	83.5	85.0
West	3,291	98.6	98.8	2,880	87.2	86.5
Unknown	40	2.2	1.9	658	35.8	30.5

[1] Based on ECLS-K survey data and not data from the sampling frame.
[2] Reading, math, or science assessment was scorable or child was disabled.
[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 third grade data collection, school year 2001–02.

Table 5-10. Number of completed child-level cases and child-level completion rates among 1998–99 kindergartners in spring-third grade, by survey type and school characteristics: School year 2001–02

School characteristic[1]	School administrator questionnaire			School fact sheet		
		Completion rate			Completion rate	
	Completes[2]	Weighted	Unweighted	Completes[2]	Weighted	Unweighted
All school types	12,361	66.1	73.3	14,064	76.5	83.4
School type						
Public	9,559	78.4	80.7	10,976	91.1	92.6
Private	2,352	85.8	88.7	2,545	94.6	96.0
Catholic	1,515	88.5	90.5	1,623	95.9	97.0
Other private	837	81.4	85.7	922	92.5	94.4
Unknown school type	450	12.0	19.0	543	14.5	22.9
Type of locale						
Large city	1,832	69.0	74.1	2,246	88.6	90.8
Mid-size city	2,138	82.2	85.8	2,349	93.1	94.3
Urban fringe of large city	3,271	73.9	76.9	3,894	89.7	91.6
Urban fringe of mid-size city	792	78.0	82.4	903	91.0	94.0
Large town	374	99.4	99.7	375	100.0	100.0
Small town	933	88.0	89.7	1,015	97.2	97.6
Rural – outside MSA	1,401	93.0	93.0	1,480	97.7	98.2
Rural – inside MSA	1,053	81.7	86.5	1,132	89.9	93.0
Unknown	567	15.3	22.2	670	17.4	26.3
School size (total enrollment)						
1 to 299	2,883	90.0	92.6	3,024	95.9	97.1
300 to 499	3,934	81.9	85.2	4,328	92.4	93.8
500 to 749	3,234	76.5	79.0	3,807	90.4	93.0
750 or more	1,859	72.8	72.6	2,320	90.2	90.6
Unknown	451	11.6	18.1	585	15.5	23.5

See notes at end of table.

Table 5-10. Number of completed child-level cases and child-level completion rates among 1998–99 kindergartners in spring-third grade, by survey type and school characteristics: School year 2001–02—Continued

School characteristic[1]	School administrator questionnaire			School fact sheet		
		Completion rate			Completion rate	
	Completes[2]	Weighted	Unweighted	Completes[2]	Weighted	Unweighted
Percent non-White enrolled						
0 – 10	4,017	85.2	87.2	4,412	94.6	95.8
11 – 49	3,973	82.1	85.0	4,379	91.1	93.6
50 – 89	1,970	73.1	75.6	2,389	90.6	91.6
90 – 100	1,921	74.0	78.2	2,265	90.9	92.2
Unknown	480	12.2	19.0	619	16.2	24.5
Region						
Northeast	2,035	72.8	75.4	2,460	89.5	91.1
Midwest	3,349	86.9	89.7	3,575	94.4	95.7
South	3,864	79.0	81.3	4,454	91.4	93.7
West	2,668	76.0	80.1	3,039	89.9	91.3
Unknown	445	11.8	18.9	536	14.4	22.8

[1] Based on ECLS-K survey data and 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 third grade data collection, school year 2001–02.

Table 5-11. Number of completed child-level cases and child-level completion rates among 1998–99 kindergartners in spring-third grade, by survey type and school characteristics: School year 2001–02

School characteristic[1]	Teacher questionnaire part A			Teacher questionnaire part B			Teacher questionnaire part C		
		Completion rate			Completion rate			Completion rate	
	Completes[2]	Weighted	Unweighted	Completes[2]	Weighted	Unweighted	Completes[2]	Weighted	Unweighted
All school types	11,770	62.4	69.8	11,741	62.3	69.6	11,802	62.7	70.0
School type									
Public	9,274	75.7	78.3	9,242	75.5	78.0	9,301	76.0	78.5
Private	2,416	87.3	91.1	2,418	87.2	91.2	2,425	87.4	91.5
Catholic	1,557	90.2	93.0	1,564	90.5	93.4	1,575	91.4	94.1
Other private	859	82.7	87.9	854	82.0	87.4	850	81.3	87.0
Unknown school type	80	2.2	3.4	81	2.3	3.4	76	2.2	3.2
Type of locale									
Large city	1,649	62.0	66.7	1,647	62.0	66.6	1,654	62.4	66.9
Mid-size city	2,087	80.0	83.7	2,079	79.5	83.4	2,096	80.2	84.1
Urban fringe of large city	3,322	74.5	78.1	3,295	73.9	77.5	3,329	74.6	78.3
Urban fringe of mid-size city	801	78.1	83.4	804	78.9	83.7	808	79.5	84.1
Large town	363	96.0	96.8	370	98.2	98.7	368	97.7	98.1
Small town	985	92.9	94.7	982	92.8	94.4	977	92.1	93.9
Rural—outside MSA	1,356	85.3	90.0	1,363	85.8	90.4	1,369	86.6	90.8
Rural—inside MSA	1,027	80.1	84.4	1,020	79.4	83.8	1,023	79.5	84.1
Unknown	180	5.4	7.1	181	5.5	7.1	178	5.4	7.0
School size (total enrollment)									
1 to 299	2,832	87.9	90.9	2,817	87.3	90.5	2,829	87.8	90.8
300 to 499	3,911	81.2	84.7	3,920	81.4	84.9	3,924	81.5	85.0
500 to 749	3,181	75.3	77.7	3,164	75.0	77.2	3,199	75.8	78.1
750 or more	1,730	65.1	67.6	1,723	65.0	67.3	1,735	65.6	67.8
Unknown	116	3.2	4.7	117	3.3	4.7	115	3.3	4.6

See notes at end of table.

Table 5-11. Number of completed child-level cases and child-level completion rates among 1998–99 kindergartners in spring-third grade, by survey type and school characteristics: School year 2001–02—Continued

School characteristic[1]	Teacher questionnaire part A			Teacher questionnaire part B			Teacher questionnaire part C		
		Completion rate			Completion rate			Completion rate	
	Completes[2]	Weighted	Unweighted	Completes[2]	Weighted	Unweighted	Completes[2]	Weighted	Unweighted
Percent non-White enrolled									
0 – 10	4,216	88.2	91.5	4,206	87.9	91.3	4,211	88.0	91.4
11 – 49	3,839	77.9	82.1	3,835	77.9	82.0	3,872	78.6	82.8
50 – 89	1,963	73.4	75.3	1,951	73.2	74.8	1,960	73.9	75.2
90 – 100	1,605	62.2	65.3	1,601	62.0	65.2	1,611	62.3	65.6
Unknown	147	4.0	5.8	148	4.0	5.9	148	4.1	5.9
Region									
Northeast	2,168	77.1	80.3	2,155	76.7	79.8	2,170	77.8	80.4
Midwest	3,272	85.0	87.6	3,275	85.1	87.7	3,292	85.3	88.2
South	3,822	76.2	80.4	3,815	75.9	80.3	3,842	76.6	80.8
West	2,433	69.2	73.1	2,420	68.9	72.7	2,427	69.1	72.9
Unknown	75	2.0	3.2	76	2.1	3.2	71	2.0	3.0

[1] Based on ECLS-K survey data and 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 third grade data collection, school year 2001–02.

schools in large towns), except for the “unknown” category. Similarly, the completion rates for the parent interviews were uniform across school characteristics (ranging from 80.9 percent for children in large cities to 93.4 percent for children in large towns), except for the “unknown” category. The “unknown” category aside, both the child assessments and the parent interview completion rates increased between spring-first grade and spring-third 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 66.1 percent for the school administrator questionnaire and 76.5 percent for the school fact sheet. The completion rate for the school fact sheet is about 10 percent higher than that for the school administrator questionnaire due to the continued data collection in fall 2002 that affected the school fact sheet and the student record abstract. The completion rate for the school administrator questionnaire is about 10 percentage points lower than the corresponding rate in spring-first grade. Note that there was no school fact sheet in spring-first grade. Once again, the increase in the movers is largely responsible for the lower rates, as discussed below. The completion rates for the school administrator questionnaire range from 69.0 percent for children in large cities to 99.4 percent for those in large towns (ignoring the “unknown” category). Rates for the school fact sheet follow the same pattern. Excluding the movers, the completion rate for the school administrator questionnaire is much higher, with an overall rate of 86.7 percent, only slightly lower than the spring-first grade rate of 88.7 percent. In the case of the school fact sheet, the rate for nonmovers is 95.8 percent. 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 in the first-grade year and in spring-third grade compared to the base year, reflecting the success of increased data collection efforts targeted toward these schools.

All three of the teacher questionnaires were completed at an overall rate of 62 to 63 percent. The completion rates have substantial variation when broken out by school characteristics, even when the “unknown” category is ignored. The completion rates are 90 percent or more for Catholic schools and for schools in large and small towns. Schools in large cities, schools with 750 students or more, and schools with 90 percent or more minority enrollment have completion rates in the 60s. The “unknown” categories have, by far, the lowest completion rates.

As noted above, the rate at which the survey instruments were completed varies markedly by mover status and within movers, by whether or not the child was located and followed. As shown in table 5-12 the completion rate for the child assessments was 94.5 percent for children still enrolled in their base year school. For movers it dropped by close to 9 points to 85.6 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 84.5 percent for nonmovers to 74.8 percent for movers who were located and followed for the purposes of the child assessments, to 51.2 percent for movers who could either not be located or were not followed for the purposes of the child assessments. Even though children who had moved to a non-ECLS-K PSU were not administered the child assessments, a parent interview was conducted by telephone wherever possible, leading to the 51.2 percent response rate for this category.

Table 5-13 shows that the school administrator questionnaire completion rate is about 30 points lower for movers compared to nonmovers, even when the children who had moved were located and followed. For the school fact sheet, it is about 20 points lower for movers than for nonmovers. 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 86.7 percent for the school administrator questionnaire and 95.8 percent for the school fact sheet. For located and followed movers it was 56.0 and 74.5 percent for the school administrator questionnaire and for the school fact sheet, respectively.

For all three teacher questionnaires the completion rates were approximately 82 percent if the child had not moved; about 54 percent if the child moved, was located, and followed; and just about 0 percent if not located or followed (table 5-14). A handful of children who could not be located but had teacher data was due to the fact that, if they move during the term, and teachers in their old schools already filled out the questionnaires but teachers in the new schools did not, the teacher data from the old schools were attached to these children. The reasons for lower completion rates from teachers if the child moved are similar to the reasons that affected the school administrator questionnaire and school fact sheet completion rates for movers.

Table 5-12. Number of completed child-level cases and child-level completion rates among 1998–99 kindergartners in spring-third grade, by survey type and child's mover status: School year 2001–02

Mover status[1]	Child assessments			Parent interview		
		Completion rate			Completion rate	
	Completes[2]	Weighted	Unweighted	Completes[3]	Weighted	Unweighted
All students	14,349	80.8	86.1	13,392	77.8	80.3
Mover status						
Mover	2,791	60.9	62.6	3,063	68.0	68.8
Located, followed	2,791	85.6	85.1	2,451	74.8	74.7
Other[4]	0	0.0	0.0	612	51.2	52.1
Nonmover	11,558	94.5	94.6	10,329	84.5	84.6

[1] This is the mover status used in weighting, which does not consider children who moved into identified destination schools as movers.
[2] Reading, math or science assessment was scorable or child was disabled.
[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 third grade data collection, school year 2001–02.

Table 5-13. Number of completed child-level cases and child-level completion rates among 1998–99 kindergartners in spring-third grade by survey type and child's mover status: School year 2001–02

Mover status[1]	School administrator questionnaire			School fact sheet		
		Completion rate			Completion rate	
	Completes[2]	Weighted	Unweighted	Completes[2]	Weighted	Unweighted
All students	12,361	66.1	73.3	14,064	76.5	83.4
Mover status						
Mover	1,739	37.2	37.4	2,327	49.6	50.0
Located, followed	1,739	56.0	54.5	2,327	74.5	72.9
Other[3]	0	0.0	0.0	0	0.0	0.0
Nonmover	10,622	86.7	87.0	11,737	95.8	96.1

[1] This is the mover status used in weighting, which does not consider children who moved into identified destination schools as movers.
[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 third grade data collection, school year 2001–02.

Table 5-14. Number of completed child-level cases and child-level completion rates among 1998–99 kindergartners in spring-third grade, by survey type and child's mover status: School year 2001–02

Mover status[1]	Teacher questionnaire part A			Teacher questionnaire part B			Teacher questionnaire part C		
		Completion rate			Completion rate			Completion rate	
	Completes[2]	Weighted	Unweighted	Completes[2]	Weighted	Unweighted	Completes[2]	Weighted	Unweighted
All students	11,770	62.4	69.8	11,741	62.3	69.6	11,802	62.7	70.0
Mover status									
Mover	1,656	35.7	35.6	1,645	35.6	35.3	1,650	35.7	35.5
Located, followed	1,653	53.7	51.8	1,641	53.4	51.4	1,646	53.7	51.6
Other[3]	3	0.1	0.2	4	0.2	0.3	4	0.2	0.3
Nonmover	10,114	81.5	82.8	10,096	81.4	82.7	10,152	81.9	83.1

[1] This is the mover status used in weighting, which does not consider children who moved into identified destination schools as movers.
[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 third grade data collection, school year 2001–02.

Tables 5-15 to 5-17 present child-level completion rates for the spring-third grade data collection, this time broken out by child characteristics for children who were sampled as part of the kindergarten cohort in the base year. When the “unknown” categories are not included, the differences in completion rates by sex and by year of birth are very small, but for race and ethnicity they are more substantial. For the child assessments the completion rate was highest for Asians and Pacific Islanders (84.1 percent and 84.9 percent, respectively) and lowest for American Indians or Alaska Natives (75.5 percent). For the parent interview it was highest for Whites (82.9 percent), and lowest for Blacks (67.0 percent) and Asian students (68.6 percent). The ECLS-K sample of Pacific Islanders is very clustered and has unusually high completion rates for the instruments filled out by school personnel, 80.9 percent for the school administrator questionnaire, 84.6 percent for the school fact sheet, and about 74 percent for each of the teacher questionnaires. The lowest completion rate for the school administrator questionnaire is for Blacks (57.1 percent); for the school fact sheet it is for American Indians or Alaska Natives (66.0 percent). For the teacher questionnaires the lowest rates are in the 53 to 55 percent range and are associated with Blacks, Hispanics, and American Indians or Alaska Natives. Since 60 percent of the Black and Hispanic students fielded in spring-third grade are enrolled in high minority schools (50 percent or higher), this may be associated with lower levels of response for the school administrator questionnaire from high minority schools. Of the 32 percent of Black and Hispanic students with no school administrator questionnaire data, roughly half are enrolled in high minority schools.

In addition to the child assessments, parent interview, teacher questionnaires, school administrator questionnaires and school fact sheets whose completion rates have been summarized in the preceding tables, various other types of data were collected during spring-third grade as well. Table 5-18 presents counts of completes and weighted and unweighted completion rates at the overall student level for these other data collection efforts. The facilities checklist has a 77.5 percent completion rate, which is about 11 points higher than that for the school administrator questionnaire but only 1 point higher than the rate for the school fact sheet, the two other school-level survey instruments. The student record abstract, which was to have been completed for all students except for those who moved and could not be located, has a 67.0 percent completion rate. Data were also collected during spring-third grade from the special education teachers for children in special education programs (fewer than 1,200). The completion rates for these instruments are higher than for the regular teacher questionnaires, 73.0 percent for part A, which captures teacher information, and 72.8 percent for part B, which relates to individual students who receive special education services.

Table 5-15. Number of completed child-level cases and child-level completion rates among 1998–99 kindergartners in spring-third grade, by survey type and selected child characteristics: School year 2001–02

Child characteristic[1]	Child assessments			Parent interview		
		Completion rate			Completion rate	
	Completes[2]	Weighted	Unweighted	Completes[3]	Weighted	Unweighted
All students	14,349	80.8	86.1	13,392	77.8	80.3
Sex						
Male	7,285	80.3	85.7	6,858	78.3	80.7
Female	7,064	81.5	86.6	6,534	77.3	80.1
Unknown sex	0	0.0	0.0	0	0.0	0.0
Race/ethnicity						
White (not Hispanic)	8,119	81.2	86.9	8,000	82.9	85.6
Black (not Hispanic)	1,872	78.0	83.6	1,570	67.0	70.2
Hispanic	2,575	82.2	85.9	2,295	73.9	76.6
Asian	963	84.1	86.4	757	68.6	67.9
Pacific Islander	171	84.9	87.2	161	82.3	82.1
American Indian or Alaska Native	248	75.5	81.3	242	75.5	79.3
Other	382	81.7	88.4	360	78.8	83.3
Unknown race/ethnicity	19	42.9	48.7	7	7.8	17.9
Year of birth						
1992	4,179	80.5	85.7	3,922	77.3	80.5
1993	10,086	81.1	86.4	9,397	78.2	80.5
Other/unknown	84	59.6	71.2	73	53.3	61.9

[1] Based on ECLS-K survey data and not on data from the sampling frame.
[2] Reading, math, or science assessment was scorable or child was disabled.
[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 third grade data collection, school year 2001–02.

Table 5-16. Number of completed child-level cases and child-level completion rates among 1998–99 kindergartners in spring-third grade, by survey type and selected child characteristics: School year 2001–02

Child characteristic[1]	School administrator questionnaire			School fact sheet		
		Completion rate			Completion rate	
	Completes[2]	Weighted	Unweighted	Completes[2]	Weighted	Unweighted
All students	12,361	66.1	73.3	14,064	76.5	83.4
Sex						
Male	6,241	65.0	72.6	7,119	75.7	82.9
Female	6,108	67.2	73.9	6,933	77.3	83.9
Unknown sex	12	100.0	100.0	12	100.0	100.0
Race/ethnicity						
White (not Hispanic)	7,309	70.8	78.2	8,061	79.1	86.2
Black (not Hispanic)	1,489	57.1	64.6	1,810	70.5	78.5
Hispanic	2,062	61.1	66.5	2,484	75.2	80.1
Asian	786	64.9	69.6	920	78.1	81.4
Pacific Islander	172	80.9	86.9	178	84.6	89.9
American Indian or Alaska Native	212	61.9	67.1	221	66.0	69.9
Other	305	60.4	70.3	358	72.4	82.5
Unknown race/ethnicity	26	68.2	66.7	32	83.7	82.1
Year of birth						
1992	3,681	68.1	74.8	4,115	76.9	83.6
1993	8,602	65.4	72.7	9,856	76.4	83.3
Other/unknown	78	54.5	65.0	93	69.4	77.5

[1] Based on ECLS-K survey data and not on data from the sampling frame.
[2] A completed questionnaire was defined as one that was not completely left blank.
SOURCE: U.S. Department of Education, National Center for Education Statistics, Early Childhood Longitudinal Study, Kindergarten Class of 1998–99 third grade data collection, school year 2001–02.

Table 5-17. Number of completed child-level cases and child-level completion rates among 1998–99 kindergartners in spring-third grade, by survey type and selected child characteristics: School year 2001–02

Child characteristic[1]	Teacher questionnaire part A			Teacher questionnaire part B			Teacher questionnaire part C		
		Completion rate			Completion rate			Completion rate	
	Completes[2]	Weighted	Unweighted	Completes[2]	Weighted	Unweighted	Completes[2]	Weighted	Unweighted
All students	11,770	62.4	69.8	11,741	62.3	69.6	11,802	62.7	70.0
Sex									
Male	5,930	61.4	69.0	5,921	61.4	68.9	5,951	61.7	69.3
Female	5,840	63.7	70.7	5,820	63.4	70.4	5,851	63.9	70.8
Unknown sex	0	0.0	0.0	0	0.0	0.0	0	0.0	0.0
Race/ethnicity									
White (not Hispanic)	7,174	68.6	76.8	7,158	68.4	76.6	7,186	68.8	76.9
Black (not Hispanic)	1,405	53.7	61.0	1,402	53.6	60.8	1,409	54.0	61.1
Hispanic	1,819	54.7	58.7	1,820	54.7	58.7	1,837	55.2	59.3
Asian	726	59.8	64.2	714	58.7	63.2	720	59.1	63.7
Pacific Islander	158	74.2	79.8	157	73.8	79.3	157	73.8	79.3
American Indian or Alaska Native	189	53.3	59.8	189	53.3	59.8	191	53.1	60.4
Other	288	54.3	66.4	289	55.0	66.6	289	55.1	66.6
Unknown race/ethnicity	11	21.8	28.2	12	27.7	30.8	13	29.0	33.3
Year of birth									
1992	3,491	63.3	71.0	3,488	63.2	70.9	3,497	63.5	71.1
1993	8,215	62.3	69.4	8,188	62.2	69.2	8,241	62.6	69.7
Other/unknown	64	38.6	53.3	65	40.2	54.2	64	39.6	53.3

[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.
SOURCE: U.S. Department of Education, National Center for Education Statistics, Early Childhood Longitudinal Study, Kindergarten Class of 1998–99 third grade data collection, School year 2001–02.

Table 5-18. Number of completed instruments and child-level completion rates among 1998–99 kindergartners for additional data collected in spring-third grade, by survey type: School year 2001–02

Survey instrument	Completes	Completion rate	
		Weighted	Unweighted
Facilities checklist[1]	14,280	77.5	84.7
Student record abstract[1]	12,359	67.0	73.3
Special education part A[1]	875	73.0	75.0
Special education part B[1]	870	72.8	74.6

[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 third grade data collection, school year 2001–02.

5.6.2 Students 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 did not attend kindergarten in the United States or were 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 students was followed into spring-third grade. 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 students in spring-third grade (second component). The first component alone can further be decomposed into two sources: attrition due to entire schools refusing to implement the freshening procedure (*the school term*), and attrition due to ECLS-K sampled children moving to other schools (*the child term*). To contain costs, students who transferred from schools targeted for freshening were not used as links to identify freshened students, even when they were otherwise followed for data collection. These movers were considered freshening nonrespondents in the *child term*.

Table 5-19 presents weighted completion rates for freshened students. The two components of the completion rates are presented separately in table 5-19. The actual completion rates are the products of the two components. Since no freshening was done in the third grade, the first component is identical to that for spring-first grade. It 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

Table 5-19. Number of completed child-level cases and child-level completion rates among freshened 1999–2000 first graders in spring-third grade, by freshening component and survey type: School year 2001–02

Survey instrument	Completes	Completion rate[1]	
		Weighted	Unweighted
First component of completion rate	7,135	64.3	77.9
School term[2]	7,192	65.4	78.5
Child term[3]	7,135	98.3	99.2
Second component of completion rate			
Child assessments[4]	121	78.3	76.1
Parent interview[5]	97	63.7	61.0
Teacher questionnaire part A[6]	86	53.1	54.1
Teacher questionnaire part B[6]	85	52.2	53.5
Teacher questionnaire part C[6]	82	50.9	51.6
Special education part A[6]	12	67.6	70.6
Special education part B[6]	13	73.3	76.5
School administrator questionnaire[6]	102	65.8	64.2
School fact sheet[6]	124	78.9	78.0
Facilities checklist[6]	126	80.3	79.2
Student records abstract[6]	103	68.7	64.8
Overall completion rate			
Child assessments[4]	121	50.4	59.3
Parent interview[5]	97	41.0	47.5
Teacher questionnaire part A[6]	86	34.2	42.1
Teacher questionnaire part B[6]	85	33.6	41.6
Teacher questionnaire part C[6]	82	32.7	40.2
Special education part A[6]	12	43.5	55.0
Special education part B[6]	13	47.1	59.6
School administrator questionnaire[6]	102	42.3	50.0
School fact sheet[6]	124	50.8	60.7
Facilities checklist[6]	126	51.7	61.7
Student records abstract[6]	103	44.2	50.5

[1] The first component is the completion rate for sample freshening in first grade. The second component 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 school term is the completion rate for schools targeted for freshening.
[3] The child term is the completion rate for children in schools that agreed to the freshening procedure.
[4] Reading, math, or science assessment was scorable or child was disabled.
[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 third grade data collection, school year 2001–02.

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 65.4 percent. The reasons that schools did not participate in the freshening process included refusing or being unable to provide the requested information in order to complete the procedures. Within the schools that agreed to freshen, the freshening completion rate is 98.3 percent, the slight loss due to students who transferred to other schools (*the child term*). Multiplying these two terms together gives a first component completion rate of 64.3 percent. Note that the first component rate for spring-third grade is identical to the first component rate for spring-first grade.

The second component varies by survey instrument. The rates for the paper-and-pencil instruments range from 50.9 percent for the child-level teacher questionnaire to 80.3 percent for the facilities checklist and are uniformly lower than for the kindergarten sample. The child assessments at 78.3 percent are 3 points lower than for the kindergarten sample and the parent interview, at 63.7 percent, is 14 points lower. 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 51.7 percent for the facilities checklist to a low of 32.7 percent for the child-level teacher questionnaire.

5.6.3 Spring-Third Grade Completion Rates—All Children

Table 5-20 presents final spring-third grade completion rates for children sampled in kindergarten, children sampled in first grade, and all children combined. Because children sampled in first grade represent such a small fraction of the total population of children, their inclusion in the computation of the completion rate brings down the rates for all children by less than one percent relative to the children sampled in kindergarten rates, even though the completion rates for children sampled in first grade are lower than the kindergarten rates.

Table 5-20. Number of completed child-level cases and child-level completion rates in spring-third grade, by sampling timeframe and survey type: School year 2001–02

Survey instrument	Children sampled in kindergarten			Children sampled in first grade			All children		
		Completion rate			Completion rate			Completion rate	
	Completes	Weighted	Unweighted	Completes	Weighted	Unweighted	Completes	Weighted	Unweighted
Child assessments[1]	14,349	80.8	86.1	121	50.4	59.3	14,470	80.1	85.9
Parent interview[2]	13,392	77.8	80.3	97	41.0	47.5	13,489	76.9	80.1
Teacher questionnaire part A[3]	11,770	62.4	69.8	86	34.2	42.1	11,856	61.7	69.6
Teacher questionnaire part B[3]	11,741	62.3	69.6	85	33.6	41.6	11,826	61.6	69.4
Teacher questionnaire part C[3]	11,802	62.7	70.0	82	32.7	40.2	11,884	62.0	69.7
Special education part A[3]	875	73.0	75.0	12	43.5	55.0	887	72.3	74.8
Special education part B[3]	870	72.8	74.6	13	47.1	59.6	883	72.2	74.5
School administrator questionnaire[3]	12,361	66.1	73.3	102	42.3	50.0	12,463	65.5	73.1
School fact sheet[3]	14,064	76.5	83.4	124	50.8	60.7	14,188	75.9	83.2
Facilities checklist[3]	14,280	77.5	84.7	126	51.7	61.7	14,406	76.9	84.5
Student records abstract[3]	12,359	67.0	73.3	103	44.2	50.5	12,462	66.5	73.1

[1] Reading, math, or science assessment was scorable or child was disabled.
[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 third grade data collection, school year 2001–02.

5.6.4 Spring-Third Grade Completion Rates Conditioned on Child Assessments

Table 5-21 shows the completion rates for the child assessments, the parent interviews, and the school and teacher instruments for children who have nonzero child weights (C5CW0>0). These are children whose spring-third grade reading, math, or science assessments were scorable, or children who could not be assessed because of disabilities. For these children, the completion rate for the child assessments should be 100 percent. The less than 100 percent rate shown when children sampled in kindergarten are combined with children sampled in first grade is due to the school freshening nonresponse for children sampled in first grade.

When the completion rates are conditioned on the presence of the child weight (i.e., completion rates for children with C5CW0>0), they are at least 9 points higher than the unconditioned completion rates for all instruments but the special education questionnaires. For these last two instruments, the difference between the number of completes for the conditioned and unconditioned rates is very small; hence the conditioned rates are not affected as much as for the other instruments. For all the other instruments, the conditioned completion rates are higher by 9 points for the parent interviews and as high as 17 points for the student record abstract. These numbers are the differences between the unconditioned rates (table 5-20) and conditioned rates (table 5-21), hence not shown in any table.

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 are also different by school and child characters. A separate report examines the potential for bias in estimates produced from the ECLS-K third grade data. Since analysis of the third grade data is conditioned on the base year—only base year respondents were included in the collection of first grade and third grade data—the analysis of nonresponse bias is built on the base year nonresponse bias analysis (see *Analysis of Nonresponse Bias in the Base Year Early Childhood Longitudinal Survey, Kindergarten Class of 1998–99*, [U.S. Department of Education, National Center for Education Statistics, forthcoming (a)]).

Table 5-21. Number of completed child-level cases and child-level completion rates in spring-third grade that includes children with scorable reading, math, or science assessments or children not assessed due to disabilities, by sampling timeframe and survey type: School year 2001–02

Survey instrument	Children sampled in kindergarten			Children sampled in first grade			All children		
		Completion rate			Completion rate			Completion rate	
	Completes	Weighted	Unweighted	Completes	Weighted	Unweighted	Completes	Weighted	Unweighted
Child assessments[1]	14,349	100.0	100.0	121	64.3	77.9	14,470	99.2	99.8
Parent interview[2]	12,564	86.8	87.6	90	49.7	57.9	12,654	85.9	87.4
Teacher questionnaire part A[3]	11,644	78.0	81.4	86	43.7	55.4	11,730	77.2	81.2
Teacher questionnaire part B[3]	11,614	77.8	81.2	85	43.0	54.7	11,699	77.0	81.0
Teacher questionnaire part C[3]	11,684	78.4	81.7	82	41.9	52.8	11,766	77.5	81.5
Special education part A[3]	854	74.1	75.8	12	43.5	55.0	866	73.4	75.6
Special education part B[3]	847	73.7	75.2	13	47.1	59.6	860	73.1	75.1
School administrator questionnaire[3]	11,846	80.1	82.8	93	50.6	59.9	11,939	79.4	82.6
School fact sheet[3]	13,447	92.5	94.0	113	60.4	72.7	13,560	91.7	93.8
Facilities checklist[3]	13,670	94.0	95.5	116	62.0	74.7	13,786	93.3	95.3
Student records abstract[3]	12,282	84.2	85.8	103	56.5	66.3	12,385	83.6	85.6

[1] Reading, math, or science assessment was scorable or child was disabled.
[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 third grade data collection, school year 2001–02.

REFERENCES

U.S. Department of Education, National Center for Education Statistics (forthcoming [a]). *Analysis of Nonresponse Bias in the Base Year Early Childhood Longitudinal Study, Kindergarten Class of 1998–99 (ECLS-K).* By Mike Brick, John Burke, and Thanh Lê. Washington, DC: U.S. Department of Education.

U.S. Department of Education, National Center for Education Statistics (forthcoming [b]). *Early Childhood Longitudinal Study, Kindergarten Class of 1998–99 (ECLS-K), Psychometric Report for the Third Grade.* Washington, DC: U.S. Department of Education.

U.S. Department of Education, National Center for Education Statistics (forthcoming [c]). *Early Childhood Longitudinal Study, Kindergarten Class of 1998–99 (ECLS-K), Third Grade Methodology Report.* Washington, DC: U.S. Department of Education.

This page intentionally left blank.

6. DATA PREPARATION

As described in chapter 5, two types of data collection instruments were used for the Early Childhood Longitudinal Study, Kindergarten Class of 1998–99 (ECLS-K) data collection in the spring-third grade: computer-assisted and self-administered paper forms (hard copy). The data preparation approach differed with the mode of data collection. The direct child assessments and parent interview were conducted using computer-assisted interviewing (CAI) techniques. Editing specifications were built into the computer programs used by assessors to collect these data. The teacher and school administrator forms were self-administered. When the field supervisors returned these forms, coders recorded the receipt of these forms into a project-specific forms tracking system. Coders reviewed the questionnaires to ensure data readability for transfer into an electronic format. The visual review included changing (upcoding) any "Other, specify" responses that actually fit within the available response categories of the question. There were some items for which upcoding was conducted after the data were keyed due to the large volume of "Other" responses. Once they finished this review, the coders sent the instruments to data entry to be manually transferred to an electronic format and reviewed for range and logic consistency. The following sections describe the data preparation activities for both modes of data collection in more detail.

6.1 Coding and Editing Specifications for Computer-Assisted Interviews (CAI)

The very nature of designing a computer-assisted interview forces decisions about edit specifications to be made up front. Both acceptable ranges and logic consistency checks were preprogrammed into the electronic questionnaire. The next few sections describe the coding and editing of the data collected using CAI. Though the child assessments and the parent interviews were both collected using CAI, the child assessments did not contain some of the additional range and edit checks contained in the parent interview. The following sections describe the coding and editing that were conducted on the CAI parent interview.

6.1.1 Range Specifications

Within the CAI parent interview instruments, respondent answers were subjected to both “hard” and “soft” range edits during the interviewing process. (The child assessment did not have such hard and soft ranges.) A “soft range” is one that represents the reasonable expected range of values but does not include all possible values. Responses outside the soft range were confirmed with the respondent and entered a second time. For example, the number of hours each week a child attended a day care center on a regular basis had a soft range of 1 to 50. A value outside this range could be entered and confirmed as correct by the assessor as long as it was within the hard range of values (1 to 70).

“Hard ranges” are those that have a finite set of parameters for the values that can be entered into the computer, for example, “0-5 times” for the number of times the child, in the previous 5 days, ate a breakfast that was not school provided. Out-of-range values for closed-ended questions were not accepted. If the respondent insisted that a response outside the hard range was correct, the assessor could enter the information in a comments data file. Data preparation and project staff reviewed these comments. Out-of-range values were accepted and entered into the data file if the comments supported the response.

6.1.2 Consistency Checks (Logical Edits)

Consistency checks, or logical edits, examine the relationship between responses to ensure that they do not conflict with one another or that the response to one item does not make the response to another item unlikely. For example, in the household roster, one could not be recorded as a mother and male. When a logical error such as this occurred during a session, the interviewer saw a message requesting verification of the last response and a resolution of the discrepancy. In some instances, if the verified response still resulted in a logical error, the assessor recorded the problem either in a comment or on a problem report. Consistency checks were not applicable to the child assessments.

6.1.3 Coding

Additional coding was required for some of the items collected in the CAI instruments. These items included “Other, specify” text responses, occupation, and race/ethnicity. Interviewers keyed verbatim responses to these items. Once the data were keyed, staff were trained to code these data using coding manuals designed by Westat and the National Center for Education Statistics (NCES) to support the coding process. In this section, we describe the coding activities for the CAI instruments.

6.1.3.1 Review of "Other, specify" Items

The "Other, specify" open-ended parent interview responses were reviewed to determine if they should be coded into one of the existing response categories. During data collection, when a respondent selected an "other" response in the parent interview, the interviewer entered the text into a "specify" overlay that appeared on the screen. The data preparation staff reviewed these text "specify" responses and, where appropriate, coded them into one of the existing response categories. There were no "Other, specify" items in the child assessments.

6.1.3.2 Parent Occupation Coding

As in the base year and first grade data collections, occupations were coded using the *Manual for Coding Industries and Occupations* (U.S. Department of Education, 1999). This coding manual was created for the Adult Education Survey of the National Household Education Surveys Program (AE-NHES: 1999) and used an aggregated version of industry and occupation codes. The industry and occupation codes used by NHES were originally developed for the 1989–90 National Postsecondary Student Aid Study (NPSAS: 90) and contained one to four digits. Analysis of the NPSAS categories revealed that some categories had very small numbers of cases and some categories that are similar had similar participation rates, suggesting that the separate codes could be collapsed without significant loss of information. The NHES industry and occupation code categories use a two-digit code, the highest level of aggregation, to have sufficient numbers of cases to support analysis without collapsing categories. There are 13 industry codes and 22 occupation codes in the NHES coding scheme. If an industry or occupation could not be coded using this manual, the *Index of Industries and Occupations—1980* (U.S. Department of Commerce, 1982) and *Standard Occupational Classification Manual—1980* (U.S. Department of Commerce, 1980) were used. Both of these manuals use an expanded coding system and at the same time are directly related to the much more condensed NHES coding scheme. These manuals were used for reference in cases where the NHES coding scheme did not adequately cover a particular situation. (See chapter 7, section 7.4.7 for an expanded description of the industry and occupation codes.)

Occupation coding began with an autocoding procedure using a computer string match program developed for the NHES. The program searched the responses for strings of text for each record/case and assigned an appropriate code. A little over half the cases were autocoded (50.4 percent).

Cases that could not be coded using the autocoding system were coded manually using a customized coding utility program designed for coding occupations. The customized coding utility program brought up each case for coders to assign the most appropriate codes. In addition to the text strings, other information, such as main duties, highest level of education, and name of the employer, was available for the coders. The coders used this information to ensure that the occupation code assigned to each case was appropriate. Almost half the cases (49.6 percent) were manually coded.

All of the cases were then verified. Verification of coding is an important tool for ensuring quality control and extending coder training. As a verification step, two coders independently assigned codes (i.e., double-blind coding) to industry and occupation cases. A coding supervisor arbitrated disagreements between the initial code and the verification code. In the early stages, 100 percent of each coder’s work was reviewed. Once the coder’s error rate had dropped to 1 percent or less, 10 percent of the coder’s work was reviewed. Almost 20 percent (19.9 percent) of the cases that were autocoded required adjudication because the verifier disagreed with the autocoding. About the same percent (21.2 percent) of the cases that were manually coded required adjudication because the manual coder and the verifier disagreed.

6.1.3.3 Race/Ethnicity Coding

The same coding rules used in the base year were used to code all race/ethnicity variables for children, resident parents, and nonresident parents. See chapter 7, section 7.4 for details on how the race variables were coded and how the race/ethnicity composite was created.

6.1.3.4 Partially Complete Parent Interviews

A “completed” parent instrument was defined by whether the section on family structure (FSQ) was completed by the respondent. Only completed interviews were retained in the final data file. A small number of interviews in each wave, approximately 103 (less than 1 percent) in third grade terminated the parent interview after the FSQ section but before the end of the instrument. These interviews were considered as “partially complete” cases and were included in the data file. All instrument items after the interview termination point were set to -9 for “not ascertained.”

6.1.3.5 Household Roster in the Parent Interview

Several tests were run on the household roster to look for missing or inaccurate information. These tests are the same tests run on the first grade files. One flag was used to identify cases that were edited for any of the reasons described below. The flag is P5EDIT; the flag was set to 1 if the case was edited in the given wave. There were 644 cases requiring edits in wave 5.

There were essentially three general types of roster tests performed to determine which cases required editing. First, the relationship of an individual to the focal child was compared to the individual’s listed age and sex. Problems found were corrected on the basis of data from prior data collections wherever possible. Second, households with more than one mother or more than one father were scrutinized for errors. While it is possible to have more than one mother in a household—for example, a household could contain one biological and one foster mother of the focal child—such cases warranted closer inspection. Corrections were made whenever clear errors and a clear resolution existed. Lastly, the relationship of an individual to both the focal child and the reference person was examined, as there were cases in which the relationship of an individual to the focal child conflicted with his status as the spouse/partner of the reference person. For example, in a household containing a child’s grandparents but not his or her parents, the grandmother may be designated the “mother” figure, and the grandfather thus becomes the “father” (for the purposes of some questions in the interview) by virtue of his marriage to the grandmother. These cases were examined but left unchanged. Both the original—and correct (grandfather)—relationship data and the new “parent-figure” designation (father) that had been constructed were kept. In the third grade data, there are 164 cases with these types of errors after the roster tests were run; the cases can be identified by the flag “P5ERRFLG.”

6.2 Coding and Editing Specifications for Hard-Copy Questionnaires

6.2.1 Receipt Control

In order to monitor the more than 50,000 documents that were to be received in the third grade year, the project-specific receipt and document control system developed in the base year was used, with some modifications. The receipt and document control system was initially loaded with the identifying information, such as identification numbers for schools, teachers, and children; the links between teachers and children; and the questionnaires that were expected from each school and teacher for each cooperating school in the sample. As data were collected in the field, field supervisors completed transmittal forms for each school to indicate which questionnaires were being mailed to the home office.

Once data collection started, receipt control clerks reviewed the questionnaires sent in from the field for accuracy and completeness. The identification number on each form was matched against the identification numbers in the tracking system to verify that the appropriate number of forms for each school was returned.

When the clerks verified that the correct questionnaires were returned, the questionnaires were scanned for missing critical items. Critical items were identified for each hard-copy questionnaire, except for the facilities checklist. Questionnaires with incomplete or missing data for critical items were not considered complete, and were processed for return to the field. Clerks completed scan edit sheets listing the missing critical items. The questionnaires were receipted in the system as “needs data retrieval” and were forwarded to the data preparation department for coding, data entry, and editing. Using the scan edit sheets, clerks identified the missing or incomplete items on data retrieval forms that were then sent to the appropriate field supervisor. The supervisor was instructed to contact the school to try to obtain the missing information.

Questionnaires that were scanned and deemed complete were logged into the receipt and document control system as “complete, no data retrieval.” Once forms were logged in, if they had any data (some forms had no data due to refusal by the respondent to complete them), they were then coded. The data were then keypunched into electronic format, after which the data were edited.

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

6.2.2 Coding

The hard-copy questionnaires required coding of race/ethnicity for teachers, review of “Other, specify” text responses, and a quick visual review of particular questions in each questionnaire. The quick visual review was to ensure that the questionnaire values were accurate, complete, and consistent across variables and that the numbers were converted to the appropriate unit of measurement prior to converting data to an electronic format. The coding staff were trained on the coding procedures and had coding manuals to support the coding process. This staff also did the data editing after data entry was complete. Senior coders verified coding. The verification rate was set at 100 percent for each coder until accuracy of less than 1 percent error rate was established. After that point, work was reviewed at a rate of 10 percent.

6.2.2.1 Review of “Other, specify” Items

The “Other, specify” text responses were reviewed by the data editing staff and, where appropriate, upcoded into one of the existing response categories. The small number of specify responses that remained after upcoding did not fit into any preexisting category.

6.2.2.2 Coding Teacher Race/Ethnicity

“Other, specify” text responses for race/ethnicity in the teacher questionnaire part B were coded using the base year and first grade procedures. Many of these “others” included more than one response (e.g., African American/Asian or American Indian/White). The open responses were coded into one or more of the following seven categories: one Hispanic category; White, non-Hispanic; Black or African American, non-Hispanic; American Indian or Alaska Native; Asian; Native Hawaiian or other Pacific Islander; and one unspecified multirace-ethnicity category.

6.2.2.3 Coding Teacher Language

"Other, specify" text responses for language in the teacher questionnaire part A were coded using the base year and first grade procedures. Languages beyond the options provided were recorded in "Other, specify." Groups of languages were created based on geographic boundaries. Additional languages included African language; Eastern European language; Native American language; sign language; Middle Eastern language; Western European language; Indian subcontinent language; Southeast Asian language; Pacific Islander language; and other language.

6.2.3 Data Entry

Westat data entry staff keyed the forms in each batch. The data were rekeyed by more senior data entry operators at a rate of 100 percent to verify the data entry. The results of the two data entry passes were compared and differences identified. The hard-copy form was pulled and examined to determine what corrections had to be made to the keyed data. These corrections were rekeyed, resulting in an accuracy rate exceeding 99 percent. The verified batches were then transmitted electronically to Westat's computer system for data editing.

6.2.4 Data Editing

The data editing process consisted of running range edits for soft and hard ranges, running consistency edits, and reviewing frequencies of the results.

6.2.4.1 Range Specifications

Hard-copy range specifications set the parameters for high and low acceptable values for a question. Where values were printed on the forms, these were used as the range parameters. For open-ended questions, such as, "Counting this school year, how many years have you taught in your **current school** including part-time teaching?," high and low ranges were established as acceptable values. Data frequencies were run on the range of values to identify any errors. Values outside the range were identified as errors and were printed on hard copy for a data editor to review. Cases identified with range errors were identified, and the original response was updated. In some cases, range violations were retained in the data because the value was checked and found to be the value reported by the teacher or school. These were marked as “keep as is” cases. Data frequencies were then rerun and reviewed. This iterative process was repeated until no further range errors were found.

6.2.4.2 Consistency Checks (Logical Edits)

By programming logical edits between variables, consistency between variables not involved in a skip pattern was confirmed. For example, in the school administrator questionnaire, the number of children eligible for free breakfast could not exceed the total number of children enrolled in the school. These logical edits were run on the whole database after all data entry and range edits were complete. The logical edits were run separately for each form. All batches of data were combined into one large data file, and data frequencies were produced. The frequencies were reviewed to ensure the data remained logically consistent within the form. When an inconsistency was found, the case was identified and the inconsistency was printed on paper for an editor to review. The original value was corrected (or checked and “kept as is” if the date had been reported, and the case was then rerun through the consistency edits. Once the case passed the consistency edits, it was appended back into the main data set. The frequencies were then rerun and reviewed. This was an iterative process; it was repeated until no further inconsistencies were found.

6.2.4.3 Frequency and Cross-Tabulation Review

Frequencies and cross-tabulations were run to determine consistency and accuracy across the various forms and matched against the data in the field management system. If discrepancies could not be explained, no changes were made to the data. For example, in teacher questionnaire part A, an item asking about languages other than English spoken in the classroom included a response option of “No language other than English.” If a respondent circled that response, but also answered (in subsequent items) that other languages besides English were spoken in the classroom, then the response was left as recorded by the respondent because the discrepancy could not be resolved.

REFERENCES

U.S. Department of Commence, Bureau of the Census (1982). *1980 Census of Population—Index of Industries and Occupations, Final Edition.* Washington, D.C.: Superintendent of Documents, U.S. Government Printing Office.

U.S. Department of Commerce, Office of Federal Statistical Policy and Planning (1980). *Standard Occupational Classification Manual.* Washington, D.C.: Superintendent of Documents, U.S. Government Printing Office.

U.S. Department of Education, National Center for Education Statistics (1999). *Manual for Coding Industries and Occupations.* Appendix to Data File User's Manual, Volume IV, of the National Household Education Survey of 1999 (NCES 2000–077). Washington, D.C.: U.S. Department of Education.

7. DATA FILE CONTENT AND COMPOSITE VARIABLES

This chapter describes the content of the Early Childhood Longitudinal Study, Kindergarten Class of 1998–99 (ECLS-K) Third Grade Public-Use Data File and focuses largely on the composite variables that have been created. The third grade data file can be used for longitudinal analysis in combination with the files from the base year (kindergarten year) and first grade (see chapter 9 for details about longitudinal analyses). For reference, the Base Year and First Grade User's Manuals are included in appendix C of the third grade Electronic Code Book (ECB).

There is one child-level third grade data file or catalog, as noted in chapter 1. Each child record contains data from the various respondents associated with the child (the child herself/himself, a parent, one or more teachers, and a school administrator), as well as from the facilities checklist, school records, and the Field Management System (FMS).

The third grade child catalog contains one record for each of the 15,305 participating students in spring-third grade. Included in the file are cases with a child assessment, a parent interview, or both. Third grade school- and teacher-level data, including composites, are also stored in the child catalog. The file, named child3p.dat, is stored in the root directory of the CD-ROM as an ASCII file. However, it is strongly recommended that users access the data using the ECB software available on the CD-ROM rather than access the ASCII file directly. Appendix B on the CD-ROM contains the record layout for the child catalog.

This chapter is divided into eight sections. Sections 7.1 through 7.3 focus on the conventions used in the study and describe identification variables, missing values, and variable names. Section 7.4 provides details about the creation of composite variables on the third grade data file. Section 7.5 focuses on the methodological variables. Section 7.6 discusses variables used to identify children who changed schools. Section 7.7 contains a table of the composite variables. Finally, section 7.8 describes masked variables.

7.1 Identification Variables

The third grade data file contains a child identification (ID) variable (CHILDID) that uniquely identifies each record. Teachers on the child records are identified with the ID variable T5_ID; schools are identified by the ID variable S5_ID (spring-third grade). The ID variable S5_ID indicates the school the child attended at the end of the spring-third grade data collection. Other identification variables indicate whether the child moved within spring-third grade and the school or teacher from which his/her questionnaire data came. Section 7.6 provides further details on school identifiers.

Each type of respondent (child, parent, regular teacher, special education teacher, and school) has a unique ID number. The original school ID number (S_ID) is the base for all the subsequent ID numbers as children, parents, and teachers were sampled from schools during the base year. The school ID number is a four-digit number assigned sequentially to sampled schools. The number has a series of ranges: 0001-1299 for originally sampled schools; 2000 series for new schools added to the sample during the first grade sample freshening process; 3000 series for substitute schools that replaced nonresponding original sample schools; and 4000 through 6000 series for transfer schools, which were assigned during processing at the home office. (See chapter 4 for a complete description of the ECLS-K sample.) There is also a 9000 series of S_ID numbers that refers to children who do not attend regular school because they are schooled at home (S_ID numbers 9101 through 9499). There are also several specific 9000 series codes for children who were not located or not followed at the end of a round. The school ID numbers start with 999 for these cases. These are described in section 7.6.

The child ID number (CHILDID) is a concatenation of the school ID where the child was sampled, a three-digit student number, and the letter “C.” For example, 0001010C is the ID number of the tenth child sampled in school 0001. The teacher ID number (T5_ID) is a concatenation of the school ID where the teacher was sampled, the letter “T,” and a two-digit teacher number. In previous rounds of the study, the numbering for the two-digit teacher number started with 01, such that 0001T01 was the ID number for the first teacher sampled in school 0001. In spring-third grade, the two-digit teacher numbers started numbering with T41 so that the teachers from this round of the study could be identified easily. Thus, in spring-third grade 0002T41 is the ID number for the first teacher sampled in school 0002. The parent ID number (PARENTID) is linked to the child ID number and is a concatenation of the four-digit school ID, the three-digit student number, and the letter “P.” It is the same number as the child ID with a letter “P” instead of a letter “C” at the end. For example, 0001010P is the ID number of the parent of the tenth child sampled in school 0001. If twins are sampled, the ID of the first child sampled is used to generate the parent ID. For twins, there are two child-level records with the same parent ID. Children with the same teacher can be identified by finding all children on the child file with the same teacher ID.

It should be noted that there is a difference in the variable names between the base year and both the first and third grade special education teacher IDs. In the base year of the study information from special education teachers was included in a separate file and was not part of the child or teacher catalogs. The ID number for special education teachers in the base year special education file was T_ID. In the third grade data file (and the first grade data file), the special education teacher information is included with the rest of the data; thus ID numbers were needed to distinguish special education teachers from regular education teachers. In the third grade file, T5_ID is used to identify regular education teachers and D5T_ID is used to identify special education teachers.

7.2 Missing Values

All variables in the ECLS-K data use a standard scheme for missing values. Codes are used to indicate item nonresponse, legitimate skips, and unit nonresponse (see exhibit 7-1).

Exhibit 7-1. Missing values codes, ECLS-K data: School years 1998–99, 1999–2000, and 2001–02

Value	Description
-1	Not Applicable, including legitimate skips
-7	Refused (a type of item nonresponse)
-8	Don’t Know (a type of item nonresponse)
-9	Not Ascertained (a type of item nonresponse)
(blank)	System Missing, including unit nonresponse

SOURCE: U.S. Department of Education, National Center for Education Statistics, Early Childhood Longitudinal Study, Kindergarten Class of 1998–99 kindergarten, first, and third grade data collections, school years 1998–99, 1999–2000, and 2001–02.

The “Not Applicable” code (-1) has two purposes. Its primary purpose is to indicate that a respondent did not answer the question due to skip instructions within the instrument or external reasons that led a respondent not to participate. In the parent interview, where the parent or guardian was a respondent, a “Not Applicable” is coded for questions that were not asked of the respondent because of a previous answer given. For example, a question about a sibling's age is not asked when the respondent has indicated that the child has no siblings. A "Not Applicable" code is also used in the direct child assessment if a child did not participate in any section due to a disability. For the teacher and school data where the instruments are self-administered, a "Not Applicable" is coded for questions that the respondent left blank because the written directions instructed him or her to skip the question due to a certain response on a previous question.

Another use of the "Not Applicable" code is the circumstance in which it is not known whether a respondent would have answered a question series following a lead question. One example of this use of "Not Applicable" is school administrator questionnaire Question 38. Question 38 asks whether there is a gifted and talented program at the school. If the answer is "yes," the questionnaire skips to question 39 about what grade levels are included in the program. If the answer is "no," the questionnaire skips to question 40. If question 38 was left blank by the respondent, question 39 is coded "Not Applicable."

The "Refused" code (-7) indicates that the respondent specifically told the interviewer that he or she would not answer the question. This, along with the "Don't Know" code and the "Not Ascertained" code, indicates item nonresponse. The "Refused" code rarely appears in the school and teacher data because it indicates the respondent specifically wrote something on the questionnaire indicating an unwillingness to answer the question.

The "Don't Know" code (-8) indicates that the respondent specifically told the interviewer that he or she did not know the answer to the question (or in rare cases on the self-administered questionnaires, "I don't know" was written in for the question). For questions where "Don't Know" is one of the options explicitly provided, a "-8" will not be coded for those that choose this option; instead the "Don't Know" response will be coded as indicated in the value label information for that question.

The "Not Ascertained" code (-9) indicates that the respondent left a question blank that he or she should have answered. For the school and teacher self-administered questionnaires, this is the primary code for item nonresponse. For data outside the self-administered questionnaires (e.g., direct assessment scores), a "-9" means that a value was not ascertained or could not be calculated due to nonresponse.

"System Missing" appears as a blank when viewing code book frequencies and in the ASCII data file. System missing codes (blanks) in the third grade data file indicate that an entire instrument or assessment is missing due to unit nonresponse. (Note that in the first grade, “system missing” also indicated that some questions were not asked in the school administrator questionnaire for returning schools but were asked in another form of a questionnaire for new schools. This issue does not apply to the third grade file because only one form of the school administrator questionnaire was used.) An example of system missing is a child’s parent not participating in the parent interview. In this case, all questions from the parent interview will be blank (system missing). These may be translated to another value when the data are extracted into specific processing packages. For instance, SAS will translate these blanks into periods (“.”) for numeric variables.

Depending on the research question being addressed, cases with missing values may need to be recoded. It is advised that users cross-tabulate all lead questions (e.g., whether the child received child care from a relative) and followup questions (e.g., hours of child care from a relative) before proceeding with any recodes or use of the data.

Missing values for composite variables were coded using the same general coding rules as those used for other variables. If a particular composite was inappropriate for a given household—as the variable P5MOMID was for a household with no resident mother—that variable was given a value of “-1” (Not Applicable). In instances where a variable was appropriate, but complete information to construct the composite was not available, the composite was given a value of –9 (Not Ascertained). The “Refused” and “Don’t Know” codes were not used for the composites, except in the calculations of the height, weight, and body mass index (BMI) composites for spring-third grade.[1]

The ECLS-K Third Grade Public-Use Data File is provided on a CD-ROM and is accessible through an ECB that allows data users to view variable frequencies, tag variables for extraction, and create the SAS, SPSS for Windows, or Stata code needed to create an extract file for analysis. The child data file on the ECB is referred to as a “catalog.” Instructions for using the CD-ROM and ECB are provided in chapter 8.

[1] Children’s height and weight measurements were each taken twice to prevent error and provide an accurate reading. Children’s BMI was calculated based on height and weight. The rules for using “Don’t Know” and “Not Ascertained” codes for these values was as follows. If both the first and second measurement of height in the child assessment were coded as -8 (Don’t Know), then the height composite was coded as -8 (Don’t Know). If both the first and second measurements of weight were coded as -8 (Don’t Know), the weight composite was coded as -8 (Don’t Know). If either the height or weight composites were coded as not ascertained (-9), the BMI composite was coded as not ascertained (-9). If neither the height nor weight composites were coded as not ascertained, and either the height or weight composite was coded as -8 (Don’t Know), then the BMI composite was coded as -8 (Don’t Know).

7.3 Variable Naming Conventions

Variables were named according to the data source (e.g., parent interview, teacher questionnaire) and the data collection point. (A number is used to indicate in which round of data collection the variable was obtained, as follows: 5 for spring-third grade, 4 for spring-first grade, 3 for fall-first grade, 2 for spring-kindergarten, and 1 for fall-kindergarten. This numbering system is used for all variables except those beginning with "W." For those variables, 3 indicates third grade, 1 first grade, and K kindergarten.) These variable prefixes are used throughout the catalog, with a few exceptions, and are presented in two categories, (1) third grade variables and (2) cross-sectional and cross-round panel weights, in exhibit 7-2.

A few exceptions that do not follow the prefix convention are as follows:[2]

- The identifiers CHILDID, PARENTID, T5_ID, and S5_ID.
- The composite variable R5R4SCHG. This variable indicates change in school between spring-first grade and spring-third grade. Source variables and other details for this and all other composite variables can be found in table 7-12.

7.4 Composite Variables

To facilitate analysis of the survey data, composite variables were created and added to the child data file. Most composite variables were created using two or more variables, each of which is named in the text that explains the composite variable. Other composite variables are recodes of single variables. Variables based on the child assessment include height, weight, and BMI. Variables based on the teacher data include class size, percentage of limited-English-proficient children in the class, and student grade level. Variables constructed from the school data include the percentage of minority students, school type, and school instructional level. Variables constructed from the parent interview data include parent identifiers, parent demographics, household composition, household income, and poverty, child care, and child demographics. Certain composites were created using data from the Field Management System (FMS).

[2] It should be noted that in past rounds derived child demographic variables for gender, race/ethnicity, and date of birth (GENDER, RACE, DOBMM, DOBDD, and DOBYY) in the kindergarten and first grade files did not follow the prefix conventions noted above because they combined information across data collection points and/or several sources. In spring-third grade these same demographic variables begin with the prefix R5 (e.g., R5RACE). This was done because reports of these variables from parent data were prioritized over other sources in round 5 and a prefix change was used to indicate the difference to users.

Exhibit 7-2. Prefixes for ECLS-K third grade variables and cross-sectional and cross-round panel weights: School year 2001–02

Category	Description
Third grade variables	
A5	Data collected/derived from spring-third grade teacher questionnaire A
B5	Data collected/derived from spring-third grade teacher questionnaire B
C5	Data/scores collected/derived from spring-third grade direct child assessment and spring-third grade weight variables
D5	Data collected from spring-third grade special education teacher questionnaire A
E5	Data collected from spring-third grade special education teacher questionnaire B
F5	Data from spring-third grade Field Management System (FMS)
IF	Imputation flags
K5	Data collected/derived from spring-third grade school facilities checklist
L5	Data collected/derived from spring-third grade school fact sheet
P5	Data/scores collected/derived from spring-third grade parent interview
R5	Derived child demographic or child status variables for spring-third grade
S5	Data collected/derived from spring-third grade school administrator questionnaire
T5	Data/scores collected/derived from spring-third grade teacher questionnaire C
U5	Data collected/derived from spring-third grade student record abstract
W3	Third grade (cross-round) parent composite variables
Cross-sectional and cross-round panel weights	
C5C	Child-level panel weight variable from spring-third grade
C5P	Child-level panel weight for parent data from spring-third grade
C5CPT	Child-level panel weight for combined parent, child, and teacher data from spring-third grade
C45C	Child-level panel weight variable from spring-first grade and spring-third grade
C245C	Child-level panel weight variable from spring-kindergarten, spring-first grade, and spring-third grade
C1_5FC	Child-level panel weight variable from fall-kindergarten, spring-kindergarten, spring-first grade, and spring-third grade
C1_5SC	Child-level panel weight variable from fall-kindergarten, spring-kindergarten, fall-first grade, spring-first grade, and spring-third grade
C45P	Child-level panel weights for parent data from spring-first grade and spring-third grade
C245P	Child-level panel weights for parent data from spring-kindergarten, spring-first grade, and spring-third grade
C1_5FP	Child-level panel weight variable for parent data from fall-kindergarten, spring-kindergarten, spring-first grade, and spring-third grade
C1_5SP	Child-level panel weight variable for parent data from fall-kindergarten, spring-kindergarten, fall-first grade, spring-first grade, and spring-third grade

SOURCE: U.S. Department of Education, National Center for Education Statistics, Early Childhood Longitudinal Study, Kindergarten Class of 1998–99 third grade data collection, school years 2001–02.

Table 7-12 lists all the composite variables for the third grade. All basic child demographic variables are presented first. Child care and household composition variables follow the basic child demographic variables. Imputed variables are listed next, followed by demographics for parents (resident father and mother characteristics are followed by characteristics of nonresident biological parents and nonresident adoptive parents). Teacher, classroom, and school variables are listed last. Once the user identifies the composites of interest, he or she can refer to exhibit 8-60 for instructions on accessing the variables from the ECB.

7.4.1 Child Composite Variables

There are many child-level composite variables on the child catalog. Table 7-12 describes all of the composites. Some of these variables are described in further detail here.

7.4.1.1 Child's Age at Assessment (R5AGE)

The child's age was calculated by determining the number of days between the date when the child completed the ECLS-K direct child assessment and the child's date of birth (R5DOBMO, R5DOBDA, R5DOBYR). The total number of days was then divided by 30 to calculate the age in months. The child assessment date was tested for the appropriate range (March to July 2002). If the assessment date fell outside these ranges, the modal assessment date for the child's school was used.

It should be noted that the date of assessment used for R5AGE may be different from the set of assessment dates and times incorporated into methodological variables that are described further in section 7.5. These variables are not edited like those for R5AGE and are text variables that note both date and time.

7.4.1.2 Gender (R5GENDER)

The third grade gender composite was derived using the gender indicated in the parent interview (INQ.016), child report (AIQ.050), and the FMS. (The composite variable is on the file for R5GENDER, although the source variables are not.) For most of the cases the data were collected in the base year. Gender was asked in the third grade parent interview only if the information was missing from previous parent data, and asked in the child assessment only if the information was missing from previous FMS data.

In the kindergarten and first grade files, the variable GENDER was derived from the parent data and, if it were missing, the FMS. However, in examining the third grade data it was noted that there were some discrepancies in reports of a child’s sex from different sources. Using the parent report, the child report, and the FMS, the most frequently reported sex was used for the child. If there were an equal number of reports for male and female from these sources, the following hierarchy of rules was used: if the data were from the parent interview in previous rounds, then R5GENDER was equal to gender from that parent data. Otherwise, gender was updated from the third grade parent interview question. If the parent interview data were missing, gender was updated from child report. Otherwise, R5GENDER was equal to the composite GENDER from a previous round (because GENDER in previous rounds incorporated the FMS, this last step meant that the FMS was used as the final source of data).

After R5GENDER was created, all cases for which reports of gender differed by variable source were printed with the child’s name and checked against the composite. This check showed that the rules for assigning gender were successful. In three cases in which the name was clearly male or female, the gender was changed.

7.4.1.3 Child’s Date of Birth (R5DOBYY, R5DOBMM, and R5DOBDD)

In the third grade, the child’s date of birth was derived from one of three sources: the parent report (CHILDDOB), the child report (AIQ.040), or the FMS. If the child’s date of birth had been reported in a parent interview from a previous round, that value was used. Otherwise, the value from the third grade parent interview was used. If those data were not available or were outside the criteria for inclusion (June 1, 1990 to March 31, 1995), the date of birth from the child interview was used. Finally, if the child report was not available or was outside the criteria for inclusion, the FMS value was used. If the date of birth given was before June 1, 1990, or after March 31, 1995, the data were excluded.

It should be noted that in the kindergarten and first grade files, the child date of birth composites (DOBYY, DOBMM, and DOBDD) were created using two rather than three sources of data. The two sources were parent interview data and, in cases in which the parent interview data did not exist or were outside reasonable boundaries, FMS data. In spring-third grade, a third source—the child—was added.

7.4.1.4 Race/Ethnicity (W3AMERIN, W3ASIAN, W3PACISL, W3BLACK, W3WHITE, W3HISP, W3MT1RAC, W3RACETH, and R5RACE)

The composites for the child's race/ethnicity are presented in the ECLS-K files in three ways: (1) as dichotomous variables for each race/ethnicity category (W3AMERIN, W3ASIAN, W3PACISL W3BLACK, W3WHITE, W3HISP, W3MT1RAC) from the parent interview data; (2) as a single race/ethnicity composite taken from the parent interview data (W3RACETH); and (3) as a race/ethnicity composite taken from either the parent data or the FMS, with FMS data used only if parent data were missing (R5RACE).

Respondents were allowed to indicate that they belonged to more than one of the five race categories (White, Black or African American, American Indian or Alaska Native, Asian, Native Hawaiian or other Pacific Islander). From these responses, a series of five dichotomous race variables were created that indicated separately whether the respondent belonged to each of the five specified race groups. In addition, one more dichotomous variable was created for those who had simply indicated that they were multiracial without specifying a race (e.g., biracial). The retention of the dichotomous variables on the file allows users to create different composites as needed.

Data were collected on ethnicity as well. Specifically, respondents were asked whether or not they were Hispanic. Using the six race dichotomous variables and the Hispanic ethnicity variable (P4HSP_1 to P4HSP_25, depending on household size), the race/ethnicity composite variables (W3RACETH and R5RACE) were created. The categories were: White, non-Hispanic; Black or African American, non-Hispanic; Hispanic, race specified; Hispanic, no race specified; Asian; Native Hawaiian or other Pacific Islander; American Indian or Alaska Native, and more than one race specified, non-Hispanic. The child composites W3RACETH (race/ethnicity) and R5RACE (race/ethnicity) both share these categories; however, FMS data were used to fill in missing parent report data for the variable R5RACE and only parent report data were used for the variable W3RACETH. A child was classified as Hispanic if a respondent indicated the child's ethnicity was Hispanic regardless of whether a race was identified and what that race was.

For W3RACETH, if the child's race/ethnicity information was available from a parent interview in a prior data collection, this value was used and copied forward.[3] If the data were missing for a child in a previous parent interview, they were collected in third grade (FSQ.190, FSQ.195) and those data were used.

For R5RACE, responses from previous parent interviews were prioritized over the FMS. This is different from the method used to derive the variable RACE in the first grade. In the first grade, the composite RACE was copied forward from previous rounds and FMS data were used if parent reports were not available. Because parent reports were expected to be more accurate than school records, if new information about race was obtained in the third grade parent interview, it was used rather than previous information obtained from the FMS. Therefore, the third grade variable R5RACE is different from RACE in previous rounds for a minority of cases.

7.4.1.5 Child's Height (C5HEIGHT)

To obtain good measurements, each child's height was measured twice. For the height composite C5HEIGHT, if the two height values from the instrument (i.e., C5HGT1 and C5HGT2 for spring-third grade) were less than two inches apart, the average of the two height values was computed and used as the composite value. Otherwise, the value that was closest to 52.5 inches (the median height for 9-year-olds as developed by the National Center for Health Statistics (NCHS) in collaboration with the National Center for Chronic Disease Prevention and Health Prevention (NCCDPHP)) was used as the composite value.

7.4.1.6 Child's Weight (C5WEIGHT)

Each child's weight was also measured twice. For the weight composite (C5WEIGHT), if the two weight values from the instrument (i.e., C5WGT1 and C5WGT2 for spring-third grade) were less than 5 pounds apart, the average of the two values was computed and used as the composite value.

[3] A number of respondents, both in this and in prior rounds, gave some variant of "biracial" as the other-specify response to child race. In previous rounds, these responses had been considered to be uncodable, and the relevant children were given a value of –9 (not ascertained) for WKRACETH and W1RACETH. In spring-third grade, these responses were treated as multiracial, and the relevant children were given a value of 8 (multiracial) for W3RACETH. This change affected 23 children with parent data from prior rounds.

Otherwise, the value that was closest to 64.0 pounds (the median weight for 9-year-olds as developed by NCHS in collaboration with the NCCDPHP) was used as the composite value.

7.4.1.7 Child's Body Mass Index (C5BMI)

Composite Body Mass Index (BMI; variable name C5BMI), which is a calculation of the child's body weight adjusted for height, was calculated by multiplying the composite weight in pounds by 703.0696261393 and dividing by the square of the child's composite height in inches.

7.4.1.8 Child's Disability Status (P5DISABL)

A composite variable was created to indicate whether a child had a disability that was diagnosed by a professional. Questions in the parent interview about disabilities in spring-third grade asked about the child's ability to pay attention and learn, overall activity level, overall behavior and relations to adults, overall emotional behavior such as anxiety or depression, ability to communicate, difficulty hearing and understanding speech, and eyesight. For each disability or behavior, a question was asked about whether a diagnosis of a problem was obtained by a professional (CHQ.050, CHQ.110, CHQ.170, CHQ.210, CHQ.300, CHQ.335, CHQ.360). A question was also asked about receipt of therapy services or participation in a program for children with disabilities (CHQ.520).

The composite variable P5DISABL was coded 1 (Yes) if any of the source variables (CHQ.050, CHQ.110, CHQ.170, CHQ.210, CHQ.300, CHQ.335, CHQ.360, CHQ.520) about diagnosis or therapy services were coded 1 (Yes). This was done even if data for some of the source variables were missing. If data for all the source variables were missing, the composite was coded -9 (Not Ascertained). Otherwise, P5DISABL was coded 2 to indicate no reported disability.

It should be noted that the spring-third grade composite is somewhat different from the composite in previous rounds of the study because questions were added about both overall behavior and relations to adults and overall emotional behavior such as anxiety or depression. In addition, (like the spring-first grade composite P4DISABL) the spring-third grade composite does not include a question used in the fall-kindergarten questionnaire and composite that asked parents about their children's coordination in using their arms or legs.

7.4.1.9 Nonparental Care (P5CARNOW)

There are several composite variables on the file that can be used to describe child care arrangements based on information from the parent interview. One of these (P5CARNOW) describes whether the child had any type of nonparental care at the time of the interview. The creation of P5CARNOW was as follows. If the child was receiving care from a relative (CCQ.010), a nonrelative (CCQ.150), or a day care center or before or after school program at a school or in a center (CCQ.260), P5CARNOW was equal to 1 (Yes). Otherwise, if any of the three variables was unknown, P5CARNOW was coded as -9 (Not Ascertained). If the respondent indicated that the child was not currently receiving any of the three types of care (CCQ.010, CCQ.150, and CCQ.260 all equaled 2 [No]), P5CARNOW was coded as 2 (No).

It should be noted that the nonparental care as defined by P5CARNOW does not have to be received on a regular basis. However, for the composite P5HRSNOW (hours per week in child care) described below, if the nonparental care is not regular, the number of hours in care is coded as zero. This is because the child must have a regular arrangement in order for hours per week in care to be reported. Users should be aware of the differences in definitions when comparing P5CARNOW with P5HRSNOW.

7.4.1.10 Hours Per Week in Child Care (P5HRSNOW)

Another set of child care composites indicates the number of hours per week the child spent in child care. P5HRSNOW indicates the total number of hours per week the focal child spent in care at the time of the spring-third grade interview. The variable combines hours in child care arrangements in which the child spent the most time with hours from additional regular child care arrangements. It was coded as follows. If the relevant child care receipt variables for relative, nonrelative, and center-based care (CCQ.010, CCQ.150, or CCQ.260) were equal to 2 (No Receipt), or if the indicator for regular receipt of that type of care (CCQ.080, CCQ.180, and CCQ.340) was equal to 2 (No Regular Receipt), the number of hours for that type of care was coded to 0. If the receipt variables or regular receipt of care variables were refused or unknown, then the number of hours for that type of care was coded as -9 (Not Ascertained). Also, if the regular receipt variable was coded as 1 (Yes), but the hours given was refused or unknown, then the number of hours for that type of care was coded as -9 (Not Ascertained). Otherwise, if the indicator for regular receipt of care was equal to 1 (Yes), and the hours given were greater than or equal to 0, then the number of hours for that type of care was coded as the number of hours given.

The composite also includes hours spent with additional regularly scheduled providers of care of the same type. This was done to include child care arrangements such as those in which two different relatives cared for the child on a regular basis or two different child care programs were attended. For each type of care, if the care receipt variables indicated no care of that type, or if the number of providers of that type of care (questions CCQ.060, CCQ.165, and CCQ.325 indicated number of regular providers of each type) was equal to 1, then additional hours were coded to 0. Otherwise, if the number of providers or the number of additional hours (questions CCQ.140, CCQ.250, and CCQ.403 indicated number of hours spent with additional providers) was refused or unknown, then the number of additional hours was coded as -9 (Not Ascertained). Otherwise the number of additional hours was coded to equal the appropriate number of additional hours variables in the instrument (CCQ.140, CCQ.250, or CCQ.403).

This process was followed three times, once each for relative care, nonrelative care, and center-based care. If any of the three primary caregiver hour variables or the three additional hours variables was missing then the total number of hours was coded as -9 (Not Ascertained). Otherwise the total number of hours in regularly scheduled child care was coded as the sum of the six hour variables.

It should be noted that in earlier rounds, if the primary care arrangement hours were not missing and the additional hours were missing, the primary caregiver hours were used for the composite. In spring-third grade, if any of the primary or additional hours variables were missing, the composite was missing. This change makes the variable represent all types of regular care rather than prioritizing primary arrangements. Because there are slightly more missing data for the composite in spring-third grade than in the previous rounds (the percentage of “not ascertained” answers was 0.8 percent in fall-kindergarten, 0.6 percent in spring-first grade, and 1.5 percent in spring-third grade), users who want to prioritize primary care hours over additional hours may want to calculate their own composite.

Although P5HRSNOW was created almost identically to the same composite variable in kindergarten (P1HRSNOW), with the exception noted above, there was one other difference. In kindergarten, questions were asked about whether the child was ever in a particular type of care. If not, P1HRSNOW was set to 0. Because questions about the child having ever been in a particular type of care were not included after the kindergarten year, they were not part of the composite variable definition for either the third or first grade variables.

7.4.1.11 Number of Child Care Arrangements (P5NUMNOW)

Another composite variable (P5NUMNOW) was used to indicate the total number of all types of care arrangements the focal child had at the time of the spring-third grade parent interview. The variable was created as follows. If any of the child care receipt variables for relative, nonrelative, or center-based care (CCQ.010, CCQ.150, or CCQ.260) was refused, unknown, or missing, then P5NUMNOW was coded as -9 (Not Ascertained). If any of the care receipt variables was equal to 1 (Yes), but its corresponding number of arrangements variable (CCQ.060, CCQ.165, and CCQ.325) was refused, unknown, or missing, then P5NUMNOW was again coded as -9 (Not Ascertained). Otherwise, the number of arrangements indicated in CCQ.060, CCQ.165, and CCQ.325 were summed to obtain the total number of current child care arrangements.

The differences in how missing data are handled for each of the child care composites are important to note when combining variables. For example, because P5NUMNOW requires that the number of child care arrangements be known, it is possible for a child to have P5CARNOW =1 (child was in nonparental care) and have P5NUMNOW be -9 (Not Ascertained).

7.4.1.12 Primary Nonparental Child Care Arrangement (P5PRIMNW)

A composite variable (P5PRIMNW) was created to indicate the primary, nonparental child care arrangement in which the child spent the most hours per week at the time of the spring-third grade interview. This variable is for children in a regular care arrangement. The values for this variable follow:

- 0=No nonparental care
- 1=Relative care in child’s home
- 2=Relative care in another home
- 3=Nonrelative care in child’s home
- 4=Nonrelative care in another home
- 5=Center-based program
- 6=Two or more programs
- 7=Location of care varies

To obtain the composite, hours were compared for relative care in the child's home (CCQ.090) or in other home (CCQ.070); nonrelative care in child's home (CCQ.190) or in other home (CCQ.170); and center/program care (CCQ.355). First, the composite P5HRSNOW, described earlier, was used to code individuals missing current hours of care (P5HRSNOW=-9) or with no hours of nonparental care (P5HRSNOW=0). Those with missing hours of care were coded as -9 (Not Ascertained); those with no hours of care or no regularly schedule care were coded as 0.

For the remaining cases, if the number of hours of either relative or nonrelative care (given in CCQ.090 and CCQ.190) were higher than all other hours of care, the variable indicating location of care for that type was examined using instrument items CCQ.070 and CCQ.170. If location of care was missing, then P5PRIMNW was coded as -9 (Not Ascertained); if P5PRIMNW was not missing, then P5PRIMNW was coded 1, 2, 3, or 4, depending on the type (relative/nonrelative) and location (child's home/other home) of care. Otherwise, if the number of hours of care in center-based programs (CCQ.355) was higher than for relative or nonrelative care, then P5PRIMNW was coded as 5. If the number of hours of care was equal for two or more types of care, P5PRIMNW was coded as 6. P5PRIMNW was coded as 7 if the location of care varied between two homes.

It should be noted that it is possible to have missing data for the primary child care arrangement (P5PRIMNW), but still have information on the number of hours of child care a child has (P5HRSNOW). This is because there must be information about the location of care in order to have a valid value for P5PRIMNW.

7.4.2 Family and Household Composite Variables

Many composites were created to capture information about the sampled children's family and household characteristics. Several of these are described below. All of the family and household composites are listed and described in table 7-12.

7.4.2.1 Number of Siblings (P5NUMSIB)

The composite P5NUMSIB indicates the total number of siblings (full, step, adoptive, or foster) with whom the child lived in the household (FSQ.160 and FSQ.170). Siblings were identified through the respondents' stated relation of the sibling to the focal child. In addition, any child that was reported to be a child of the focal child's parent/guardian was considered a sibling of the focal child.

7.4.2.2 Parent and Household Members' Age (P5LESS18, P5OVER18, P5HDAGE, and P5HMAGE)

There are several composite variables on the file that refer to the ages of adults and children in the household. These are P5LESS18 (total number of people in the household under age 18, including focal child, siblings, and other children), P5OVER18 (total number of people in the household age 18 or older, siblings, and other children), P5HDAGE (age of resident father), and P5HMAGE (age of resident mother). The ages of these persons in the household were collected during the fall of kindergarten in the household matrix. However, in subsequent years of the study, questions about age were not asked for household members who were previously in the household. This was done to save interviewing time. In the third grade, ages were collected only for new household members. Otherwise, ages were incremented by adding years based on the round in which the person joined the study. Age changes were made to increase the ages of all household members other than the focal child and twin (the ages of the focal child (and twin, if applicable) were updated based on birthdate).

The ages of all household members who were not new to the study in spring-third grade (other than the focal child and twin) were increased by the numbers shown in table 7-1. The guidelines for creating these were as follows: (1) half years could not be included, and (2) the same number of years was added for those who entered the study during the same school year. The numbers were made to err on the side of making persons older rather than younger because this would cause fewer problems with range checks and displays in the parent interview if there was a discrepancy between actual age and imputed age.

Table 7-1. Incremented ages of previous household members based on round household member entered study, spring-third grade: School year 2001–02

Round in which household member joined study	Number of years to add for spring-third grade
Fall 1998	+4
Spring 1999	+4
Fall 1999	+3
Spring 2000	+3

NOTE: In order to save interviewing time, questions about age were not asked about household members whose ages were reported in previous rounds of the study. Instead, years were added to the originally reported age. The number of years added was based on when the household member joined the study.
SOURCE: U.S. Department of Education, National Center for Education Statistics, Early Childhood Longitudinal Study, Kindergarten Class of 1998–99 third grade data collection, school year 2001–02.

7.4.2.3 Food Security Status

Food security status of the children's families was assessed based on responses to the 18 food security questions (P5WORRFD through P5NOMONY) in the spring-third grade parent interview. The questions measured a wide range of food insecurity and reduced food intake issues. They were combined into a scale using statistical methods based on the Rasch measurement model. The items and the food security scale based on them have been validated using both ethnographic and statistical methods. For spring-third grade, composites were created for Household Food Security scale variables and Children's Food Security scale variables (for spring-kindergarten, composites were created only for Household Food Security scale variables). Calculations of the Household Food Security Scale variables were carried out in accordance with the standard methods described in *Guide to Measuring Household Food Security, Revised 2000* (U.S. Department of Agriculture, 2000). Calculations of the Children's Food Security Scale variables were carried out in accordance with the standard methods described in *Measuring Children's Food Security in U.S. Households, 1995-99* (U.S. Department of Agriculture, 2002). Analysis of the ECLS-K data using Rasch methods indicated that use of the standard benchmark household scores was appropriate.

7.4.2.4 Food Security Status: Continuous Measures (P5FSSCAL and P5FSCHSC)

P5FSSCAL is the scale score presentation of the Household Food Security items. It is a continuous, interval-level measure of food insecurity and is appropriate for linear models. This scale score is a Rasch transformation of the raw score (P5FSRAW). Valid values range from 1.4 to 13.0, with higher values indicating more severe food insecurity. Under Rasch-model assumptions, the scale score for families that affirm no items (in other words, they did not provide "yes" answers to the questions and the raw score = 0) is indeterminate. It is less than the lowest measured value (1.4), but its precise value is unknown and may vary substantially among families. P5FSSCAL for such cases is assigned a value of -6. If these cases (a substantial majority of all cases) are included in linear models, appropriate methods must be used to take into account this indeterminacy.

P5FSCHSC is similar to P5FSSCAL but is the Children's Food Security scale score. This is a measure of the severity of food insecurity or hunger experienced by children in the household in the previous 12 months. Valid values range from 4.1 to 12.2, with higher values indicating more severe food deprivation. The scale score is undefined for households that affirmed no child-referenced items (see discussion of P5FSSCAL above).

7.4.2.5 Food Security Status: Categorical Measures (P5FSSTAT and P5FSCHST)

P5FSSTAT is a categorical measure of Household Food Security status formed by dividing P5FSSCAL into four ordered categories: food secure, food insecure without hunger, food insecure with hunger (moderate), and food insecure with hunger (severe). P5FSSTAT is appropriate for comparing prevalence rates of food insecurity and hunger across subpopulations and can be used as a categorical variable in associative models. There are few cases in the most severe category, so for most prevalence reporting purposes, the two categories of food insecure with hunger (moderate and severe) should be collapsed and reported as a single category. When interpreting food security statistics, users should remember that food security status is a household-level characteristic. In most households classified as food insecure with hunger, the children in the household were not hungry.

P5FSCHST is a categorical measure of Children's Food Security status that identifies households with hunger among children at some time during the 12 months prior to the survey. This variable is appropriate for comparing prevalence rates of hunger among children across subpopulations.

There were few households (n=27, 0.2 percent) that reported hunger among children, so the analytic utility of this variable is limited. However, for analytic purposes, other categories of children's food insecurity delineated by less severe thresholds (based on children's food security raw scores or scale scores) may be useful. For example, Nord and Bickel (2001) suggested a threshold of 2 or more affirmative responses as representing reduced quality and variety of children's diets. When interpreting children's food security statistics, users should remember that these variables represent conditions among all children in the household and may not reflect experiences of the child in the ECLS-K study if there are other children in the household.

7.4.2.6 Food Security Status: Raw Scores (P5FSRAW and P5FSCHRA)

The Household Food Security raw score, P5FSRAW, is a count of affirmative responses to the 18 items. This is an ordinal-level measure of food insecurity and is not recommended for use in analysis. The Children's Food Security raw score, P5FSCHRA, is a count of affirmative responses to child-referenced items. Responses to items skipped because of screening are assumed to be negative. Families with no valid responses are coded as missing (-9). Missing item responses of families with one or more valid responses are imputed as negative responses (only 0.13 percent of the sample).

7.4.2.7 Socioeconomic Status (SES) and Poverty (W3DADSCR, W3MOMSCR, W3SESL, W3SESQ5, W3INCCAT, W3POVRTY)

Socioeconomic status (SES) was computed at the household level using data for the set of parents who completed the parent interview in spring-third grade. The SES variable reflects the socioeconomic status of the household at the time of data collection for spring-third grade (spring 2002). The components used to create the SES variable were as follows:

- Father/male guardian's education;
- Mother/female guardian's education;
- Father/male guardian's occupation;
- Mother/female guardian's occupation; and
- Household income.

Occupation was recoded to reflect the average of the 1989 General Social Survey (GSS) prestige score. This was computed as the average of the corresponding prestige scores for the 1980 Census occupational categories covered by the ECLS-K occupation. Table 7-12 provides details on the prestige score values (W3DADSCR, W3MOMSCR).

The variables were collected as follows:

- Income. The information about income was collected in spring-third grade. Broad-range and detailed-range income questions were asked of all participants. The broad range classifies household income as $25,000 and less per year, or as greater than $25,000. The detailed range classifies household income as shown in table 7-2.

- Households that were determined to meet the size and income criteria related to poverty shown in table 7-3 were asked to report income to the nearest $1,000. (We call this exact income for simplicity.) Because not all households were asked to report exact income, the midpoint of the detailed income range was used to compute the SES composite variable.

- Parent's education. The information about parent's education was collected or updated in spring-third grade.

- Parent's occupation. The information about parent's occupation was collected or updated in spring-third grade.

Table 7-2. Levels of detailed income range, spring-third grade: School year 2001–02

Detailed income range	Total household income
1	$5,000 or less
2	$5,001 to $10,000
3	$10,001 to $15,000
4	$15,001 to $20,000
5	$20,001 to $25,000
6	$25,001 to $30,000
7	$30,001 to $35,000
8	$35,001 to $40,000
9	$40,001 to $50,000
10	$50,001 to $75,000
11	$75,001 to $100,000
12	$100,000 to $200,000
13	$200,001 or more

SOURCE: U.S. Department of Education, National Center for Education Statistics, Early Childhood Longitudinal Study, Kindergarten Class of 1998–99 third grade data collection, school year 2001–02.

Table 7-3. Households asked to report income to the nearest $1,000, spring-third grade: School year 2001–02

Household size	Total household income
Two or three	$15,000 or less
Four	$20,000 or less
Five or six	$25,000 or less
Seven	$30,000 or less
Eight	$35,000 or less
Nine or more	$40,000 or less

SOURCE: U.S. Department of Education, National Center for Education Statistics, Early Childhood Longitudinal Study, Kindergarten Class of 1998–99 third grade data collection, school year 2001–02.

Not all parents completed the parent interview; among those who did, not all responded to every question. Therefore, there were missing values for some of the components of the SES composite variable. Only a small percentage of values for the education and occupation variables were missing; a larger proportion of households had missing values for the detailed income range (see table 7-4).

Table 7-4. Missing data for SES source variables, spring-third grade: School year 2001–02

Variable	Number missing	Percent
Mother’s education	256	1.92
Father’s education	265	1.98
Mother’s occupation	308	2.31
Father’s occupation	367	2.75
Detailed income range	1,482	11.10

SOURCE: U.S. Department of Education, National Center for Education Statistics, Early Childhood Longitudinal Study, Kindergarten Class of 1998–99 third grade data collection, school year 2001–02.

A two-stage procedure was used to impute missing values for each component of the SES composite variable. First, if a parent had completed an interview in the kindergarten or first grade year, missing values for the spring-third grade education, occupation, and detailed income range were filled in with values from the previous years. The rationale for this approach was that the best source of data for an individual or a household was the data from a previous year.

This first imputation stage was implemented as follows:

- Education level was brought forward from the most recent previous round. This was done only if the same person was the parent figure both in spring-third grade and in the earlier round.
- Occupation was brought forward only if the individual was in the labor force (i.e., was working at a paid job, on vacation from a paid job, or looking for a job). It was also required that the same person be the parent figure both in spring-third grade and in the earlier round. NOTE: Prestige scores were not assigned to individuals unless they were in the labor force, regardless of whether they reported an occupation.
- Detailed income category was brought forward from the most recent previous round.

Second, data still missing after this initial step were imputed using a hot deck methodology. In hot deck imputation, the value reported by a respondent for a particular item is assigned or “donated” to a “similar” person who failed to respond to that question. Auxiliary information known for both donors and nonrespondents is used to form groups of persons having similar characteristics. These groups of similar respondents and nonrespondents are called “imputation cells.” The imputed value for a case with a missing value is taken from a randomly selected donor among the respondents within the cell.

Imputation cells were defined by respondent characteristics that were the best predictors of the variables to be imputed. These relationships had been determined previously by CHAID (Chi-squared Automatic Interaction Detector) analyses of the base year data. Missing values for the education, occupation, and detailed income range variables were imputed by the hot deck method for all households. Hot deck imputation was done in a sequential order, separately, by type of household (female single parent, male single parent, and both parents present). For households with both parents present, the mother’s and father’s variables were imputed separately. Imputed as well as reported values were used to define imputation cells; missing values for donor characteristics were treated as a separate category. No imputed value was used as a donor. No donor was used more than once. The order of hot deck imputation for all the variables was from the lowest percent missing to the highest.

Occupation imputation involved two steps. First, the labor force status of the parent was imputed (i.e., whether the parent was employed). Then the parent’s occupation was imputed only for those parents whose status was identified as employed either through the parent interview or the first imputation step. The detailed income range was imputed in two steps: first for cases where the broad income range was known, and second for cases where it was unknown.

For households where both parents were present, the order of hot deck imputation was as follows:

- Mother's education;
- Father's education;
- Mother's labor force status;
- Mother's occupation;
- Father's labor force status;
- Father's occupation;
- Detailed income range, where the broad income range was known; and
- Detailed income range, where the broad income range was unknown.

At this point, all of the missing values had been imputed. However an exact income value was still required to construct the SES composite. The midpoint of the detailed income range was assigned for this purpose to all households.

The log of the detailed income range midpoint was then used to compute the SES composite. This value does not vary widely within the levels of the detailed income range, so the midpoint was a reasonable choice. It was used only for the purpose of computing the SES composite and was not retained in the data file.

All missing values of the SES components were imputed by the process described above. Tables 7-5 through 7-8 summarize the results.

Table 7-5. Selected statistics on imputed education variables, spring-third grade: School year 2001–02

SES component	Total missing	Number of values filled from previous rounds	Number of values imputed by hot deck	Number of cases resolved
Mother's education	256	223	33	256
Father's education	265	220	45	265

SOURCE: U.S. Department of Education, National Center for Education Statistics, Early Childhood Longitudinal Study, Kindergarten Class of 1998–99 third grade data collection, school year 2001–02.

Table 7-6. Selected statistics on imputed labor force status, spring-third grade: School year 2001–02

Labor force status	Number of values filled from previous rounds	Number of values imputed by hot deck	Number of cases resolved
Mother			
Total missing			227
In labor force	148	24	172
Not in labor force	51	4	55
Father			
Total missing			189
In labor force	147	27	174
Not in labor force	11	4	15

SOURCE: U.S. Department of Education, National Center for Education Statistics, Early Childhood Longitudinal Study, Kindergarten Class of 1998–99 third grade data collection, school year 2001–02.

Table 7-7. Selected statistics on imputed occupation variables, spring-third grade: School year 2001–02

Occupation	Number of values filled from previous rounds	Number of values imputed by hot deck	Number of cases resolved
Mother			308
Total missing			
Occupation	41	212	253
Not in labor force[1]	51	4	55
Father			367
Total missing			
Occupation	83	269	352
Not in labor force[1]	11	4	15

[1] No occupation was imputed if "not in labor force" was filled from previous rounds or imputed by hot deck.
SOURCE: U.S. Department of Education, National Center for Education Statistics, Early Childhood Longitudinal Study, Kindergarten Class of 1998–99 third grade data collection, school year 2001–02.

Table 7-8. Selected statistics on imputed detailed income range, spring-third grade: School year 2001–02

SES component	Total missing	Number of values filled from previous round	Number of values imputed by hot deck Broad income range		Number of cases resolved
			Known	Unknown	
Detailed income range	1482	1441	12	29	1482

SOURCE: U.S. Department of Education, National Center for Education Statistics, Early Childhood Longitudinal Study, Kindergarten Class of 1998–99 third grade data collection, school year 2001–02.

Once the components of the SES variable were imputed, their corresponding *z*-scores or normalized values were computed. The expression of *z*-score z_{hi} for the *h*-th component in the *i*-th household is

$$z_{hi} = \frac{x_{hi} - \bar{x}_w}{se(\bar{x}_w)},$$

where

x_{hi}	is	the value of the *h*-th SES component for the *i*-th household;
w_i	is	the base weight for the *i*-th record;
$\bar{x}_w$	is	the weighted mean of x_{hi}; and
$se(\bar{x}_w)$	is	the standard error of $\bar{x}_w$.

Thus, each component was converted to a *z*-score with mean of 0 and a standard deviation of one. For income, the component x_i is the logarithm of the income for *i*-th household. The logarithm of income was used because the distribution of the logarithm of income is less skewed than the direct income values. The SES value for the *i*-th household was then computed as

$$SES_i = \frac{\sum_{h=1}^{m_i} z_{hi}}{m_i},$$

where m_i is the number of nonmissing SES components for the *i*-th household. W3SESL is the continuous variable for the SES composite that ranges from –2.49 to 2.58. As described, the SES composite is the average of up to five measures, each of which was standardized to have a mean of 0 and a standard deviation of 1, hence the negative values. For analyses that require a continuous SES measure, such as multivariate regressions, W3SESL is the variable to use. A categorical SES variable (W3SESQ5) was created that contains the quintile for the value of the composite SES for the child. Quintile 1 represents the lowest SES category and quintile 5 represents the highest SES category. The quintiles were computed at the child level using the spring-third grade parent weights. For categorical analyses, use W3SESQ5 and the parent weight.

Note that for households with only one parent present, not all the components were defined. In these cases, the SES was computed by averaging the available components.

The imputed detailed income range variable (W3INCCAT) was also used to create a household-level poverty variable (W3POVRTY). Income was compared to Census poverty thresholds for 2001, which vary by household size. Table 7-9 shows the detailed income categories used in the ECLS-K parent interview for determining whether to ask a more detailed question about income to the nearest 1,000. For comparison, the table also shows weighted poverty thresholds from Census data.[4] Households whose income fell below the appropriate threshold were classified as poor. For example, if a household contained two members, and the household income was lower than $11,569, then the household was considered to be below the poverty threshold.

Table 7-9. ECLS-K spring-third grade and Census poverty thresholds for 2001: School year 2001–02

Household size	ECLS-K income categories	Census weighted average thresholds for 2001
Two	Less than or equal to $15,000	$11,569
Three	Less than or equal to $15,000	$14,128
Four	Less than or equal to $20,000	$18,104
Five	Less than or equal to $25,000	$21,405
Six	Less than or equal to $25,000	$24,195
Seven	Less than or equal to $30,000	$27,517
Eight	Less than or equal to $35,000	$30,627
Nine or more	Less than or equal to $40,000	$36,286

SOURCE: U.S. Census Bureau, Current Population Survey. http://www.census.gov/hhes/poverty/threshld/thresh01.html; U.S. Department of Education, National Center for Education Statistics, Early Childhood Longitudinal Study, Kindergarten Class of 1998–99 third grade data collection, school year 2001–02.

7.4.2.8 Parent Education (W3PARED, W3DADED, and W3MOMED)

There are three parent education composites on the file. These are W3PARED (the highest level of education for the child's parents or nonparent guardians who reside in the household), W3DADED (father's highest level of education), and W3MOMED (mother's highest level of education). The variables include both parent (birth, adoptive, step, and foster) and nonparent guardians. For example, if the child had no parents but had a guardian, the education of the guardian and his or her spouse were used in the creation of the composites if the guardian was specified as such in the relationship variable or if the guardian was the respondent/respondent's spouse and there were no other parent figures in the household.

[4] The ECLS-K provides an approximate, but not exact measure of poverty. Income category thresholds used in the parent questionnaire are similar, but not identical to those from weighted Census averages.

In spring-third grade, parent education level was updated from spring-first grade if it was a household that had been part of that round of the study. Respondents were asked if they or their corresponding parent figures, if applicable, completed any additional grades of school or received any diplomas or degrees (PEQ.010). If so, PEQ.020 asked what grade the parent completed or what degree was received. If there was no education information to update from spring-first grade, respondents were asked for their highest education level in PEQ.020. If this education level was less than the education level reported in a previous round, the higher education level was kept for the composite.[5]

If both parents/guardians resided in the household, W3PARED was the highest value for education level from either the mother/guardian in W3MOMED or the father/guardian in W3DADED. If the household only had one parent or guardian, then W3PARED was equal to either W3MOMED or W3DADED depending on which parent or guardian resided with the child. If the education data for either of the parents were missing[6] it was imputed, and the composite W3PARED was created based on both the reported and imputed data.

7.4.2.9 Parent Race/Ethnicity (P5HDRACE and P5HMRACE)

The composites for race/ethnicity for the parents were calculated in the same way as those for the child, except that there is not a variable that supplements parent reported race/ethnicity with FMS data similar to the variable R5RACE for children. All data on parent race/ethnicity come from the parent interview. Race/ethnicity for parents is presented in the spring-third grade data file as a categorical race/ethnicity composite (for the father/male guardian it is P5HDRACE, and for the mother/female guardian it is P5HMRACE).

---

[5] Because of a programming issue, many respondents were asked the education question in PEQ.020 rather than asked to update education information obtained in a previous round. For 1,385 mothers and 1,124 fathers, the spring-third grade education levels were lower than the education levels provided in the base year of the study. One source of the discrepancy may be that the question structures were different in the base year of the study and spring-third grade. In the base year, if a respondent answered any grade less than 12, he or she was then asked if the person had received a high school diploma or its equivalency, such as a GED. For example, if he or she answered that the highest grade completed was 9th grade, but that he or she had completed a GED, then the highest education level would not be 9th grade, it would be high school diploma/equivalent. However, in spring-third grade, the follow up question on high school equivalency was not asked. Thus, a base year respondent (who was a nonrespondent in spring-first grade) who had answered 9th grade in the base year, would answer 9th grade again in the spring-third grade data collection, but this time due to the absence of the followup question, the highest level of education completed would be 9th grade. Based on a review of the cases, the higher of the two education levels was used. This solution took into consideration that the base year cases had a followup item that collected the information in a more informative way.

[6] Missing data were due to “refused” or “don’t know” answers from respondents, in addition to program issues that caused a few cases to have missing data.

Respondents were allowed to indicate that they belonged to more than one of the five race categories (White, Black or African American, American Indian or Alaska Native, Asian, Native Hawaiian or other Pacific Islander). From these responses, a series of five dichotomous race variables were coded that indicated separately whether the respondent belonged to each of the five specified race groups. In addition, one more variable was coded for those who had simply indicated that they were multiracial without specifying the race (e.g., biracial).[7] The dichotomous codes for each of the race variables are not provided on the spring-third grade file, but the composite derived from the responses is provided.

Parent race/ethnicity was obtained for all parents and spouses of respondent parents, but may or may not have been collected for a parent's boyfriend or girlfriend. For example, in a family with a birth mother and stepfather the race/ethnicity of both parents was obtained. However, in a family with a birth mother and the mother's boyfriend, if the mother's boyfriend was not identified as a spouse or partner of the mother, the race/ethnicity of the mother was obtained but that of the boyfriend was not.

7.4.3 Teacher Composite Variables

Details about how two of the teacher composites, child grade level and class type, were created are provided here. All of the teacher composites are listed and described in table 7-12.

7.4.3.1 Grade-Level Composite (T5GLVL)

To create the grade-level composite (T5GLVL), five possible sources of information were used: (1) teacher questionnaire part C (Q1 T5GRADE for grade level); (2) special education teacher part B (Q2 E5GRADE for grade level); (3) child assessment introductory section (AIQ.030 C5INGRAD); (4) child assessment closing section (ACQ.005 C5THIRD and ACQ.010 C5GRADE, completed by interviewer), and (5) FMS information about grade level.

If conflicts existed among these five sources, the grade level indicated by the majority of the nonmissing sources was used for T5GLVL. When there were five, four, or three sources of information

[7] In a previous round of the study, respondents who reported they were "biracial" in the "other" category were classified as "uncodable." These responses were reclassified as "multiracial" in spring-third grade. Eighteen parents had race changes due to this change in coding.

and three were in agreement, the grade level indicated by the three sources was taken. When there were four sources of information and only two were in agreement, the grade level indicated by the two sources in agreement was taken. When there were three sources of information and two were in agreement, the grade level indicated by the two sources in agreement was taken. When there were four or five sources of data, and two sources indicated one option and the other two indicated another option, the grade indicated in a particular source was selected, according to the hierarchy presented below.

- Classroom teacher, T5GRADE
- Special education teacher, E5GRADE
- Assessment introduction, C5INGRAD
- Assessment closing, C5THIRD and C5GRADE
- FMS

In establishing this hierarchy, it was assumed that teachers had the best knowledge and that school records (on which the FMS are based) were more apt to be in error. It was also assumed that children were reliable reporters of their own grade level, so their reports were prioritized over the FMS. When equal numbers of sources were in conflict (1 vs. 1) or (2 vs. 2) or (1 vs. 1 vs. 1), the decision was made by using the information from the source highest on the list above.

One exception to this hierarchy was made. Because the FMS and AIQ grade-level information did not allow for ungraded classrooms, the FMS and AIQ information were not considered in any case in which at least one source indicated an “ungraded” classroom.

It should be noted that in spring-first grade, there was information about grade level from the student record abstract; however, there were no grade-level questions in the child assessment at that time. In spring-third grade, grade level was not asked in the student record abstract, but was included as part of the child assessment instead.

7.4.3.2 Class Size (A5CLSZ)

The composite for class size was created from class totals provided in three different questions in the teacher questionnaire, part A. The totals for race/ethnicity (Q4, A5TOTRA), age (Q3, A5TOTAG), and sex (Q5, A5BOYS+A5GIRLS) were compared. If one of the totals differed, but two totals matched, the total shown by the two matching sources was used. If there were no matches among the totals, the total for the composite was set using, in order of priority, the sex, age, or race/ethnicity total. Otherwise, A5CLSZ was coded as -9 (Not Ascertained).

It should be noted that the class size composite A5CLSZ was used in spring-third grade as the denominator for the composite variables A5PMIN, A5PHIS, A5PBLK, and A5PLEP. In previous years, the total class size used in the calculation of these variables was based on the total number of children in the question about numbers of children by race/ethnicity.

7.4.4 School and Class Composite Variables

Variables on school and class characteristics were constructed from the teacher and school data and the sample frame. Details on how some of the variables were created follow.

7.4.4.1 School Type (S5SCTYP)

In spring-third grade, S5SCTYP was created as follows. Questions 2 (L5PUBLIC) (whether school is public) and 4 (L5CATHOL, L5OTHREL) (type of private school) from the school fact sheet, along with school sample frame data, were used to create the school type composite variable. If the response to question 2 (Is this a public school?) was “Yes,” then S5SCTYP was coded “public.” If the response to question 4.a. (L5CATHOL) (Is your school a Catholic school) was “Yes,” then the school was coded as “Catholic.” Otherwise, if the response to question 4.b. (L5OTHREL) (Is your school private with another religious affiliation?) was “Yes,” then S5SCTYP was coded as “private, other religious.” Otherwise, because the skip pattern to question 4 was used only if the school was private, if the response to question 4.a. (L5NAISKL, private school accredited by NAIS), question 4.d. (L5OTHPRI, other private), question 4.e. (L5PVTSPD, special education school-primarily serves children with disabilities), or question 4.f. (L5PVTEAR, an early childhood center-school or center includes preschool and/or early elementary grades) was “Yes,” then S5SCTYP was coded as “private, non-religious.”

If S5SCTYP could not be coded from the school fact sheet, reports of school type from the same school in previous rounds were used (in previous rounds, school type was asked in the school administrator questionnaire and the variable names were S4SCTYP, S3SCTYP, S2KSCTYP, and CS_TYPE2). If those sources were unavailable, a variable from the school master file (taken from the 1999–2000 PSS/2000–01 CCD frame) was used to code S5SCTYP. If S5SCTYP could not be coded, S5SCTYP was coded as -9 (Not Ascertained). If the child was schooled at home, the composite was coded as -1 (Not Applicable).

7.4.4.2 Public or Private School (S5PUPRI)

This variable is a less detailed version of school type (with only two categories—public and private) and is derived from the school type composite S5SCTYP described above. In spring-third grade, it was created as follows. If S5SCTYP was 4 (public), then S5PUPRI was coded as public (1). If S5SCTYP was 1-3 (Catholic, other religious, other private) then S5PUPRI was coded as private (2). If S5SCTYP was coded as Not Ascertained (-9), then S5PUPRI was -9 (Not Ascertained). If S5SCTYP was coded “Not Applicable,” then S5PUPRI was coded “Not Applicable.”

7.4.4.3 School and Grade-Level Enrollment (S5ENRLS, S5ENRLT)

There are two composite enrollment variables on the third grade file: total school enrollment (S5ENRLS) and third grade enrollment (S5ENRLT). Total school enrollment was created using the school enrollment variable from the school administrator questionnaire (S5ANUMCH). If this variable was missing, data for private schools were taken from the 1999–2000 Private School Survey (PSS) and data for public schools were taken from the 2000–01 Common Core of Data (CCD) public school universe. If these were also missing, the variable was coded -9 (Not Ascertained). If the child was schooled at home, the composites were coded -1 (Not Applicable).

Third grade enrollment was not obtained during data collection. The third grade enrollment data for private schools came from the 1999–2000 PSS data. The enrollment data for public schools came from the 2000–01 CCD public school universe data.

7.4.4.4 Percent Minority Students in the School (S5MINOR)

The composite variable S5MINOR indicates the percentage of minority students in a school in spring-third grade. The composite is based on a question in the school administrator questionnaire (Q3) that was used to ask about the number or percentage of students in the following categories: Hispanic, regardless of race; Black, not of Hispanic origin; White, not of Hispanic origin; Asian or Pacific Islander; American Indian or Native Alaskan; and other. The composite was based on the sum of percentages for all categories except White, not of Hispanic origin. In some cases, the composite could not be obtained from the data because of missing data or errors. If the composite could not be derived from the data, percent minority was obtained from the CCD (for public schools) or the PSS (for private schools). If these data were missing, the composite was coded -9 (Not Ascertained). If the child was schooled at home, the composite was coded as -1 (Not Applicable).

In spring-third grade, school administrators were allowed to report their answers to the student racial composition questions as either numbers or percents, whereas in spring-kindergarten they were asked to report those answers as percents. All answers recorded as numbers in spring-third grade were converted to percentages for the composite variable. The sum of the answers across all categories was allowed to add within +/- 5 percent of the reported total. In a few cases, this produced answers slightly over 100 percent. These were topcoded to 100 percent.

A flag for each race/ethnicity variable indicates whether the answer was reported as a number or a percent.[8] Because the composite is calculated as a percent, these flags will not be needed by users unless the analyst is interested in examining how answers were reported. If the flags (S5ASNFL, S5HSPFL, S5BLKFL, S5WHTFL, S5INDFL, and S5OTHFL) were equal to 1 for each of the race variables S5ASNPCT, S5HISPPCT, S5BLKPCT, S5WHTPCT, S5INDPCT, S5OTHPCT, these 6 race/ethnicity variables were reported by the respondent as percentages.

It should be noted that the spring-third grade composite was created in the same way as the composite for spring-first grade. However, both the spring-third grade and first grade composites are

[8] There were also other questions in the school administrator questionnaire that allowed for answers to be recorded as either a number or percent. The flags for these variables are S5ADAFLG (average daily attendance reported as number/percent), S5ASNFLG (question about Asian or Pacific Islander teachers reported as number or percent), S5HSPFLG (question about Hispanic teachers reported as number or percent), S5BLKFLG (question about black teachers reported as number or percent), S5WHTFLG (question about white teachers reported as number or percent), S5INDFLG (question about American Indian or Native Alaskan teachers reported as number or percent), and S5OTHFLG (question about teachers of other races reported as number or percent). In all cases, the final variables related to these flags are reported as percentages, but the flags indicate how the answers were originally recorded by respondents.

slightly different from the one used in spring-kindergarten (S2MINOR) because the school administrator questionnaire item that asked about the percent of minority students in the school had different response options. In spring-kindergarten, the percent of minority students was derived from answers to the school administrator questionnaire by determining the percentage of children who were of either Hispanic or Latino origin (question 14) and the percentage of children who were American Indian or Alaska Native, Asian, black or African American, or Native Hawaiian or Other Pacific Islander (question 15) to create the percent minority composite. In spring-third grade, Hispanic or Latino origin and race were included in the same question.

7.4.4.5 School Instructional Level (S5SCLVL)

The purpose of this composite is to classify schools based on the highest grade taught in the school. This composite is taken in spring-third grade from the school fact sheet (Q1, L5PRKNDR, L5KINDER, L5GRADE1, L5SECOND, L5THIRD, L5FOURTH, L5FIFTH, L5SIXTH, L57TH, L58TH, L5NINTH, L5TENTH, L511TH, L512TH). The highest grade level circled on the form was determined, and the grade level was classified accordingly. If data were missing, data were used from the school master file (based on the 1999–2000 PSS and the 2000–01 CCD) to fill in instructional level. If school master file data were unavailable for a particular school, data from previous school administrator questionnaires from spring-first grade or spring-kindergarten schools (S4SCLVL and S2KSCLVL) were used to determine instructional level. If those sources were also not available, S5SCLVL was coded as -9 (Not Ascertained). If the child was schooled at home, the composite was coded as -1 (Not Applicable).

In previous rounds of the study, this composite was taken from the school administrator questionnaire rather than the school fact sheet. Also, in previous rounds, if the question about grade levels in the school was left blank, another question from the school administrator questionnaire about grade levels that participated in special programs was used. If the respondent did not answer either of these questions, then school sample frame data were used to determine the value for the composite in previous rounds.

7.4.4.6 School Year Start and End Dates (L5SCHBDD, L5SCHBMM, L5SCHBYY, L5SCHEDD, L5SCHEMM, L5SCHEYY)

The composite for school year start and end dates was taken from the school fact sheet (Q5, L5SYRSMM, L5SYRSDD, L5SYRSYY, L5SYREMM, L5SYREDD, L5SYREYY). If those data were missing, the values were taken from the FMS.

- L5SCHBDD L5 School Year Starting Date, Day
- L5SCHBMM L5 School Year Starting Date, Month
- L5SCHBYY L5 School Year Starting Date, Year
- L5SCHEDD L5 School Year Ending Date, Day
- L5SCHEMM L5 School Year Ending Date, Month
- L5SCHEYY L5 School Year Ending Date, Year

It should be noted that in past rounds, the composites for school year start and end dates were created differently because they were based on different questions. In past rounds, the question was in the student record abstract rather than the school fact sheet and was based on responses to multiple questions about start and end dates for school terms (e.g., semesters, trimesters). Composite variable names in past rounds started with a "U" prefix (rather than an "L" prefix as in spring-third grade) because they were taken from the student record abstract (variables for spring-first grade were U4SCHBDD, U4SCHBMM, U4SCHBYY, U4SCHEDD, U4SCHEMM, U4SCHEYY). If the start and end dates varied for children in the same school, the composite was created by using the school start and end dates reported for the majority of children in a school. Because school start and end dates were collected only once in the spring-third grade school fact sheet, discrepancies in questionnaire reports for children in the same school were not an issue.

7.4.5 Student Record Abstract and Field Management System Composite Variables

The composite variables created from the student record abstract and FMS data follow.

7.4.5.1 Year-Round Schools (F5YRRND)

This composite was created using data from the FMS. The FMS flag was “1” if the child was in year-round school. The values for the year-round school composite variable are 1 (Yes) and 2 (No). If the child was schooled at home, the composite was coded as -1 (Not Applicable).

7.4.5.2 Indicator of Whether Child Received Special Education Services (F5SPECS)

The composite variable F5SPECS indicates whether or not the child received special education services in the spring of third grade, based on the presence or absence of a link to a special education teacher in the FMS in spring-kindergarten. The values are 1 if the child received special education services, 2 if the child did not receive special education services, and -9 if the link was missing between the child and his or her teacher in the FMS.

7.4.5.3 Indicator of Whether Child Had an Individualized Education Plan (IEP) on Record at School (U5RIEP)

The variable U5RIEP indicates whether or not the child had an IEP or Individualized Family Service Plan (IFSP) on record at his/her school or another school in the spring of third grade. The values for the variable are 1 (child has an IEP/IFSP on record at his or her school, or at another school) and 2 (child does not have an IEP/IFSP on record at his or her school). If the information was missing, U5RIEP was coded as -9 (Not Ascertained).

7.4.6 Parent Identifiers and Household Composition (P5DADID, P5MOMID, P5HPARNT, P5HDAD, P5HMOM, P5HFAMIL, P5MOMTYP, P5DADTYP)

The construction of parent identifiers and the household composition variables from the parent interview data was a two-step process. First, individuals identifying themselves as the child’s mother/father were located within the household roster, and the type of their relationship to the child (biological, adoptive, foster, step, partner of parent, or unknown) was established. For households containing more than one father or mother, a hierarchy was used to designate the “current,” or residential, parent of each gender. The biological parent, if present, was always the current mother or father. In the absence of a biological parent, the current mother/father designation was assigned to the adoptive, step, foster/guardian, partner, or “unknown-type” parent. If there were more than one father or mother of the same type, the one with the lower person number on the household roster was selected. Person number refers to the number each household member has on the roster list. Household members are listed in the order they are reported by the respondent. Information about parents in the household, along with household size and presence or absence of grandparents, siblings, and other relatives was used to construct the household composition variables P5HPARNT, P5HDAD, P5HMOM, and P5HFAMIL and parent-type variables P5MOMTYP, and P5DADTYP.

After the residential parents were identified and the composite variables were constructed, in any household without a parent, the household respondent (and his or her spouse/partner, if applicable) was assigned as a “parent figure.” Parent demographic variables (including age, race/ethnicity, and education) were then constructed for all parents/parent figures. It should be noted, however, that these parent figures were not defined as parents (meaning biological, step-, adoptive, or foster) in the construction of the household composition composite variables described earlier. For example, for P5HFAMIL, composite values are as follows:

- 1=two parents and sibling(s)
- 2=two parents, no siblings
- 3=one parent and sibling(s)
- 4=one parent, no siblings
- 5=other

Parent figures were placed in the “other” category for this composite. Likewise, for the composite P5HPARNT, parent figures were placed in categories 8 or 9 for related and unrelated guardians, respectively. Similarly, parent figures were included in the category ‘no resident mother’ for P5HMOM and “no resident father” for P5HDAD. Thus, although persons reported as children’s parents/guardians and the spouses/partners of the parents/guardians were included in the definitions of all the household composites, individuals later identified as parent figures in households in which no parents were present were not considered to be parents in the coding of the household composites.

Some parent-specific variables do include persons who were later identified as parent figures. These are as follows (variables for fathers are listed below but those for mothers are created in the same way):

- P5DADID (Household roster number of resident father, male guardian, or father figure);
- P5HDAGE (Age of resident father, male guardian, or father figure);
- P5HDRACE (Race and ethnicity of the father, male guardian, or father figure in the household);
- P5HDEMP (The work status of the father, male guardian, or father figure in the household);
- P5DADOCC (Father, male guardian, or father figure's occupation);
- W3DADED (The father, male guardian, or father figure's highest level of education); and
- W3DADSCR (Father, male guardian, or father figure's occupation prestige score)

It should be noted that because the composite construction identifies only one resident mother or one resident father, same-sex parents are not readily identified in the composites themselves. Two approaches can be used to identify these couples. First, the user should search the relationship variables (P5REL_1, etc.) to identify households in which more than one person is identified as a father/mother to the focal child. Second, since not all same-sex partners identify themselves as "mother" or "father" to the focal child, the user should also search for households in which the respondent (identified by P5PER_1, etc.) is the child's parent and the respondent's spouse/partner (identified from P5SPOUSE) is the same sex as the respondent.

There are two sections in the parent interview that asked parent-figure-specific questions:

- PEQ Parent education
- EMQ Employment

Each of these sections was completed during the parent interview for up to two parents or parent figures. To indicate which household member or members were the subject of each section, "pointer" variables that hold the original number of the household member on the household roster were used. To illustrate how the pointer variables work, suppose there is a household with both a mother and a father who were listed third and fourth in the household roster. If household member #3, the mother, was the first person to receive the PEQ education section, then the pointer variable P5EDUP1 will equal “3.” The answers to the education questions for the mother will be contained in interview items in this section that end with the suffix “_1” (e.g., P5NDEG_1, P5DEGT_1, P5ENR_1, etc.). The suffix “_1” indicates that the data are for the first subject of the questions. Similarly, if household member #4, the father, was the second person to receive the PEQ education section, then the pointer variable P5EDUP2 will equal “4.” The answers to the education questions for the father will be contained in interview items in this section that end with the suffix “_2” (e.g., P5NDEG_2, P5DEGT_2, P5ENR_2, etc.). The suffix “_2” indicates that the data are for the second subject of the questions. Table 7-10 identifies the pointer variables.

7.4.7 Industry and Occupation Codes Used in ECLS-K

This section describes the aggregated categories that were used for coding occupation in the ECLS-K.

1. Executive, Administrative, and Managerial Occupations

This category includes senior-level and middle management occupations and occupations that directly support management. Senior-level managers are persons concerned with policymaking, planning, staffing, directing, and/or controlling activities. Middle managers include persons who plan, organize, or direct and/or control activities at the operational level. Workers in this category are not directly concerned with the fabrication of products or with the provision of services. Other officials and administrators include consultants, library directors, custom house builders, and location managers. Legislators are also included in this category.

2. Engineers, Surveyors, and Architects

The category includes occupations concerned with applying principles of architecture and engineering in the design and construction of buildings, equipment and processing systems, highways and roads, and land utilization.

3. Natural Scientists and Mathematicians

This category includes those engaged primarily in the application of scientific principles to research and development. Natural scientists are those in the physical sciences (e.g., chemistry, physics) and the life sciences (e.g., biology, agriculture, medicine). In addition, this category includes those in computer science, mathematics (including statistics), and operations research.

Table 7-10. Pointers to parent figure questions, spring-third grade: School year 2001–02

Person pointer[1]		Interview item	
P5EDUP1	P5 PEQ010-060 HH PERSON POINTER 1	P5NDEG_1	P5 PEQ010 PERS 1 COMPLETED NEW DEGREE
		P5DEGT_1	P5 PEQ020 PERS 1 DEGREE TYPE COMPLETED
		P5ENR_1	P5 PEQ030 IF PERS 1 ENROLLED IN COURSES
		P5FPT_1	P5 PEQ040 PERS 1 COURSE FULL/PART TIME
		P5TRN_1	P5 PEQ050 IF PERS 1 GETS JOB TRAINING
		P5HRTR_1	P5 PEQ060 PERS 1 HR/WK SPEND ON TRAINING
P5EDUP2	P5 PEQ010-060 HH PERSON POINTER 2	P5NDEG_2	P5 PEQ010 PERS 2 COMPLETED NEW DEGREE
		P5DEGT_2	P5 PEQ020 PERS 2 DEGREE TYPE COMPLETED
		P5ENR_2	P5 PEQ030 PERS 2 ENROLLED IN COURSES
		P5FPT_2	P5 PEQ040 PERS 2 COURSE FULL/PART TIME
		P5TRN_2	PR PEQ050 IF PERS 2 GETS JOB TRAINING
		P5HTR_2	PR PEQ060 PERS 2 HR/WK SPEND ON TRAINING
P5EMPP1	P5 EMQ010-150 HH PERSON POINTER 1	P5CHJB_1	P5 EMQ010 PERS 1 CHNGD JOB SNC SPR 2000
		P5PAY_1	P5 EMQ020 PERS 1 HAD PAID JOB LAST WEEK
		P5VAC_1	P5 EMQ030 IF PERS 1 ON LEAVE PAST WEEK
		P5JOB_1	P5 EMQ040 PERSON 1 NUMBER OF ALL JOBS
		P5HRS_1	P5 EMQ050 PERSON 1 HOURS/WK AT ALL JOBS
		P5LOK_1	P5 EMQ060 PERS 1 SOUGHT JOB LAST 4 WEEKS
		P5DO1_1	P5 EMQ070 PERS 1 CHKD W/PUB EMPL AGENCY
		P5DO2_1	P5 EMQ070 PERS 1 CHKD W/PRIV EMP AGENCY
		P5DO3_1	P5 EMQ070 PERS 1 CHKD W/EMPLOYER DIRECTLY
		P5DO4_1	P5 EMQ070 PERS 1 CHKD W/FRIENDS & REL
		P5DO5_1	P5 EMQ070 PERS 1 PLACED OR ANSWERED ADS
		P5DO6_1	P5 EMQ070 PERS 1 READ WANT ADS
		P5DO7_1	P5 EMQ070 PERS 1 DID SOMETHING ELSE
		P5DOW_1	P5 EMQ080 WHAT PERSON 1 DOING LAST WEEK
		P5TAK_1	P5 EMQ100 PERS 1 JOB AVAILABLE LAST WEEK
		P5OCC_1	P5 EMQ130-50 1ST PERSON OCCUPATION CODE
P5EMPP2	P5EMQ010-150 HH PERSON POINTER 2	P5CHJB_2	P5 EMQ010 PERS2 CHNGD JOB SNC SPRING 2000
		P5PAY_2	P5 EMQ020 PERS 2 HAD PAID JOB LAST WEEK
		P5VAC_2	P5 EMQ030 IF PERS 2 ON LEAVE PAST WEEK
		P5JOB_2	P5 EMQ040 PERSON 2 NUMBER OF ALL JOBS
		P5HRS_2	P5 EMQ050 PERSON 2 HOURS/WK AT ALL JOBS
		P5LOK_2	P5 EMQ060 PERS 2 SOUGHT JOB LAST 4 WEEKS
		P5DO1_2	P5 EMQ070 PERS 2 CHKD W/PUB EMPL AGENCY
		P5DO2_2	P5 EMQ070 PERS 2 CHKD W/PRIV EMP AGENCY
		P5DO3_2	P5 EMQ070 PERS 2 CHKD W/EMPLOYER DIRECTLY
		P5DO4_2	P5 EMQ070 PERS 2 CHKD W/FRIENDS & REL
		P5DO5_2	P5 EMQ070 PERS 2 PLACED OR ANSWERED ADS
		P5DO6_2	P5 EMQ070 PERS 2 READ WANT ADS
		P5DO7_2	P5 EMQ070 PERS 2 DID SOMETHING ELSE
		P5DOW_2	P5 EMQ080 WHAT PERSON 2 DOING LAST WEEK
		P5TAK_2	P5 EMQ100 PERS 2 JOB AVAILABLE LAST WEEK
		P5OCC_2	P5 EMQ130-50 2ND PERSON OCCUPATION CODE

[1] Pointer variables hold the original number of the household member on the household roster.
SOURCE: U.S. Department of Education, National Center for Education Statistics, Early Childhood Longitudinal Study, Kindergarten Class of 1998–99 third grade data collection, school year 2001–02.

4. **Social Scientists, Social Workers, Religious Workers, and Lawyers**

This category includes occupations concerned with the social needs of people and in basic and applied research in the social sciences.

5. **Teachers: College, University, and Other Postsecondary Institution; Counselors, Librarians, and Archivists**

This category includes those who teach at higher education institutions and at other postsecondary (after high school) institutions, such as vocational institutes. In addition, vocational and educational counselors, librarians, and archivists are included here.

6. **Teachers, except Postsecondary Institution**

This category includes prekindergarten and kindergarten teachers, elementary and secondary teachers, special education teachers, instructional coordinators, and adult education teachers (outside postsecondary).

7. **Physicians, Dentists, and Veterinarians**

This category includes health care professionals who diagnose and treat patients. In addition to physicians, dentists, and veterinarians, this category includes optometrists, podiatrists, and other diagnosing and treating professionals, such as chiropractors, hypnotherapists, and acupuncturists.

8. **Registered Nurses, Pharmacists, Dieticians, Therapists, and Physician's Assistants**

This category includes occupations concerned with the maintenance of health, the prevention of illness, and the care of the ill through the provision and supervision of nursing care; compounding drugs, planning food service or nutritional programs; providing assistance to physicians; and the provision of therapy and treatment as directed by physicians.

9. **Writers, Artists, Entertainers, and Athletes**

This category includes occupations concerned with creating and executing artistic works in a personally interpreted manner by painting, sculpturing, drawing, engraving, etching, and other methods; creating designs for products and interior decorations; designing and illustrating books, magazines, and other publications; writing; still, motion picture and television photography/filming; producing, directing, staging, acting, dancing, singing in entertainment; and participating in sports and athletics as a competitor or player and administering and directing athletic programs.

10. **Health Technologists and Technicians**

This category includes occupations concerned with providing technical assistance in the provision of health care. For example, clinical laboratory technologists and technicians, dental hygienists, radiologic technicians, licensed practical nurses (LPNs), and other health technologists are included here.

11. Technologists and Technicians, except Health

This category includes those providing technical assistance in engineering and scientific research, development, testing, and related activities, as well as operating and programming technical equipment and systems.

12. Marketing and Sales Occupations

This category includes occupations involving selling goods or services, purchasing commodities and property for resale, and conducting wholesale or retail business.

13. Administrative Support Occupations, including Clerks

This category includes occupations involving preparing, transcribing, transferring, systematizing, and preserving written communications and records; collecting accounts; gathering and distributing information; operating office machines and data processing equipment; operating switchboards; distributing mail and messages; and other support and clerical duties such as bank teller, data entry keyer, etc.

14. Service Occupations

The category includes occupations providing personal and protective services to individuals, and current maintenance and cleaning for building and residences. Some examples include food service, health service (e.g., aides or assistants), cleaning services other than household, and personal services.

15. Agricultural, Forestry, and Fishing Occupations

This category is concerned with the production, propagation (breeding/growing), gathering, and catching of animals, animal products, and plant products (timber, crop, and ornamental); the provision of services associated with agricultural production; and game farms, fisheries, and wildlife conservation. "Other agricultural and related occupations" include occupations concerned with the production and propagation of animals, animal products, plants, and products (crops and ornamental).

16. Mechanics and Repairers

Mechanics and repairers are persons who do adjustment, maintenance, part replacement, and repair of tools, equipment, and machines. Installation may be included if installation is usually done in conjunction with other duties of the repairers.

17. Construction and Extractive Occupations

This category includes occupations that normally are performed at a specific site, which will change over time, in contrast to production workers, where the work is usually at a fixed location. Construction workers include those in overall construction, brick masons, stonemasons, carpenters, electricians, drywall installers, paperhangers and painters, etc. Extractive occupations include oil well drillers, mining machine operators, and so on.

18. Precision Production Occupations

Precision production includes occupations concerned with performing production tasks that require a high degree of precision or attainment of rigid specification and operating plants or large systems. Examples are tool and die makers, pattern and model makers, machinists, jewelers, engravers, and so on. Also included are some food-related occupations including butchers and bakers. Plant and system operators include water and sewage, gas, power, chemical, petroleum, and other plant or system operators.

19. Production Working Occupations

This category includes occupations concerned with setting up, operating, and tending of machines and hand production work usually in a factory or other fixed place of business.

20. Transportation and Material Moving Occupations

This category includes occupations concerned with operating and controlling equipment used to facilitate the movement of people or materials and the supervising of those workers.

21. Handlers, Equipment Cleaners, Helpers, and Laborers

This category includes occupations that involve helping other workers and performing routine nonmachine tasks. A wide variety of helpers, handlers, etc., are included in this category. Examples include construction laborers, freight, stock, and material movers, garage and service station related occupations, parking lot attendants, and vehicle washers and equipment cleaners.

22. Unemployed, Retired, Disabled, or Unclassified Workers

This category includes persons who are unemployed, have retired from the work force, or are disabled. It also includes unclassified occupations that do not fit into the categories above (e.g., occupations that are strictly military, such as “tank crew member” and “infantryman”).

7.5 Methodological Variables

To facilitate methodological research, eleven new variables were added to the third grade file. The identifiers for parent interview work area (F5PWKARE), parent interviewer (F5PINTVR), child assessment work area (F5CWKARE), and child assessor (F5CASSOR) were extracted from the field management system.

Start and end times for both the child assessment (C5ASMSTM, C5ASMETM) and the parent interview (P5INTSTM, P5INTETM) were created from keystroke-by-keystroke records of each parent interview and child assessment. All four are text variables in the form *MM/DD/YY hour:minute:second AM/PM*. It should be noted that there may be more than one attempt to complete an interview or assessment, and those attempts could span several days. For example, an interviewer could begin a parent interview on one evening and complete the remainder of the interview several days later. For this reason, variables to indicate the number of attempts necessary to complete the parent interview (P5ATTMPT) and the number of attempts necessary to complete the child assessment (C5APPMPT) have also been included on the file.

Finally, an indicator variable (F5PREFCV, Parent Interview Refusal Conversion) was created to flag cases that had, at any time, refused to respond to the parent interview but then agreed to participate. The values for F5PREFCV are 1=YES (refused but were converted to be a participant) and 2=NO (did not refuse).

7.6 Children Who Changed Schools

There are several variables in the file that can be used to determine if a child moved to a different school between rounds or moved to a different school during the third grade data collection period.

7.6.1 Children Who Changed Schools During Third Grade Data Collection

The variable S5_ID is a school identification number that indicates which school the child was in at the end of the third grade data collection. There is another school ID variable, S5_ST_ID, that indicates where the child was at the beginning of the round. By comparing school ID variables, users can determine whether the child physically moved from one school to another during round 5. For the vast majority of the children these two variables will be identical, but for those who moved during the data collection period they will be different. If it was not known where the child was at the beginning or the end of the round, the scheme shown in table 7-11 for assigning ID numbers was used.[9] In previous rounds of the study, the code "9994" was used to indicate that the student was deceased. This code was not needed in third grade and thus does not appear on the spring-third grade file.

Table 7-11. Case status and school ID numbers for children not followed or located, spring-third grade: School year 2001–02

Case status	School identification variables S5_ID/S5_ST_ID
Not in the United States. The student now lives outside the U.S.	9993
Unlocatable. Field staff were unable to locate a transfer student in his/her new school.	9995
End of field period. Information on the transfer student's new school was identified too late in the field period for the case to be re-fielded for the assessment.	9996
Moved to nonsampled PSU. The transfer student enrolled in a school that was outside of ECLS-K's sampled PSUs—field staff did not attempt to collect the assessment but did attempt to collect the parent interview.	9997
Do not follow. The transfer student was flagged by the statisticians as "do not follow" because of subsampling of transfer students due to cost constraints. If the child moved from his/her original school, field staff did not "follow" him or her to the new school and did not collect a child assessment or parent interview.	9998

SOURCE: U.S. Department of Education, National Center for Education Statistics, Early Childhood Longitudinal Study, Kindergarten Class of 1998–99 third grade data collection, school year 2001–02.

In addition, there are other variables on the file that identify the origin of data for children who moved within the current round. These are helpful, for example, if a child changed schools within the 2001–02 school year and there are data from the first school the child attended (the starting school for the round, S5_ST_ID) but not the second school (the ending school for the round, S5_ID). The procedures for locating children within schools were as follows: Children's schools were identified at the beginning of the 2001–02 school year. If data collectors went to a school in the spring and found that a child was no longer there, the child was followed to his/her transfer school and data collection was attempted at the new school. If data were obtained from the new school, they were used. If data were not obtained and teacher or school data from the first school were available, those data were used.

[9] It should be noted that there were some children who could not be located during the field period for the beginning of the round but were located during the field period for the end of the round. Children who could not be located at the beginning of the round, but who were located and enrolled in sampled schools at the end of the round, will have S5_ST_ID values that begin with 999 and S5_ID values that are ID numbers for schools. Others who could not be located at the beginning or end round will have 999 codes for both S5_ST_ID and S5_ID.

The names and labels for the variables that can be used to determine whether data came from starting or ending schools are as follows:

- A5_T_ID A5 TQA RESPONDING TEACHER ID
- B5_T_ID B5 TQB RESPONDING TEACHER ID
- T5_T_ID T5 TQC RESPONDING TEACHER ID
- L5_S_ID L5 SFS RESPONDING SCHOOL ID
- S5_S_ID S5 SAQ RESPONDING SCHOOL ID
- K5_S_ID K5 FACILITY CHECKLIST RESPONDING SCHL ID
- D5_T_ID D5 SPECIAL ED A RESPONDING TEACHER ID
- E5_T_ID E5 SPECIAL ED B RESPONDING TEACHER ID
- U5_S_ID U5 SRA RESPONDING SCHOOL ID
- F5NOTEND F5 SOME DATA FR OTHER THAN ENDING SOURCE

The first 9 variables are instrument-specific ID numbers that indicate the school or teacher identification number from which the child or school data were obtained. For users interested in the source of the data for children who moved within the round, they should match the original school or teacher IDs with the first 9 instrument-specific IDs above to determine if the data for a given instrument were collected from the starting school or the ending school. The flag F5NOTEND can also be used to identify children for whom some data were from a school or teacher other than the ending school or ending teacher. There are 147 children with “F5NOTEND=TRUE” flags.

The following scenarios illustrate how these variables may appear in the files:

- Child has same starting and ending school. In this case, the data from the school and teachers are all from the same location and the 9 instrument-specific IDs above will match the appropriate ending school and teacher IDs. The “overall” IDs containing the correct end-of-round data are S5_ID (school), T5_ID (regular teacher), and D5T_ID (special education teacher). Thus, if the child had the same starting and ending school, S5_S_ID would match S5_ID; A5_T_ID, B5_T_ID, and T5_T_ID would match T5_ID, etc.

- Child has same starting and ending school, no data from either source. In this case, the school/teacher IDs will be from the ending school. The 9 instrument-specific IDs above will be system missing, as will the rest of the data from those instruments. For instance, most children do not have a special education teacher. The special education data are all missing, including the Special Education A ID (D5_T_ID) and the Special Education B ID (E5_T_ID). The same is true if there are no data from other sources, such as the teacher questionnaire A. If there were no hard-copy data (but there was a child assessment or parent interview), then all the hard-copy instrument data including the instrument IDs would be system missing.

- Child has different starting and ending schools, all data from ending school. The 9 instrument-specific school/teacher IDs above will be from the ending school.

- Child has different starting and ending schools, some data from ending school, some data from starting school. The school ID (S5_ID) or teacher ID will be from the ending school, and each of the 9 instrument-specific flags will indicate the ending school or the starting school depending on which school was the source of the instrument.

- Child has different starting and ending schools, ending school is unknown, child has a few instruments from starting school. If there is nonresponse from the ending school, S5_ID will be coded with a "999n" number described in table 7-11 to appropriately reflect the type of nonresponse. If there are instrument-specific data, the instrument IDs will be filled. If not, the instrument-specific IDs will be blank.

- Child has same starting and ending school. User is interested in whether there are special education data for the child. If there is no special education teacher, D5T_ID will be missing. If there is a special education teacher, D5T_ID will be filled. In either case, it should be noted that there could be missing data for special education data in the part B questionnaire. It is left to users to determine how they would like to set "Not Applicable" versus "Not Ascertained" codes for such combinations. For users interested in links to special education services, regardless of whether the source of the information was the starting or ending school, the composite variable F5SPECS that is based on information from the FMS system rather than the receipt of particular special education questionnaires can be used.

7.6.2 Children Who Changed Schools Between Rounds (R5DEST, R5R4SCHG)

Children moved between schools for a variety of reasons, but one factor was that a school terminated before the third grade and most of the students went to third grade at another particular school. This is known as a "destination school" and the move is known as a "destination move." Destination schools were schools for which it was determined (during a fall data collection conducted to locate children prior to the spring data collection) that at least four ECLS-K children would move into them from a school that ended before the third grade or a school that had closed. The variable on the file that indicates destination moves is R5DEST (moved to a spring-third grade destination school).

It should be noted that the destination school may also have been an originally sampled school; in this case, the school was a destination school only for the new students, not for the originally sampled students. The variable R5DEST was set to 1 (True) if a child made a school change and destination move to a spring-third grade destination school. If a child did not move to a spring-third grade destination school or did not move between schools at all, the composite is coded 0 (False). If the data are missing about whether the move was a destination move, the composite was coded -9 (Not Ascertained). If the child was schooled at home, the composite was coded as -1 (Not Applicable).

Another variable on the file that will be of interest to users examining school change is R5R4SCHG (school type change between spring-first grade and spring-third grade). It is used in the creation of R5DEST (R5R4SCHG must indicate a school change for R5DEST to be set to "1"). It indicates whether the child changed schools and, if so, what the school type was in the previous and new school (e.g., whether the change was from public to private school, private to private school, etc.). R5R4SCHG is created by comparing the school IDs from spring-first grade and spring-third grade for children who were in the spring-first grade data collection. A difference in IDs indicated a change. If there was no difference in ID's, R5R4SCHG was coded 1 (child did not change schools). For children who changed schools, the spring-first grade school type variable S4SCTYP was compared to the spring-third grade school type variable S5SCTYP. Categories were assigned as appropriate (2 = child transferred from public to public; 3 = child transferred from private to private; 4 = child transferred from public to private; 5 = child transferred from private to public; and 6 = child transferred, other). Category six was used for those children who transferred schools, but school type was unknown. Children who were not in the spring-first grade data collection were coded "Not Ascertained" on R5R4SCHG. Children who were home-schooled in spring-first grade or spring-third grade were coded "Not Applicable" for R5R4SCHG.

7.7 Composite Table

Table 7-12 describes the composite and derived variables that are on the ECLS-K child catalog. Please note that a few of the variables specified in the "derived from" column are intermediary variables that were not included in the final data sets. An example of an intermediary variable is the child gender variable from the parent questionnaire, CHILDGEN. If this variable was missing, or had conflicting information across rounds of the study, information about gender was used from the Field Management System (FMS) or child report. The variable CHILDGEN is not included in the final data set, but the composite R5GENDER is included. Other intermediary variables are taken from either the FMS or the school master file and are not included on the data file.

The "derived from" column also contains the item numbers from the questionnaire, which help in identifying the items that were used in the creation of these composites. This information allows a user to decide if he or she would like to use the composite based on how it was defined.

Some variables in table 7-12 have been recoded or suppressed. Reasons for these data changes are discussed in section 7.8. The new recoded categories are noted for applicable variables in table 7-12.

7.8 Masked Variables

All the variables from the ECLS-K restricted-use file are included in the same order on the ECLS-K public-use file. For some of the variables, certain categories were modified. The value labels for those masked variables were updated from the restricted-use variables to reflect the new categories that were created during the masking process.

Variables on the restricted-use files were modified in different ways based on the disclosure analysis NCES conducted in order to protect the identity of the respondents and children. There are several types of modifications on the public-use files.

- Outliers are top- or bottom- coded to prevent identification of unique schools, teachers, parents, and children without affecting overall data quality.
- Certain schools identified as at risk for disclosure have a 5 to 10 percent noise introduced in those variables that pose a risk for disclosure.
- Variables with too few cases and a sparse distribution are suppressed in the public-use files. The values for these variables were set to -2 and labeled "suppressed" in the ECB.
- For one group of variables, values were modified by "data swapping." This process removes a reported value and replaces it with a reported value from a different respondent for a subset of the records.

Table 7-12. Spring-third grade composite variables: School year 2001–02

Variable name	Category	Description	Derived from	Values
R5AGE	Child	Child's age in months at the time the direct child assessment occurred	R5DOBMM, R5DOBDD, R5DOBYY (composites), assessment date (taken from assessment audit trails, variables not on file)	Recoded from a continuous variable to the following: 1=Less than 105 2=105 to less than 108 3=108 to less than 111 4=111 to less than 114 5=114 to less than 117 6=117 or more
R5GENDER	Child	Child's gender	CHILDGEN (INQ.016, not on file), FMS (variable not on file), INGENDER from child report (variable not on file), GENDER (composite from previous rounds)	1=Male; 2=Female
R5DOBMM	Child	Child's date of birth month	DOBMM, CHILDDOB (not on file) from first data collection in which reported in parent interview, INDOB from child report (variable not on file), FMS date of birth variable (not on file)	1-12
R5DOBDD	Child	Child's date of birth day	DOBDD, CHILDDOB (not on file) from first data collection in which reported in parent interview, INDOB from child report (variable not on file), FMS date of birth variable (not on file)	1-31
R5DOBYY	Child	Child's date of birth year	DOBYY, CHILDDOB (not on file) from first data collection in which reported in parent interview, INDOB from child report (variable not on file), FMS date of birth variable (not on file)	Recoded from values of 1990-1995 to a minimum value of 1992 and a maximum value of 1993

See note at end of table.

Table 7-12. Spring-third grade composite variables—Continued

Variable name	Category	Description	Derived from	Values
W3RACETH	Child	Race and ethnicity of the focal child	W1RACETH, WKRACETH (composites), RACE1, RACE2, RACE3, RACE4, RACE5, RACE6 (variables coded in parent interview based on P5RC1_1 through P5RC6_1 up to P5RC1_25 through P5RC6_25 (FSQ.195), and P5HSP_1 through P5HSP_25 (FSQ.190))	1=White, 2=Black or African American, 3=Hispanic, race specified, 4=Hispanic, no race specified, 5=Asian, 6=Native Hawaiian or other Pacific Islander, 7=American Indian or Alaska Native, 8=More than 1 race, non-Hispanic
R5RACE	Child	Child race and ethnicity	W3RACETH, W1RACETH, WKRACETH, RACE from previous round (composites), C_RACE (FMS, not on file), HI_PSU (FMS, not on file)	1=White, 2=Black or African American, 3=Hispanic, race specified, 4=Hispanic, no race specified, 5=Asian, 6=Native Hawaiian or other Pacific Islander, 7=American Indian or Alaska Native, 8=More than 1 race, non-Hispanic
W3AMERIN	Child	Child is American Indian or Alaska Native	W1AMERIN, WKAMERIN (composites), RACE1, RACE2, RACE3, RACE4, RACE5, RACE6 (variables coded in parent interview based on P5RC1_1 through P5RC6_1 up to P5RC1_25 through P5RC6_25 (FSQ.195), and P5HSP_1 through P5HSP_25 (FSQ.190))	1=Yes, 2=No
W3ASIAN	Child	Child is Asian	W1ASIAN, WKASIAN (composites), RACE1, RACE2, RACE3, RACE4, RACE5, RACE6 (variables coded in parent interview based on P5RC1_1 through P5RC6_1 up to P5RC1_25 through P5RC6_25 (FSQ.195), and P5HSP_1 through P5HSP_25 (FSQ.190))	1=Yes, 2=No

See note at end of table.

Table 7-12. Spring-third grade composite variables—Continued

Variable name	Category	Description	Derived from	Values
W3BLACK	Child	Child is African American	W1BLACK, WKBLACK (composites), RACE1, RACE2, RACE3, RACE4, RACE5, RACE6 (variables coded in parent interview based on P5RC1_1 through P5RC6_1 up to P5RC1_25 through P5RC6_25 (FSQ.195), and P5HSP_1 through P5HSP_25 (FSQ.190))	1=Yes, 2=No
W3PACISL	Child	Child is Native Hawaiian or other Pacific Islander	W1PACISL, WKPACISL (composites), RACE1, RACE2, RACE3, RACE4, RACE5, RACE6 (variables coded in parent interview based on P5RC1_1 through P5RC6_1 up to P5RC1_25 through P5RC6_25 (FSQ.195), and P5HSP_1 through P5HSP_25 (FSQ.190))	1=Yes, 2=No
W3WHITE	Child	Child is White	W1WHITE, WKWHITE (composites), RACE1, RACE2, RACE3, RACE4, RACE5, RACE6 (variables coded in parent interview based on P5RC1_1 through P5RC6_1 up to P5RC1_25 through P5RC6_25 (FSQ.195), and P5HSP_1 through P5HSP_25 (FSQ.190))	1=Yes, 2=No
W3MT1RAC	Child	Child is more than one race	W1MT1RAC, WKMT1RAC (composites), RACE1, RACE2, RACE3, RACE4, RACE5, RACE6 (variables coded in parent interview based on P5RC1_1 through P5RC6_1 up to P5RC1_25 through P5RC6_25 (FSQ.195), and P5HSP_1 through P5HSP_25 (FSQ.190))	1=Yes, 2=No
W3HISP	Child	Child is Hispanic	W1HISP, WKHISP (composites), RACE1, RACE2, RACE3, RACE4, RACE5, RACE6 (variables coded in parent interview based on P5RC1_1 through P5RC6_1 up to P5RC1_25 through P5RC6_25 (FSQ.195), and P5HSP_1 through P5HSP_25 (FSQ.190))	1=Yes, 2=No

See note at end of table.

Table 7-12. Spring-third grade composite variables—Continued

Variable name	Category	Description	Derived from	Values
C5BMI	Child	Child’s spring-third grade body mass index	C5HEIGHT, C5WEIGHT (composites)	Continuous
C5HEIGHT	Child	Child’s spring-third grade composite height	HEIGHTIN, HEIGHTI2 (not on file)	Continuous
C5WEIGHT	Child	Child’s spring-third grade composite weight	CHILDWEI, CHILDWE2 (not on file)	Continuous
P5DISABL	Child	Child currently has a disability	P5DIAGNO (CHQ.050), P5PROFFD (CHQ.110), P5COMMU2 (CHQ.170), P5DIFFH3 (CHQ.210), P5VISIO2 (CHQ.300), P5RSVTSY (CHQ520), P5DIABEH (CHQ.335), P5DIAEMP (CHQ.360)	1=Yes, 2=No
P5CARNOW	Child	Focal child is currently receiving any nonparental care	P5RELNOW (CCQ.010), P5NRNOW (CCQ.150), P5CTRNOW (CCQ.260)	1=Yes, 2=No
P5HRSNOW	Child	Total number of hours per week the focal child currently spends in all nonparental child care	P5RHRS (CCQ.090), P5NHRS (CCQ.190), P5CHRS (CCQ.355), P5RELNOW (CCQ.010), P5RELNUM (CCQ.060), P5RHROTH (CCQ.140), P5NRNOW (CCQ.150), P5NRNUM (CCQ.165), P5NHROTH (CCQ.250), P5CTRNOW (CCQ.260), P5CTRNUM (CCQ.325), P5CHROTH (CCQ403), P5RWEEK (CCQ.080), P5NWEEK (CCQ.180), P5CWEEK (CCQ.340)	Continuous
P5NUMNOW	Child	Total number of all types of care arrangements the focal child currently has on a regular basis	P5RELNUM (CCQ.060), P5NRNUM (CCQ.165), P5CTRNUM (CCQ.325), P5RELNOW (CCQ.010), P5NRNOW (CCQ.150), P5CTRNOW (CCQ.260)	Continuous

See note at end of table.

Table 7-12. Spring-third grade composite variables—Continued

Variable name	Category	Description	Derived from	Values
P5PRIMNW	Child	Primary, regular, nonparental child care arrangement in which the child currently spends the most hours per week	P5HRSNOW (composite), P5RHRS (CCQ.090), P5NHRS (CCQ.190), P5RPLACE (CCQ.070), P5NPLACE (CCQ.170), P5CHRS (CCQ.355)	0=No nonparental care, 1=Relative care in child's home, 2=Relative care in another home 3=Nonrelative care in child's home, 4=Nonrelative care in another home, 5=Center-based program, 6=2 or more programs, 7 = Location of care varies
F5SPECS	Child	This variable indicates whether or not the child received special education services based on the presence or absence of a link to a special education teacher in the FMS.	T_ID and TYPE (FMS variables not on file)	1 = Child got special education services 2 = Child did not get special education services
U5RIEP	Child	This variable indicates whether or not the child has an Individualized Education Program (IEP) or Individualized Family Service Plan (IFSP) on record at his/her school or at another school according to information from the student record abstract.	U5IEP (Student Record Abstract item 8)	1=Child has IEP/IFSP on record at his/her school or another school 2=Child does not have an IEP/IFSP
R5DEST	Child	Moved to spring-third grade destination school	DESTSCH (School Master file variable not on file), R5R4SCHG	1=Yes 2=No
R5R4SCHG	Child	School type change between spring-first grade and spring-third grade	School ID, S5SCTYP, S4SCTYP, S3SCTYP, S2KSCTYP	1=Child did not change schools 2=Child transferred from public to public 3=Child transferred from private to private 4=Child transferred from public to private 5=Child transferred from private to public 6=Child transferred, other

See note at end of table.

Table 7-12. Spring-third grade composite variables—Continued

Variable name	Category	Description	Derived from	Values
R5ELIG	Child	Eligibility status of child	Child raw assessment status, ASSESSME (not on file)	1=Eligible, 2=Ineligible, out of scope, 3=Ineligible, moved out of the country, 4=Ineligible, deceased, 5=Ineligible, mover not followed
C5ASMTST	Child	Child assessment status	C5RDGFLG, C5MTHFLG, C5SCIFLG, statistical flag SCORE_FG (not on file)	1=Completely scorable assessment data, 2=Partially completed scorable assessment data, 3=Category not in use in round 5; 4=Child with disability, not assessed, 5=Nonrespondent
P5MOMID	Family/HH	Household roster number of resident mother, female guardian, or mother figure	P5REL_1 to P5REL_25 (FSQ.130), P5UNR_1 to P5UNR_25 (FSQ.180), P5SPOUSE (FSQ.120), P5MOM_1 through P5MOM_25 (FSQ.140)	1-25
P5DADID	Family/HH	Household roster number of resident father, male guardian, or father figure	P5REL_1 to P5REL_25 (FSQ.130), P5UNR_1 to P5UNR_25 (FSQ.180), P5SPOUSE (FSQ.120), P5DAD_1 through P5DAD_25 (FSQ.150)	1-25
P5HPARNT	Family/HH	Classification of the focal child's parents who reside in the household	P5REL_1 through P5REL_25 (FSQ.130), P5UNR_1 through P5UNR_25 (FSQ.180), P5HMOM, P5HDAD (composites)	1=Biological mother and biological father, 2=Biological mother and other father (step-, adoptive, foster), 3=Biological father and other mother (step-, adoptive, foster), 4=Biological mother only, 5=Biological father only, 6=Two adoptive parents, 7=Single adoptive parent or adoptive parent and stepparent, 8=Related guardian(s), 9=Unrelated guardian(s)

See note at end of table.

Table 7-12. Spring-third grade composite variables—Continued

Variable name	Category	Description	Derived from	Values
P5HFAMIL	Family/HH	Family type categories using both parent and sibling information	P5REL_1 through P5REL_25 (FSQ.130), P5UNR_1 through P5UNR_25 (FSQ.180), P5HMOM, P5HDAD, P5NUMSIB (composites)	1=Two parents and sibling(s), 2=Two parents, no siblings, 3=One parent and sibling(s), 4=One parent, no siblings, 5=Other
P5NUMSIB	Family/HH	Total number of siblings with whom the focal child lives, including anyone reporting him/herself as the child of the focal child's foster parent/guardian	P5REL_1 to P5REL_25 (FSQ.130)	Continuous
P5LESS18	Family/HH	Total number of household members younger than 18 years old	HHNUMBER and HH18ANDOVER (parent interview flags not on file)	Continuous
P5OVER18	Family/HH	Total number of household members age 18 or older	HH18ANDOVER (parent interview flags not on file)	Continuous
P5HTOTAL	Family/HH	Total number of household members	HHNUMBER (parent interview flag not on file)	Continuous
P5TWIN	Family/HH	Household has sampled twins	P5PER_1 to P5PER_25 (person type in FSQ roster)	0=No twin in HH, 1=Twin in HH
W3POVRTY	Family/HH	Poverty indicator	P5HILOW (PAQ.100), P5INCCAT(PAQ.110), W3INCCAT, P5HTOTAL (composites), and Census-defined thresholds	1=Below poverty threshold, 2=At or above poverty threshold

See note at end of table.

Table 7-12. Spring-third grade composite variables—Continued

Variable name	Category	Description	Derived from	Values
W3INCCAT	Family/HH	Household income	P5INCCAT(PAQ.110)	1=$5,000 or less 2=$5,001 to $10,000 3=$10,001 to $15,000 4=$15,001 to $20,000 5=$20,001 to $25,000 6=$25,001 to $30,000 7=$30,001 to $35,000 8=$35,001 to $40,000 9=$40,001 to $50,000 10=$50,001 to $75,000 11=$75,001 to $100,000 12=$100,001 to $200,000 13=$200,001 or more
W3SESL	Family/HH	Socioeconomic scale	W3INCCAT, W3MOMED, W3DADED, W3MOMSCR, W3DADSCR (all composites)	Continuous
W3SESQ5	Family/HH	Quintile indicator for W3SESL	W3SESL	1=First quintile (lowest), 2=Second quintile, 3=Third quintile, 4=Fourth quintile, 5=Fifth quintile (highest)
W3PARED	Family/HH	Highest level of education for the child's parents or non-parental guardians who reside in the household. If only one parent or guardian resides in the household, W3PARED reflects that parent's education level.	W3MOMED, W3DADED (composites)	1=8th grade or below, 2=9th to 12th grades, 3=High school diploma/equivalent, 4=Voc/Tech program, 5=Some college, 6=Bachelor's Degree, 7=Graduate/professional school/no degree, 8=Master's degree, 9=Doctorate or professional degree

See note at end of table.

Table 7-12. Spring-third grade composite variables—Continued

Variable name	Category	Description	Derived from	Values
W3MOMSCR	Family/HH	Mother, female guardian, or mother figure's occupation GSS prestige score	1989 GSS prestige scores, EMQ.120 (not on file), EMQ.130 (not on file), and EMQ.140 (not on file).	29.6 Handler, Equip, Cleaner, Helpers, Labor; 33.42 Production Working Occupation; 34.95 Service Occupations; 35.63 Agriculture, Forestry, Fishing Occupations; 35.78 Marketing & Sales Occupation; 35.92 Transportation, Material Moving; 37.67 Precision Production Occupation; 38.18 Administrative Support, Including Clerk; 39.18 Mechanics & Repairs; 39.2 Construction & Extractive Occupations; 48.69 Technologists, Except Health; 52.54 Writers, Artists, Entertainers, Athletes; 53.5 Executive, Admin, Managerial Occupation; 57.83 Health Technologists & Technicians; 59 Social Scientist/Workers, Lawyers; 61.56 Registered Nurses, Pharmacists; 62.87 Natural Scientists & Mathematicians; 63.43. Teacher, Except Postsecondary; 64.89 Engineers, Surveyors, & Architects; 72.1 Teachers; College, Postsecondary Counselors, Librarians; 77.5 Physicians, Dentists, Veterinarians
P5HDAD	Family/HH	Indicates whether the birth, adoptive, step-, or foster father of the focal child resides in the household with the focal child	P5REL_1 through P5REL_25(FSQ.130), P5DAD_1 through P5DAD_25 (FSQ.150), P5UNR_1 through P5UNR_25 (FSQ.180), P5PARTNR (FSQ.110), P5SPOUSE (FSQ.120)	1=Biological, 2=Adoptive, 3=Step, 4=Foster, 5=Partner, 6=Don't know type, 7= No resident father
P5HDAGE	Family/HH	Age of resident father, male guardian, or father figure	P5AGE_1 through P5AGE_25 (FSQ.030), P5DADID	Continuous

See note at end of table.

Table 7-12. Spring-third grade composite variables—Continued

Variable name	Category	Description	Derived from	Values
P5HDRACE	Family/HH	Race and ethnicity of the father, male guardian, or father figure in the household	RACE1, RACE2, RACE3, RACE4, RACE5, RACE6 (variables coded in parent interview based on P5RC1_1 through P5RC6_1 up to P5RC1_25 through P5RC6_25 (FSQ.195), and P5HSP_1 through P5HSP_25 (FSQ.190))	1=White, 2=Black or African American, 3=Hispanic, race specified, 4=Hispanic, no race specified, 5=Asian, 6=Native Hawaiian or other Pacific Islander, 7=American Indian or Alaska Native, 8=More than one race, non-Hispanic
W3DADSCR	Family/HH	Father, male guardian, or father figure's occupation GSS prestige score	1989 GSS prestige scores, EMQ.120, EMQ.130, and EMQ.140 (not on file)	29.6 Handler, Equip, Cleaner, Helpers, Labor; 33.42 Production Working Occupation; 34.95 Service Occupations; 35.63 Agriculture, Forestry, Fishing Occupations; 35.78 Marketing & Sales Occupation; 35.92 Transportation, Material Moving; 37.67 Precision Production Occupation; 38.18 Administrative Support, Including Clerk; 39.18 Mechanics & Repairs; 39.2 Construction & Extractive Occupations; 48.69 Technologists, Except Health; 52.54 Writers, Artists, Entertainers, Athletes; 53.5 Executive, Admin, Managerial Occupation; 57.83 Health Technologists & Technicians; 59 Social Scientist/Workers, Lawyers; 61.56 Registered Nurses, Pharmacists; 62.87 Natural Scientists & Mathematicians; 63.43. Teacher, Except Postsecondary; 64.89 Engineers, Surveyors, & Architects; 72.1 Teachers; College, Postsecondary Counselors, Librarians; 77.5 Physicians, Dentists, Veterinarians

See note at end of table.

Table 7-12. Spring-third grade composite variables—Continued

Variable name	Category	Description	Derived from	Values
W3DADED	Family/HH	The father, male guardian, or father figure's highest level of education	P5HIG_1 through P5HIG_2 (PEQ.020), P5HIS_1 through P5HIS_2 (PEQ.030), P5NDEG_1 through P5NDEG_2 (PEQ.010)	1=8th grade or below, 2=9th to 12th grades, 3=High school diploma/equivalent, 4=Voc/Tech program, 5=Some college, 6=Bachelor's Degree, 7=Graduate/professional school/no degree, 8=Master's degree, 9=Doctorate or professional degree
P5HDEMP	Family/HH	The work status of the father, male guardian, or father figure in the household.	P5HRS_1, _2 (EMQ.050), P5PAY_1, _2 (EMQ.020), P5VAC_1, _2 (EMQ 030), P5LOK_1, _2 (EMQ.060), P5DO1_1, _2 (EMQ.070), P5DO2_1, _2 (EMQ.070), P5DO3_1, _2 (EMQ.070), P5DO4_1, _2 (EMQ.070), P5DO5_1, _2 (EMQ.070), P5DO6-1, _2 (EMQ.070), P5D07_1, _2 (EMQ.070), P5CHJB_1, _2 (EMQ.010)	1=35 hours or more per week, 2=Less than 35 hours per week, 3=Looking for work, 4=Not in the labor force
P5DADOCC	Family/HH	Father, male guardian, or father figure's occupation	Combination of P5CHJB_1, _2 (EMQ.010), EMQ.120, EMQ.130, and EMQ.140 (not on file)	01 Executive, Admin, Managerial Occupation 02 Engineers, Surveyors, & Architects 03 Natural Scientists & Mathematicians 04 Social Scientists/Workers, Lawyers 05 University Teachers, Postsecondary Counselors, Librarians 06 Teachers, except postsecondary 07 Physicians, Dentists, Veterinarians 08 Registered Nurses, Pharmacists 09 Writers, Artists, Entertainers, Athletes 10 Health Technologists & Technicians 11 Technologists, except Health 12 Marketing & Sales Occupation 13 Administrative Support, incl. Clerk 14 Service Occupations 15 Agriculture, Forestry, Fishing

See note at end of table.

Table 7-12. Spring-third grade composite variables—Continued

Variable name	Category	Description	Derived from	Values
P5DADOCC (continued)	Family/ HH	Father, male guardian, or father figure's occupation	Combination of P5CHJB_1, _2 (EMQ.010), EMQ.120, EMQ.130, and EMQ.140 (not on file)	16 Mechanics & Repairs 17 Construction & Extractive Occupations 18 Precision Production Occupation 19 Production Working Occupation 20 Transportation, Material Moving 21 Handler, Equip, Cleaner, Helpers, Labor 22 Unemployed or Retired
P5HMOM	Family/HH	Indicates whether the birth, adoptive, step-, or foster mother of the focal child resides in the household with the focal child	P5REL_1 through P5REL_25(FSQ.130), P5MOM_1 through P5MOM_25 (FSQ.140), P5UNR_1 through P5UNR_25 (FSQ.180), P5PARTNR (FSQ.110), P5SPOUSE (FSQ.120)	1=Biological, 2=Adoptive, 3=Step, 4=Foster, 5=Partner, 6=Don't know type, 7=No resident mother
P5HMAGE	Family/HH	Age of resident mother, female guardian, or mother figure	P5AGE_1 through P5AGE_25 (FSQ.030), P5MOMID	Continuous
P5HMRACE	Family/HH	Race and ethnicity of the mother, female guardian, or mother figure in the household	RACE1, RACE2, RACE3, RACE4, RACE5, RACE6 (These variables are coded in parent interview—see W3RACETH specs for details. The original race variables are P5RC1_1 through P5RC6_1 up to P5RC1_25 through P5RC6_25 (FSQ.195), and P5HSP_1 through P5HSP_25 (FSQ.190)).	1=White, 2=Black or African American, 3=Hispanic, race specified, 4=Hispanic, no race specified, 5=Asian, 6=Native Hawaiian or other Pacific Islander, 7=American Indian or Alaska Native, 8=More than one race, non-Hispanic
W3MOMED	Family/HH	Mother, female guardian, or mother figure's highest level of education	P5HIG_1 through P5HIG_2 (PEQ.020), P5HIS_1 through P5HIS_2 (PEQ.030), P5NDEG_1 through P5NDEG_2 (PEQ.010)	1=8th grade or below, 2=9th to 12th grades, 3=High school diploma/equivalent, 4=Voc/Tech program, 5=Some college, 6=Bachelor's Degree, 7=Graduate/professional school/no degree, 8=Master's degree, 9=Doctorate or professional degree

See note at end of table.

Table 7-12. Spring-third grade composite variables—Continued

Variable name	Category	Description	Derived from	Values
P5HMEMP	Family/HH	The work status of the mother, female guardian, or mother figure in the household	P5HRS_1, _2 (EMQ.050), P5PAY_1, _2 (EMQ.020), P5VAC_1, _2 (EMQ 030), P5LOK_1, _2 (EMQ.060), P5DO1_1, _2 (EMQ.070), P5DO2_1, _2 (EMQ.070), P5DO3_1, _2 (EMQ.070), P5DO4_1, _2 (EMQ.070), P5DO5_1, _2 (EMQ.070), P5DO6-1, _2 (EMQ.070), P5D07_1, _2 (EMQ.070), P5CHJB_1, _2 (EMQ.010)	1=35 hours or more per week, 2=Less than 35 hours per week, 3=Looking for work, 4=Not in the labor force
P5MOMOCC	Family/HH	Mother, female guardian, or mother figure's occupation	Combination of P5CHJB_1, _2, EMQ.010, EMQ.120, EMQ.130, and EMQ.140 (not on file)	01 Executive, Admin, Managerial Occupation 02 Engineers, Surveyors, & Architects 03 Natural Scientists & Mathematicians 04 Social Scientists/Workers, Lawyers 05 University Teachers, Postsecondary Counselors, Librarians 06 Teachers, except postsecondary 07 Physicians, Dentists, Veterinarians; 08 Registered Nurses, Pharmacists 09 Writers, Artists, Entertainers, Athletes 10 Health Technologists & Technicians 11 Technologists, except Health 12 Marketing & Sales Occupation 13 Administrative Support, including Clerk 14 Service Occupations 15 Agriculture, Forestry, Fishing Occupations 16 Mechanics & Repairs 17 Construction & Extractive Occupations 18 Precision Production Occupation 19 Production Working Occupation 20 Transportation, Material Moving 21 Handler, Equip, Cleaner, Helpers, Labor 22 Unemployed or Retired

See note at end of table.

Table 7-12. Spring-third grade composite variables—Continued

Variable name	Category	Description	Derived from	Values
P5ABSDAD	Family/ HH	Type of nonresident father	P5REL_1 through P5REL_25 (FSQ.130), P5CTP_N1, P5CTP_N2, P5CTP_N3, P5CTP_N4 (all from item NRQ.100)	1=Biological only, 2=Both biological and adoptive
P5ABSMOM	Family/HH	Type of nonresident mother	P5REL_1 through P5REL_25 (FSQ.130), P5CTP_N1, P5CTP_N2, P5CTP_N3, P5CTP_N4 (all from item NRQ.100)	1=Biological only, 2=Both biological and adoptive
P5FSRAW	Family/HH	Household food security raw score, a simple count of the number of food security items affirmed by the parent	P5WORRFD (FDQ.130A), P5FDLAST (FDQ.130B), P5BLMEAL (FDQ.130C), P5LOWCST (FDQ.130D), P5NOBAL (FDQ.130E), P5CANTAF (FDQ.130F), P5EVCUT2 (FDQ.140), P5EVCUT (FDQ.150), P5EATLES (FDQ.160), P5HUNGRY (FDQ.170), P5LOSEWT (FDQ.180), P5NOTEAT (FDQ.190), P5NOTEA2 (FDQ.200), P5CUTML (FDQ.210), P5CHSKIP (FDQ.220), P5OFTCUT (FDQ.230), P5CHIEVR (FDQ.240), P5NOMONY (FDQ.250)	Continuous
P5FSSCAL	Family/HH	Household food security scale score. This is a measure of the severity of food insecurity or hunger experienced in the household in the previous 12 months.	P5WORRFD (FDQ.130A), P5FDLAST (FDQ.130B), P5BLMEAL (FDQ.130C), P5LOWCST (FDQ.130D), P5NOBAL (FDQ.130E), P5CANTAF (FDQ.130F), P5EVCUT2 (FDQ.140), P5EVCUT (FDQ.150), P5EATLES (FDQ.160), P5HUNGRY (FDQ.170), P5LOSEWT (FDQ.180), P5NOTEAT (FDQ.190), P5NOTEA2 (FDQ.200), P5CUTML (FDQ.210), P5CHSKIP (FDQ.220), P5OFTCUT (FDQ.230), P5CHIEVR (FDQ.240), P5NOMONY (FDQ.250)	Continuous

See note at end of table.

Table 7-12. Spring-third grade composite variables—Continued

Variable name	Category	Description	Derived from	Values
P5FSSTAT	Family/HH	A categorical measure of household food security status that identifies households as food secure, food insecure without hunger, food insecure with hunger (moderate), and food insecure with hunger (severe)	P5WORRFD (FDQ.130A), P5FDLAST (FDQ.130B), P5BLMEAL (FDQ.130C), P5LOWCST (FDQ.130D), P5NOBAL (FDQ.130E), P5CANTAF (FDQ.130F), P5EVCUT2 (FDQ.140), P5EVCUT (FDQ.150), P5EATLES (FDQ.160), P5HUNGRY (FDQ.170), P5LOSEWT (FDQ.180), P5NOTEAT (FDQ.190), P5NOTEA2 (FDQ.200), P5CUTML (FDQ.210), P5CHSKIP (FDQ.220), P5OFTCUT (FDQ.230), P5CHIEVR (FDQ.240), P5NOMONY (FDQ.250)	1=Food secure; 2=Food insecure without hunger; 3=Food insecure with hunger (moderate); 4=Food insecure with hunger (severe)
P5FSCHRA	Family/HH	Children's food security raw score, a simple count of the number of child-referenced food security items affirmed by the parent	P5LOWCST (FDQ.130D), P5NOBAL (FDQ.130E), P5CANTAF (FDQ.130F), P5CUTML (FDQ.210), P5CHSKIP (FDQ.220), P5OFTCUT (FDQ.230), P5CHIEVR (FDQ.240), P5NOMONY (FDQ.250)	Continuous
P5FSCHSC	Family/HH	Children's food security scale score. This is a measure of the severity of food insecurity or hunger experienced by children in the household in the previous 12 months.	P5LOWCST (FDQ.130D), P5NOBAL (FDQ.130E), P5CANTAF (FDQ.130F), P5CUTML (FDQ.210), P5CHSKIP (FDQ.220), P5OFTCUT (FDQ.230), P5CHIEVR (FDQ.240), P5NOMONY (FDQ.250)	Continuous
P5FSCHST	Family/HH	A categorical measure of children's food security status that identifies households with hunger among children at some time during the 12 months prior to the survey	P5LOWCST (FDQ.130D), P5NOBAL (FDQ.130E), P5CANTAF (FDQ.130F), P5CUTML (FDQ.210), P5CHSKIP (FDQ.220), P5OFTCUT (FDQ.230), P5CHIEVR (FDQ.240), P5NOMONY (FDQ.250)	1=Food secure or food insecure without hunger among children; 2=Food insecure with hunger among children
P5RESID	Family/HH	Household roster number of respondent	P5PER_1 to P5PER_25 (parent interview household roster person type)	1-25

See note at end of table.

Table 7-12. Spring-third grade composite variables—Continued

Variable name	Category	Description	Derived from	Values
P5RESREL	Family/HH	Respondent relationship to focal child	P5REL_1 through P5REL_25(FSQ.130), P5UNR_1 through P5UNR_25 (FSQ.180), P5MOM_1 through P5MOM_25 (FSQ.140), P5DAD_1 through P5DAD_25 (FSQ.150)	1=Biological mother 2=Other mother type 3=Biological father 4=Other father type 5=Non-parent relative 6=Non-relative
P5CHLDID	Family/HH	Household roster number of respondent	P5PER_1 to P5PER_25 (parent interview household roster person type)	1-25
P5ERRFLG	Family/HH flag	Household roster has clear errors	P5REL_1 to P5REL_25 (FSQ.130), P5UNR_1 to P5UNR_25 (FSQ.180), P5JOI_1 to P5JOI_25 (round joined study), P5RDP_1 to P5RDP_25 (round departed study), P5REAS_1 to P5REAS_25 (reason left household)	0=False, 1=True
P5EDIT	Family/HH flag	Parent household matrix was edited	HOLDINGS (parent interview editing flag – not on file)	0=False, 1=True
P5SHCHG	Family/HH flag	Household roster had a change between rounds	P5JOI_1 to P5JOI_25 (round joined study), P5RDP_1 to P5RDP_25 (round departed study), P5REAS_1 to P5REAS_25 (reason left household)	0=False, 1=True
P5PARDAT	Family/HH flag	Presence of parent data	Presence or absence of parent interview	0= False, 1= True
T5GLVL	Teacher	Grade level of child	T5GRADE (TQC), E5GRADE (from SpEd-B), C_GRADE (from FMS), C5THIRD (ACQ.005), C5GRADE (ACQ.010), C5INGRAD (AIQ.030)	1=Kindergarten 2=First grade 3=Second grade 4=Third grade 5=Fourth grade 6=Fifth grade 7=Ungraded classroom

See note at end of table.

Table 7-12. Spring-third grade composite variables—Continued

Variable name	Category	Description	Derived from	Values
A5CLSZ	Teacher	Number of students in class	A5TOTRA(TQA Q4), A5TOTAG(TQA Q3), A5BOYS, A5GIRLS (TQA Q5)	Recoded from a continuous variable to a minimum value of 10 and a maximum value of 31
A5PLEP	Teacher	Percentage of limited English proficient children in the class	A5TOTLA (TQA Q14), A5LEP (TQA Q16), A5NUMLE (TQA Q17), A5CLSZ (composite)	Recoded from values of 0-100 to the following: 1=Less than 1% 2=1% to less than 5% 3=5% to less than 10% 4=10% to less than 25% 5=25% or more
A5PBLK	Teacher	Percent of Blacks in class—child-level data	A5BLACK (TQA item Q4c), A5CLSZ (composite)	Recoded from values of 0-100 to the following: 1=Less than 1% 2=1% to less than 5% 3=5% to less than 10% 4=10% to less than 25% 5=25% or more
A5PHIS	Teacher	Percent of Hispanics in class—child-level data	A5HISP (TQA item Q4b), A5CLSZ (composite)	Recoded from values of 0-100 to the following: 1=Less than 1% 2=1% to less than 5% 3=5% to less than 10% 4=10% to less than 25% 5=25% or more
A5PMIN	Teacher	Percent of minorities in class—child-level data	A5ASIAN, A5HISP, A5BLACK, A5AMRIN, A5RACEO (TQA items Q4a, b, c, e, f), A5CLSZ (composite)	Recoded from values of 0-100 to the following: 1=Less than 10% 2=10% to less than 25% 3=25% to less than 50% 4=50% to less than 75% 5=75% or more

See note at end of table.

Table 7-12. Spring-third grade composite variables—Continued

Variable name	Category	Description	Derived from	Values
A5TQUEX	Teacher flag	Presence of spring-third grade TQA data	Receipted TQAs in the FTS	0=False, 1=True
B5TQUEX	Teacher flag	Presence of spring-third grade TQB data	Receipted TQBs in the FTS	0=False, 1=True
T5TQCDAT	Teacher flag	Presence of spring-third grade TQC data	Receipted TQCs in the FTS	0=False, 1=True
D5SETQA	Teacher flag	Presence or Absence of Special Ed A data	Receipted special education instrument A's in the FTS	Changed from values of 0=False, 1=True to a suppressed variable
E5SETQB	Teacher flag	Presence or Absence of Special Ed B data	Receipted special education instrument B's in the FTS	Changed from values of 0=False, 1=True to a suppressed variable
R5REGION	School	Indicates the geographic region of the child's school	CREGION, R3REGION, R4REGION (composites), CCP and PSS files	1=Northeast: CT, ME, MA, NH, RI, VT, NJ, NY, PA; 2=Midwest: IL, IN, MI, OH, WI, IA, KS, MN, MO, NE, ND, SD; 3=South: DE, DC, FL, GA, MD, NC, SC, VA, WV, AL, KY, MS, TN, AR, LA, OK, TX; 4=West: AZ, CO, ID, MT, NV, NM, UT, WY, AK, CA, HA, OR, WA

See note at end of table.

Table 7-12. Spring-third grade composite variables—Continued

Variable name	Category	Description	Derived from	Values
R5URBAN	School	Location type for school—7 category version	KURBAN, R3URBAN, R4URBAN (composites), CCD and PSS files	1=Large city - a central city of Consolidated Metropolitan Statistical Area (CMSA) with a pop. greater than or equal to 250,000; 2=Mid-size city - a central city of a CMSA or Metropolitan Statistical Area (MSA) with a pop. Less than 250,000; 3= Large suburb; urban fringe of large city - any incorporated place, Census Designated Place, or nonplace territory within a CMSA or MSA of a large city and defined as urban by the U.S. Census Bureau; 4 = Mid-size suburb; urban fringe of mid-size city - any incorporated place, Census Designated Place, or nonplace territory within a CMSA or MSA of a mid-size city and defined as urban by the U.S. Census Bureau; 5= Large town - an incorporated place or Census Designated Place with a pop. greater than or equal to 25,000 and located outside a CMSA or MSA; 6=Small town - an incorporated place or Census Designated Place with a pop. less than 25,000 and greater than 2,500 - located outside a CMSA or MSA; 7=Rural - any incorporated place, Census Designated Place, or nonplace territory designated as rural by the U.S. Census Bureau

See note at end of table.

Table 7-12. Spring-third grade composite variables—Continued

Variable name	Category	Description	Derived from	Values
R5LOCALE	School	Location type for school—8 category version	R3LOCALE, R4LOCALE (composites), PSS and CCD files	1=Large city - a central city of Consolidated Metropolitan Statistical Area (CMSA) with a pop. greater than or equal to 250,000; 2=Mid-size city - a central city of a CMSA or Metropolitan Statistical Area (MSA) with a pop. Less than 250,000; 3= Large suburb; urban fringe of large city - any incorporated place, Census Designated Place, or nonplace territory within a CMSA or MSA of a large city and defined as urban by the U.S. Census Bureau; 4 = Mid-size suburb; urban fringe of mid-size city - any incorporated place, Census Designated Place, or nonplace territory within a CMSA or MSA of a mid-size city and defined as urban by the U.S. Census Bureau; 5= Large town - an incorporated place or Census Designated Place with a pop. greater than or equal to 25,000 and located outside a CMSA or MSA; 6=Small town - an incorporated place or Census Designated Place with a pop. less than 25,000 and greater than 2,500 - located outside a CMSA or MSA; 7 = non-MSA Rural - any incorporated place, Census Designated Place, or nonplace territory designated as rural by the U.S. Census Bureau that is not within a MSA; 8 = MSA Rural – any incorporated place, Census Designated Place, or nonplace territory designated as rural by the U.S. Census Bureau that is within a MSA

See notes at end of table.

Table 7-12. Spring-third grade composite variables—Continued

Variable name	Category	Description	Derived from	Values
S5SCTYP	School	School type from the school fact sheet	L5PUBLIC (SFS Q2), L5CATHOL, L5OTHREL, L5NAISKL, L5OTHPRI, L5PVTSPD, L5PVTEAR (all SFS Q4), CS_TYPE2, S4SCTYP, S3SCTYP, S2KSCTYP (composites), SCHL_TYP (School Master file variable derived from PSS/CCD, not on file)	1=Catholic, 2=Other Religious, 3=Other Private, 4=Public
S5PUPRI	School	Public or private school	S5SCTYP (composite)	1=Public, 2=Private
S5ENRLT	School	Total school third grade enrollment	PSS and CCD data	Recoded from a continuous variable to the following: 1=0 - 20 2=21 - 40 3=41 - 60 4=61 - 80 5=81 - 100 6=101 - 120 7=121 - 140 8= 141 - 160 9=161 - 180 10=More than 180
S5ENRLS	School	Total school enrollment	S5ANUMCH (SAQ Q1), PSS and CCD data	1=0-149 students; 2=150-299 students; 3=300-499 students; 4=500-749 students; 5= 750 and above students
S5MINOR	School	Percentage of minority students in school	PMINOR (School Master File variable derived from PSS/CCD, not on file), S5ASNPCT, S5HISPPCT, S5BLKPCT, S5INDPCT, S5OTHPCT (all from SAQ Q3)	Recoded from values of 0-100 to the following: 1=Less than 10% 2=10% to less than 25% 3=25% to less than 50% 4=50% to less than 75% 5=75% or more

See note at end of table.

Table 7-12. Spring-third grade composite variables—Continued

Variable name	Category	Description	Derived from	Values
S5FLNCH	School	Percentage of students eligible for free lunch in school	S5ELILNC (SAQ Q13), S5ANUMCH(SAQ Q1), CCD data	Recoded from values of 0 – 100 to values of 0 to 95
S5RLNCH	School	Percent of students eligible for reduced price lunch in school	S5ELIRED (SAQ Q13), S5ANUMCH (SAQ Q1), CCD data	Recoded from values of 0-100 to the following: 1=Less than 1% 2=1% to less than 5% 3=5% to less than 10% 4=10% to less than 25% 5=25% or more
S5SCLVL	School	School instructional level from school fact sheet	L5PRKNDR, L5KINDER, L5GRADE1, L5SECOND, L5THIRD, L5FOURTH, L5FIFTH, L5SIXTH, L57TH, L58TH, L5NINTH, L5TENTH, L511TH, L512TH (all from SFS Q1); S4SCLVL, S2SCLVL, GRSPAN (School Master file variable derived from PSS/CCD, not on file)	1=Less than first grade; 2=Primary school, 3=Elementary school, 4=Combined school
S5SCINC	School	Poverty level of school-50% or more students eligible for free or reduced-price lunch	S5PUPRI, S5FLNCH, S5RLNCH (composites), S5TT1SW (SAQ Q15), CCD data for free/reduced price lunch and school-wide Title 1 Status	1=50% or more low-income children 2=Less than 50% low-income children
L5SCHBDD	School	L5 School Year Starting Date, Day	L5SYRSDD (SFS Q5), FMS (variable not on file)	Changed from values of 1-31 to a suppressed variable
L5SCHBMM	School	L5 School Year Starting Date, Month	L5SYRSMM (SFS Q5), FMS (variable not on file)	Changed from values of 1-12 to a suppressed variable
L5SCHBYY	School	L5 School Year Starting Date, Year	L5SYRSYY (SFS Q5), FMS (variable not on file)	Changed from values of 2001-02 to a suppressed variable

See note at end of table.

Table 7-12. Spring-third grade composite variables—Continued

Variable name	Category	Description	Derived from	Values
L5SCHEDD	School	L5 School Year Ending Date, Day	L5SYREDD (SFS Q5), FMS (variable not on file)	Changed from values of 1-31 to a suppressed variable
L5SCHEMM	School	L5 School Year Ending Date, Month	L5SYREMM (SFS Q5), FMS (variable not on file)	Changed from values of 1-12 to a suppressed variable
L5SCHEYY	School	L5 School Year Ending Date, Year	L5SYREYY (SFS Q5), FMS (variable not on file)	Changed from values of 2001-02 to a suppressed variable
F5YRRND	School	Year round school	S_YRRNDFLG (FMS variable not on file)	1=Year round school, 2=Not year round school
K5INFAC	School flag	Presence or absence of facilities checklist data	Receipted facilities checklists in the FTS	0=False, 1=True
S5INSAQ	School flag	Presence or absence of SAQ data	Receipted SAQs in the FTS	0=False, 1=True
L5INSFS	School flag	Presence of School Fact Sheet data	Receipted school facts sheets in the FTS	0=False, 1=True
U5SRABS	School flag	Presence of spring-third grade SRA data	Receipted student record abstracts in the FTS	Changed from values of 0=False, 1=True to a suppressed variable

SOURCE: U.S. Department of Education, National Center for Education Statistics, Early Childhood Longitudinal Study, Kindergarten Class of 1998–99 third grade data collection, school year 2001–02.

- Certain continuous variables are modified into categorical variables, and certain categorical variables have their categories collapsed in the public-use file. While this protects the cases from a disclosure risk, these variables can still be used in all different kinds of analysis such as regression analysis.

There is a comment field in the variable frequency distribution view screen of the electronic code book that displays a comment for each masked variable indicating whether the variable from the restricted-use file has been recoded or suppressed in the public-use file. Variables that were recoded in any way during the data masking process display the comment, “These data recoded for respondent confidentiality.” Variables that were suppressed on the public-use file for protection of the respondent or child from identification display the comment, “These data suppressed for respondent confidentiality” and all values for the variable are set to equal –2 for that variable.

Table 7-13 presents the list of the masked variables. The table displays the variable name, variable label, and the comment displayed in the electronic code book indicating if the variable was recoded or suppressed. The table is sorted sequentially by the variable Field ID (see section 8.3.1.1 for how to use the variable Field ID.)

All variables from the special education teacher questionnaire part A (i.e., all variables with the prefix D5), from the special education teacher questionnaire part B (i.e., all variables with the prefix E5), and from the student record abstract (i.e., all variables with the prefix U5) have been suppressed in the third grade public-use file. Included in this group of suppressed variables are all teacher and school identifiers, which have last two characters “ID” and prefix D5, E5, or U5. For brevity, these variables are not included in table 7-13.

Table 7-13. Recoded and suppressed data on the ECLS-K Third-Grade Public-Use Data Fill: School year 2001–02

Field ID	Variable	Field Label	Comment
26	CS_TYPE2	TYPE OF SCHOOL IN BASE YEAR SAMPLE FRAME	These data recoded for respondent confidentiality.
28	R5URBAN	R5 LOCATION TYPE - 7 CATEGORIES	These data recoded for respondent confidentiality.
29	R5LOCALE	R5 LOCATION TYPE - 8 CATEGORIES	These data suppressed for respondent confidentiality.
31	R5FIPSST	R5 SCHOOL FIPS STATE CODE	These data suppressed for respondent confidentiality.
32	R5FIPSCT	R5 SCHOOL FIPS COUNTY CODE	These data suppressed for respondent confidentiality.
33	R5CCDLEA	R5 CCD LEA\SCHOOL DIST ID (PUBLIC)	These data suppressed for respondent confidentiality.
34	R5CCDSID	R5 CCD SCHOOL ID (PUBLIC)	These data suppressed for respondent confidentiality.
35	R5STSID	R5 STATE SCHOOL ID (PUBLIC)	These data suppressed for respondent confidentiality.
36	R5SCHZIP	R5 SCHOOL ZIP CODE	These data suppressed for respondent confidentiality.
37	R5SCHPIN	R5 SCHOOL PIN (PRIVATE)	These data suppressed for respondent confidentiality.
40	R5DOBYY	R5 CHILD COMPOSITE DOB YEAR	These data recoded for respondent confidentiality.
43	R5AGE	R5 COMPOSITE CHILD ASSESSMENT AGE(MNTHS)	These data recoded for respondent confidentiality.
278	A5CLSZ	A5 NUMBER OF STUDENTS IN CLASS	These data recoded for respondent confidentiality.
279	A5PBLK	A5 PERCENT OF BLACKS IN CLASS	These data recoded for respondent confidentiality.
280	A5PHIS	A5 PERCENT OF HISPANICS IN CLASS	These data recoded for respondent confidentiality.
281	A5PMIN	A5 PERCENT OF MINORITIES IN CLASS	These data recoded for respondent confidentiality.
282	A5PLEP	A5 PERCENT OF LEP STUDENTS IN CLASS	These data recoded for respondent confidentiality.
284	L5SCHBMM	L5 SCHOOL YEAR BEGINNING DATE MONTH	These data suppressed for respondent confidentiality.
285	L5SCHBDD	L5 SCHOOL YEAR BEGINNING DATE DAY	These data suppressed for respondent confidentiality.
286	L5SCHBYY	L5 SCHOOL YEAR BEGINNING DATE YEAR	These data suppressed for respondent confidentiality.
287	L5SCHEMM	L5 SCHOOL YEAR ENDING DATE MONTH	These data suppressed for respondent confidentiality.
288	L5SCHEDD	L5 SCHOOL YEAR ENDING DATE DAY	These data suppressed for respondent confidentiality.
291	S5ENRLT	S5 TOTAL SCHOOL THIRD GRADE ENROLLMENT	These data recoded for respondent confidentiality.
292	S5ENRLS	S5 TOTAL SCHOOL ENROLLMENT	These data recoded for respondent confidentiality.
293	S5MINOR	S5 PERCENT MINORITY STUDENTS	These data recoded for respondent confidentiality.
294	S5FLNCH	S5 PCT FREE LUNCH ELIGIBLE STUDENTS	These data recoded for respondent confidentiality.
295	S5RLNCH	S5 PCT REDUCED LUNCH ELIGIBLE STUDENTS	These data recoded for respondent confidentiality.
301	C5INGRAD	C5 AIQ030 GRADE CHILD REPORTED	These data recoded for respondent confidentiality.
302	C5THIRD	C5 ACQ005 CHILD IN THIRD GRADE	These data suppressed for respondent confidentiality.
313	C5SPECAC	C5 ACQ270 SPECIAL ACCOMMODATION LISTED	These data suppressed for respondent confidentiality.
329	C5ATTMPT	C5 CHILD ASSESSMENT NUMBER OF ATTEMPTS	These data recoded for respondent confidentiality.
330	P5HOMZIP	P5 HOME ZIP CODE	These data suppressed for respondent confidentiality.
741	P5DIVSEP	P5 FSQ310A DIVORCE/SEPARATION IN FAMILY	These data suppressed for respondent confidentiality.
742	P5MARREM	P5 FSQ310B MARRIAGE/REMARRIAGE IN FAMILY	These data suppressed for respondent confidentiality.
743	P5PARDIE	P5 FSQ310C PARENT IN FAMILY DIED	These data suppressed for respondent confidentiality.
745	P5HOMELS	P5 FSQ310E FAMILY HOMELESS FOR A TIME	These data suppressed for respondent confidentiality.
1089	P5DGNATT	P5 CHQ060 1ST DIAGNOSIS-LEARNING ABILITY	These data suppressed for respondent confidentiality.
1091	P5YYDIAG	P5 CHQ075 YR AT 1ST DIAGNOSIS-LRN ABLTY	These data suppressed for respondent confidentiality.
1095	P5PROFFD	P5 CHQ110 IF ACTIVITY PROBLEM DIAGNOSED	These data suppressed for respondent confidentiality.
1096	P5DGNACT	P5 CHQ120 WHAT 1ST DIAGNOSIS - ACTIVITY	These data suppressed for respondent confidentiality.
1098	P5YYDIA2	P5 CHQ135 YR AT 1ST DIAGNOSIS-ACTIVITY	These data suppressed for respondent confidentiality.
1103	P5YYDIA4	P5 CHQ185 YEAR AT 1ST DIAGNOSIS-SPEECH	These data suppressed for respondent confidentiality.
1108	P5YYDIA5	P5 CHQ225 YR AT 1ST DIAGNOSIS-HEARING	These data suppressed for respondent confidentiality.
1109	P5HEARS	P5 CHQ230 DEGREE OF CHILD’S DEAFNESS	These data suppressed for respondent confidentiality.
1110	P5HEARAI	P5 CHQ240 IF CHILD WEARS HEARING AID	These data suppressed for respondent confidentiality.

See note at end of table.

Table 7-13. Recoded and suppressed data on the ECLS-K Third-Grade Public-Use Data File—Continued

Field ID	Variable	Field Label	Comment
1111	P5COCHLE	P5 CHQ250 IF CHILD HAS COCHLEAR IMPLANTS	These data suppressed for respondent confidentiality.
1112	P5HEARS2	P5 CHQ260 DEVICE EFFECT ON CHD'S HEARING	These data suppressed for respondent confidentiality.
1117	P5DIA6YY	P5 CHQ313 YR AT 1ST DIAGNOSIS-VISION	These data suppressed for respondent confidentiality.
1118	P5CORREC	P5 CHQ315 IF CHD'S VISION IS CORRECTABLE	These data suppressed for respondent confidentiality.
1119	P5BESTEY	P5 CHQ320 WHAT CAN CHILD BEST SEE	These data suppressed for respondent confidentiality.
1123	P5DIABEH	P5 CHQ335 BEHAVIOR PROBLEM DIAGNOSED	These data suppressed for respondent confidentiality.
1124	P5DGNBEH	P5 CHQ337 1ST DIAGNOSIS-BEHAVIOR	These data suppressed for respondent confidentiality.
1126	P5DGBEYY	P5 CHQ345 YR AT 1ST DIAGNOSIS-BEHAVIOR	These data suppressed for respondent confidentiality.
1134	P5SPECIL	P5 CHQ510 IF CHD USES SPECIAL EQUIPMENT	These data suppressed for respondent confidentiality.
1140	P5SPECND	P5 CHQ545 CHILD SPECIAL NEEDS/EDUCATION	These data suppressed for respondent confidentiality.
1150	P5HOSINJ	P5 CHQ692 HOSPITALIZED FOR INJURY	These data suppressed for respondent confidentiality.
1280	P5RECFRE	P5 WPQ215 DOES CHILD REC FREE REDUCED BF	These data suppressed for respondent confidentiality.
1281	P5FRERED	P5 WPQ216 FREE OR REDUCED BREAKFAST	These data suppressed for respondent confidentiality.
1287	P5HOWPAY	P5 PAQ137 HOW MUCH PAID IN TUITION ($)	These data recoded for respondent confidentiality.
1306	A5UNGRAD	A5 Q1A TCH UNGRADED CLASS	These data suppressed for respondent confidentiality.
1307	A5T1ST	A5 Q1B TCH 1ST GRADE LEVEL	These data suppressed for respondent confidentiality.
1308	A5T2ND	A5 Q1C TCH 2ND GRADE LEVEL	These data suppressed for respondent confidentiality.
1309	A5T3RD	A5 Q1D INCLUDES 3RD GRADE	These data suppressed for respondent confidentiality.
1310	A5T4TH	A5 Q1E INCLUDES 4TH GRADE	These data suppressed for respondent confidentiality.
1311	A5T5TH	A5 Q1F INCLUDES 5TH GRADE	These data suppressed for respondent confidentiality.
1312	A5CLSORG	A5 Q2 CLASS ORGANIZED	These data recoded for respondent confidentiality.
1313	A57YRSLS	A5 Q3A # 7 YEARS OLD OR LESS IN CLASS	These data recoded for respondent confidentiality.
1314	A58YROL	A5 Q3B HOW MANY 8-YEAR-OLDS IN CLASS	These data recoded for respondent confidentiality.
1315	A59YROL	A5 Q3C HOW MANY 9-YEAR-OLDS IN CLASS	These data recoded for respondent confidentiality.
1316	A510YROL	A5 Q3D HOW MANY 10-YEAR-OLDS IN CLASS	These data recoded for respondent confidentiality.
1317	A511YRMO	A5 Q3E # 11-YEAR OLDS & OLDER IN CLASS	These data recoded for respondent confidentiality.
1318	A5TOTAG	A5 Q3F TOTAL CLASS ENROLLMENT (AGE)	These data recoded for respondent confidentiality.
1319	A5ASIAN	A5 Q4A # ASIAN/PACIFIC ISLANDERS	These data recoded for respondent confidentiality.
1320	A5HISP	A5 Q4B # HISPANICS (ALL RACES)	These data recoded for respondent confidentiality.
1321	A5BLACK	A5 Q4C # NON-HISPANIC BLACKS	These data recoded for respondent confidentiality.
1322	A5WHITE	A5 Q4D # NON-HISPANIC WHITES	These data recoded for respondent confidentiality.
1323	A5AMRIN	A5 Q4E # AMERICAN INDIANS	These data recoded for respondent confidentiality.
1324	A5RACEO	A5 Q4F # OF STUDENTS OF OTHER RACES	These data recoded for respondent confidentiality.
1325	A5TOTRA	A5 Q4G TOTAL CLASS ENROLLMENT (RACES)	These data recoded for respondent confidentiality.
1326	A5BOYS	A5 Q5A NUMBER OF BOYS IN CLASS	These data recoded for respondent confidentiality.
1327	A5GIRLS	A5 Q5B NUMBER OF GIRLS IN CLASS	These data recoded for respondent confidentiality.
1328	A5ELFSBR	A5 Q6A ELIGIBLE FREE SCHOOL BREAKFAST	These data recoded for respondent confidentiality.
1329	A5ELRPBR	A5 Q6B ELIGIBLE REDUCED-PRICE BREAKFAST	These data recoded for respondent confidentiality.
1330	A5ELFSLU	A5 Q6C ELIGIBLE FREE SCHOOL LUNCH	These data recoded for respondent confidentiality.
1331	A5ELRPLU	A5 Q6D ELIGIBLE REDUCED-PRICE LUNCH	These data recoded for respondent confidentiality.
1337	A5GIFT	A5 Q9A # CLASSIFIED AS GIFTED/TALENTED	These data recoded for respondent confidentiality.
1338	A5PRTGF	A5 Q9B # TAKE PART IN GIFTED/TALENTED	These data recoded for respondent confidentiality.
1339	A5REPEAT	A5 Q9C # REPEATING THIS GRADE	These data recoded for respondent confidentiality.
1347	A5DISAB	A5 Q11 NUMBER WITH DISABILITIES	These data recoded for respondent confidentiality.
1348	A5IMPAI	A5 Q12A COMMUNICATION IMPAIRMENTS	These data recoded for respondent confidentiality.
1349	A5LRNDI	A5 Q12B LEARNING DISABILITIES	These data recoded for respondent confidentiality.
1350	A5EMPRB	A5 Q12C SERIOUS EMOTIONAL PROBLEMS	These data recoded for respondent confidentiality.
1351	A5RETAR	A5 Q12D MENTAL RETARDATION	These data suppressed for respondent confidentiality.

See note at end of table.

Table 7-13. Recoded and suppressed data on the ECLS-K Third-Grade Public-Use Data File—Continued

Field ID	Variable	Field Label	Comment
1352	A5DELAY	A5 Q12E DEVELOPMENTAL DELAY	These data recoded for respondent confidentiality.
1353	A5VIS	A5 Q12F VISION IMPAIRMENTS	These data recoded for respondent confidentiality.
1354	A5HEAR	A5 Q12G HEARING IMPAIRMENTS	These data recoded for respondent confidentiality.
1355	A5ORTHO	A5 Q12H ORTHOPEDIC IMPAIRMENTS	These data recoded for respondent confidentiality.
1356	A5OTHER	A5 Q12I OTHER HEALTH IMPAIRMENTS	These data recoded for respondent confidentiality.
1357	A5MULTI	A5 Q12J MULTIPLE DISABILTIES	These data suppressed for respondent confidentiality.
1358	A5AUTSM	A5 Q12K AUTISM	These data suppressed for respondent confidentiality.
1359	A5TRAUM	A5 Q12L TRAUMATIC BRAIN INJURIES	These data recoded for respondent confidentiality.
1360	A5DEAF	A5 Q12M DEAFNESS AND BLINDNESS	These data suppressed for respondent confidentiality.
1361	A5OTDIS	A5 Q12N OTHER SPECIFY DISABILITIES	These data suppressed for respondent confidentiality.
1362	A5NCDIS	A5 Q12O NOT CLASSIFIED DISABILITIES	These data suppressed for respondent confidentiality.
1363	A5SPCIA	A5 Q13A SPECIAL DISABILITY SERVICES	These data recoded for respondent confidentiality.
1364	A5IEP	A5 Q13B IEP FOR CHILDREN W/ DISABILITY	These data recoded for respondent confidentiality.
1365	A5SC504	A5 Q13C CHILDREN W/ SECTION 504 PLAN	These data recoded for respondent confidentiality.
1369	A5CASN	A5 Q15B STUDENTS SPEAK AN ASIAN LANGUAGE	These data suppressed for respondent confidentiality.
1370	A5OTLNG	A5 Q15C STUDENTS SPEAK ANOTHER LNG	These data suppressed for respondent confidentiality.
1371	A5LANOS	A5 Q15C SPECIFY STUDENTS’ OTHER LANG	These data suppressed for respondent confidentiality.
1373	A5NUMLE	A5 Q17 NUMBER LEP STUDENTS IN CLASS	These data recoded for respondent confidentiality.
1374	A5NOESL	A5 Q18A LEP STUDENTS GET NO ESL	These data recoded for respondent confidentiality.
1375	A5ESLRE	A5 Q18B LEP STUDENTS IN-CLASS ESL	These data recoded for respondent confidentiality.
1376	A5ESLOU	A5 Q18C LEP STUDENTS OUTSIDE ESL	These data recoded for respondent confidentiality.
1377	A5TENG	A5 Q19A TCHRS SPEAK ENGLISH	These data suppressed for respondent confidentiality.
1379	A5TASN	A5 Q19C TCHRS SPEAK AN ASIAN LANGUAGE	These data suppressed for respondent confidentiality.
1380	A5TOTLG	A5 Q19D TCHRS SPEAK ANOTHER LANGUAGE	These data suppressed for respondent confidentiality.
1381	A5LEPOS	A5 Q19D SPECIFY TCHRS OTHER LANGUAGE	These data suppressed for respondent confidentiality.
1382	A5IENG	A5 Q20A INSTRUCTION- ENGLISH	These data suppressed for respondent confidentiality.
1383	A5ISPNH	A5 Q20B INSTRUCTION - SPANISH	These data suppressed for respondent confidentiality.
1384	A5IASN	A5 Q20C INSTRUCTION - ASIAN LNG	These data suppressed for respondent confidentiality.
1385	A5IOTLN	A5 Q20D INSTRUCTION - OTHER LNG	These data suppressed for respondent confidentiality.
1386	A5LNGOS	A5 Q20D LANGUAGE OF INSTRUCTION-OTHER	These data suppressed for respondent confidentiality.
1396	A5COMPUT	A5 Q23A # OF COMPUTERS IN CLASS	These data recoded for respondent confidentiality.
1397	A5INET	A5 Q23B # OF COMPUTERS WITH INTERNET	These data recoded for respondent confidentiality.
1398	A5COMUSE	A5 Q23C # OF COMPUTERS CHILDREN USE	These data recoded for respondent confidentiality.
1670	B5TGEND	B5 Q10 TEACHER’S GENDER	These data suppressed for respondent confidentiality.
1671	B5YRBORN	B5 Q11 TEACHER’S YEAR OF BIRTH	These data recoded for respondent confidentiality.
1672	B5HISP	B5 Q12 HISPANIC OR LATINO	These data suppressed for respondent confidentiality.
1673	B5RACE1	B5 Q13A AMERICAN INDIAN OR ALASKA NATIVE	These data suppressed for respondent confidentiality.
1674	B5RACE2	B5 Q13B ASIAN	These data suppressed for respondent confidentiality.
1675	B5RACE3	B5 Q13C BLACK OR AFRICAN AMERICAN	These data suppressed for respondent confidentiality.
1676	B5RACE4	B5 Q13D NATIVE HAWAIIAN OR OTHER PAC IS	These data suppressed for respondent confidentiality.
1678	B5YRSTC	B5 Q14 NUMBER YEARS BEEN SCHOOL TEACHER	These data recoded for respondent confidentiality.
1679	B5YRSGRA	B5 Q15 YEARS TAUGHT THIS GRADE	These data recoded for respondent confidentiality.
1680	B5YRSCH	B5 Q16 YRS TCHR TAUGHT AT THIS SCHOOL	These data recoded for respondent confidentiality.
1681	B5ASSIGN	B5 Q17 MAIN ASSIGNMENT AT SCHOOL	These data recoded for respondent confidentiality.
1682	B5HGHSTD	B5 Q18 HIGHEST ED LEVEL TEACHER ACHIEVED	These data recoded for respondent confidentiality.
1714	T5RLATCH	T5 Q1 READING/LANGUAGE ARTS TEACHER	These data suppressed for respondent confidentiality.
1725	T5MATCH	T5 Q1 MATHEMATICS TEACHER	These data suppressed for respondent confidentiality.
1745	T5SSTCH	T5 Q1 SOCIAL STUDIES TEACHER	These data suppressed for respondent confidentiality.
1753	T5GRADE	T5 Q1 GRADE CHILD ENROLLED	These data suppressed for respondent confidentiality.

See note at end of table.

Table 7-13. Recoded and suppressed data on the ECLS-K Third-Grade Public-Use Data File—Continued

Field ID	Variable	Field Label	Comment
1759	T5PLLESL	T5 Q3E PULL-OUT ESL PROGRAM	These data suppressed for respondent confidentiality.
1760	T5INCESL	T5 Q3F IN-CLASS ESL	These data suppressed for respondent confidentiality.
1762	T5GFTRD	T5 Q3H GIFTED PROGRAM IN READING	These data suppressed for respondent confidentiality.
1763	T5GFTMTH	T5 Q3I GIFTED PROGRAM IN MATHEMATICS	These data suppressed for respondent confidentiality.
1766	T5MENTOR	T5 Q3L MEET W/MENTOR NOT PROF PSYCH	These data suppressed for respondent confidentiality.
1767	T5BEFORE	T5 Q4A INSTR SERVICES BEFORE SCHOOL	These data suppressed for respondent confidentiality.
1769	T5WKEND	T5 Q4C INSTR SERVICES WEEKENDS	These data suppressed for respondent confidentiality.
1773	T5TT1ENG	T5 Q5C TITLE 1 ENGLISH/LANGUAGE ARTS	These data suppressed for respondent confidentiality.
1774	T5TT1CMB	T5 Q5D TITLE 1 COMBINED SUBJECTS	These data suppressed for respondent confidentiality.
1775	T5TT1ES	T5 Q5E TITLE 1 ESL/BILINGUAL	These data suppressed for respondent confidentiality.
1776	T5TT1SP	T5 Q5F TITLE 1 HANDICAPPED/SPECIAL ED	These data suppressed for respondent confidentiality.
1783	T5ACCOM	T5 Q10 SPECIAL TEST ACCOMMODATIONS	These data suppressed for respondent confidentiality.
1785	T5FLBHND	T5 Q12 FELL BEHIND DUE TO HEALTH	These data suppressed for respondent confidentiality.
1787	T5CHRDGP	T5 Q14 CHILDS PLACEMENT IN READING GRP	These data recoded for respondent confidentiality.
1797	T5LNGTM	T5 Q20 LENGTH OF TIME IN CLASSROOM	These data suppressed for respondent confidentiality.
1798	T5PROMOT	T5 Q21 RECOMMEND PROMOTION/YR END	These data suppressed for respondent confidentiality.
1799	T5SUMPRG	T5 Q22 IF RETAINED ELIG FOR SUMMER PROG	These data suppressed for respondent confidentiality.
1803	L5UNGRAD	L5 Q1A GRADE LEVEL-UNGRADED	These data suppressed for respondent confidentiality.
1806	L5KINDER	L5 Q1D GRADE LEVEL-KINDERGARTEN	These data suppressed for respondent confidentiality.
1807	L5GRADE1	L5 Q1E GRADE LEVEL-FIRST GRADE	These data suppressed for respondent confidentiality.
1808	L5SECOND	L5 Q1F GRADE LEVEL-SECOND GRADE	These data suppressed for respondent confidentiality.
1809	L5THIRD	L5 Q1G GRADE LEVEL-THIRD GRADE	These data suppressed for respondent confidentiality.
1810	L5FOURTH	L5 Q1H GRADE LEVEL-FOURTH GRADE	These data suppressed for respondent confidentiality.
1815	L5NINTH	L5 Q1M GRADE LEVEL-NINTH GRADE	These data suppressed for respondent confidentiality.
1816	L5TENTH	L5 Q1N GRADE LEVEL-TENTH GRADE	These data suppressed for respondent confidentiality.
1817	L511TH	L5 Q1O GRADE LEVEL-ELEVENTH GRADE	These data suppressed for respondent confidentiality.
1818	L512TH	L5 Q1P GRADE LEVEL-TWELFTH GRADE	These data suppressed for respondent confidentiality.
1820	L5REGSKL	L5 Q3A IS IT REGULAR PUBLIC SCHOOL	These data suppressed for respondent confidentiality.
1821	L5MAGSKL	L5 Q3B IS IT A MAGNET SCHOOL	These data suppressed for respondent confidentiality.
1823	L5BIASKL	L5 Q3D IS IT A TRIBAL SCHOOL	These data suppressed for respondent confidentiality.
1824	L5SPEDSK	L5 Q3E IS IT A SPECIAL ED SCHOOL	These data suppressed for respondent confidentiality.
1829	L5PRIVRD	L5 Q4A4 IS IT A PRIVATE ORDER	These data suppressed for respondent confidentiality.
1831	L5NAISKL	L5 Q4C PRIVATE-ACCREDITED BY NAIS	These data suppressed for respondent confidentiality.
1832	L5OTHPRI	L5 Q4D IS IT OTHER PRIVATE	These data suppressed for respondent confidentiality.
1833	L5PVTSPD	L5 Q4E IS IT SPECIAL EDUCATION	These data suppressed for respondent confidentiality.
1835	L5SYRSMM	L5 Q5A SCH START MONTH	These data suppressed for respondent confidentiality.
1836	L5SYRSDD	L5 Q5B SCH START DAY	These data suppressed for respondent confidentiality.
1837	L5SYRSYY	L5 Q5C SCH START YEAR	These data suppressed for respondent confidentiality.
1838	L5SYREMM	L5 Q5D SCH END MONTH	These data suppressed for respondent confidentiality.
1839	L5SYREDD	L5 Q5E SCH END DAY	These data suppressed for respondent confidentiality.
1841	L5DAYSYR	L5 Q6 DAYS IN SCH YR	These data recoded for respondent confidentiality.
1842	L5ATNREC	L5 Q7 SCH KEEPS ATTENDANCE RECORDS	These data suppressed for respondent confidentiality.
1845	L5MM02	L5 Q10A MONTH COMPLETED	These data suppressed for respondent confidentiality.
1848	S5ANUMCH	S5 Q1A # ENROLLED AROUND 10/1/2001	These data recoded for respondent confidentiality.
1849	S5BNUMCH	S5 Q1B # ENROLLED SINCE 10/1/2001	These data recoded for respondent confidentiality.
1850	S5CNUMCH	S5 Q1C # WHO LEFT - DIDN'T RETURN	These data recoded for respondent confidentiality.
1851	S5ADA	S5 Q2 % AVERAGE DAILY ATTENDANCE FOR YR	These data recoded for respondent confidentiality.
1853	S5ASNPCT	S5 Q3A PERCENT OF ASIAN STUDENTS	These data suppressed for respondent confidentiality.
1855	S5HSPPCT	S5 Q3B PERCENT OF HISPANIC STUDENTS	These data recoded for respondent confidentiality.

See note at end of table.

Table 7-13. Recoded and suppressed data on the ECLS-K Third-Grade Public-Use Data File—Continued

Field ID	Variable	Field Label	Comment
1857	S5BLKPCT	S5 Q3C PERCENT OF BLACK STUDENTS	These data recoded for respondent confidentiality.
1859	S5WHTPCT	S5 Q3D PERCENT OF WHITE STUDENTS	These data suppressed for respondent confidentiality.
1861	S5INDPCT	S5 Q3E PERCENT OF AMERICAN INDIANS	These data suppressed for respondent confidentiality.
1863	S5OTHPCT	S5 Q3F PERCENT OF OTHER STUDENTS	These data recoded for respondent confidentiality.
1865	S5AMBUSF	S5 Q4 TIME FIRST BUS AM	These data recoded for respondent confidentiality.
1866	S5AMBUSL	S5 Q5 TIME LAST BUS AM	These data recoded for respondent confidentiality.
1867	S5STRTAM	S5 Q6 OFFICIAL SCHOOL START TIME AM	These data recoded for respondent confidentiality.
1875	S5BRKSTR	S5 Q9A TIME BREAKFAST START	These data recoded for respondent confidentiality.
1876	S5BRKEND	S5 Q9B TIME BREAKFAST END	These data recoded for respondent confidentiality.
1879	S5PRABRK	S5 Q12A2 PARTICIPATE ANY SCH BREAKFAST	These data suppressed for respondent confidentiality.
1880	S5ELIBRK	S5 Q12B1 ELIGIBLE FOR FREE BREAKFAST	These data suppressed for respondent confidentiality.
1881	S5PARBRK	S5 Q12B2 PARTICIPATES IN BREAKFAST	These data suppressed for respondent confidentiality.
1882	S5ELRPBK	S5 Q12C1 ELIGIBLE RED-PRICE BREAKFAST	These data suppressed for respondent confidentiality.
1883	S5PARPBK	S5 Q12C2 PARTICIPATE RED-PRICE BREAKFAST	These data suppressed for respondent confidentiality.
1884	S5PAALUN	S5 Q13A2 PARTICIPATE ANY SCH LUNCH	These data suppressed for respondent confidentiality.
1885	S5ELILNC	S5 Q13B1 ELIGIBLE FOR FREE LUNCH	These data suppressed for respondent confidentiality.
1886	S5PARLNC	S5 Q13B2 PARTICIPATES IN FREE LUNCH	These data suppressed for respondent confidentiality.
1887	S5ELIRED	S5 Q13C1 ELIGIBLE IN REDUCED-PRICE LUNCH	These data suppressed for respondent confidentiality.
1888	S5PARRED	S5 Q13C2 PARTICIPATES IN RED-PRICE LUNCH	These data suppressed for respondent confidentiality.
1899	S5CHLDNM	S5 Q17 # OF CHILDREN SITE ACCOMMODATES	These data recoded for respondent confidentiality.
1900	S5PORTBL	S5 Q18 # PORTABLE CLASSROOMS	These data recoded for respondent confidentiality.
1903	S5LBRYOK	S5 Q19C DOES LIBRARY MEET NEEDS	These data suppressed for respondent confidentiality.
1907	S5PLAYOK	S5 Q19G DOES PLAYGROUND MEET NEEDS	These data suppressed for respondent confidentiality.
1911	S5INSTCM	S5 Q20A # INSTRUCTIONAL COMPUTERS	These data recoded for respondent confidentiality.
1912	S5ADMNCM	S5 Q20B # INSTRUCT/ADMIN COMPUTERS	These data recoded for respondent confidentiality.
1913	S5TOTCM	S5 Q20C TOTAL NUMBER OF COMPUTERS	These data recoded for respondent confidentiality.
1916	S5CDRSCH	S5 Q21B1 COMPUTERS W/ CD-ROM ACCESS	These data suppressed for respondent confidentiality.
1917	S5CDRSTU	S5 Q21B2 CD-ROM FOR STUDENT USE	These data suppressed for respondent confidentiality.
1918	S5NETSCH	S5 Q21C1 COMPUTERS W/INTERNET ACCESS	These data suppressed for respondent confidentiality.
1919	S5NETSTU	S5 Q21C2 INTERNET FOR STUDENT USE	These data suppressed for respondent confidentiality.
1941	S5NOTEST	S5 Q26 NO GRADE TESTED	These data suppressed for respondent confidentiality.
1957	S5RETAIN	S5 Q28 CAN CHILD BE RETAINED IN GRADE	These data suppressed for respondent confidentiality.
1961	S5ACADEM	S5 Q29D RETAINED- BELOW GRADE LEVEL	These data suppressed for respondent confidentiality.
1984	S5LEPSCH	S5 Q34A PERCENT OF LEP CHILDREN	These data recoded for respondent confidentiality.
1985	S5LEPTHR	S5 Q34B % LEP IN THIRD GRADE	These data recoded for respondent confidentiality.
1986	S5BILING	S5 Q35A BILINGUAL SERVICES PERCENT -3RD	These data suppressed for respondent confidentiality.
1987	S5ESLONL	S5 Q35B ESL SERVICES PERCENT -3RD	These data suppressed for respondent confidentiality.
1988	S5ESLBIL	S5 Q35C ESL AND BILINGUAL PERCENT -3RD	These data suppressed for respondent confidentiality.
2003	S5RTCHFL	S5 Q40A1 # REG CLASSROOM TCHR-FULL	These data recoded for respondent confidentiality.
2004	S5RTCHPT	S5 Q40A2 # REG CLASSROOM TCHR-PART	These data recoded for respondent confidentiality.
2005	S5MSARFL	S5 Q40B1 # GYM DRAMA MUSIC ART TCHR-FULL	These data recoded for respondent confidentiality.
2006	S5MSARPT	S5 Q40B2 # GYM DRAMA MUSIC ART TCHR-PART	These data recoded for respondent confidentiality.
2007	S5SPEDFL	S5 Q40C1 # SPECIAL ED TCHR-FULL	These data recoded for respondent confidentiality.
2008	S5SPEDPT	S5 Q40C2 # SPECIAL ED TCHR-PART	These data recoded for respondent confidentiality.
2009	S5ESLFL	S5 Q40D1 # ESL/BILINGUAL TCHR-FULL	These data recoded for respondent confidentiality.
2010	S5ESLPT	S5 Q40D2 # ESL/BILINGUAL TCHR-PART	These data recoded for respondent confidentiality.
2011	S5READFL	S5 Q40E1 # READING TCHR/SPECIALIST-FULL	These data recoded for respondent confidentiality.
2012	S5READPT	S5 Q40E2 # READING TCHR/SPECIALIST-PART	These data recoded for respondent confidentiality.
2013	S5GIFTFL	S5 Q40F1 # GIFTED/TALENTED TCHR-FULL	These data recoded for respondent confidentiality.

See note at end of table.

Table 7-13. Recoded and suppressed data on the ECLS-K Third-Grade Public-Use Data File—Continued

Field ID	Variable	Field Label	Comment
2014	S5GIFTPT	S5 Q40F2 # GIFTED/TALENTED TCHR-PART	These data recoded for respondent confidentiality.
2015	S5NURSFL	S5 Q40G1 # SCH NURSE HEALTH PROF-FULL	These data recoded for respondent confidentiality.
2016	S5NURSPT	S5 Q40G2 # SCH NURSE HEALTH PROF-PART	These data recoded for respondent confidentiality.
2017	S5PSYCFL	S5 Q40H1 SCH PSYCH/ SOCIAL WORKER-FULL	These data recoded for respondent confidentiality.
2018	S5PSYCPT	S5 Q40H2 SCH PSYCH/SOCIAL WORKER-PART	These data recoded for respondent confidentiality.
2019	S5PARAFL	S5 Q40I1 # PARAPROFESSIONALS-FULL	These data recoded for respondent confidentiality.
2020	S5PARAPT	S5 Q40I2 # PARAPROFESSIONALS-PART	These data recoded for respondent confidentiality.
2021	S5LIBRFL	S5 Q40J1 # LIBRARIANS-FULL	These data recoded for respondent confidentiality.
2022	S5LIBRPT	S5 Q40J2 # LIBRARIANS-PART	These data recoded for respondent confidentiality.
2023	S5TEAASN	S5 Q41A % ASIAN TEACHERS	These data suppressed for respondent confidentiality.
2025	S5ETHNIC	S5 Q41B % HISPANIC/LATINO TEACHERS	These data recoded for respondent confidentiality.
2027	S5TEABLK	S5 Q41C % BLACK TEACHERS	These data recoded for respondent confidentiality.
2029	S5TEAWHT	S5 Q41D % WHITE TEACHERS	These data recoded for respondent confidentiality.
2031	S5TEAIND	S5 Q41E % AMERICAN INDIAN TEACHERS	These data suppressed for respondent confidentiality.
2033	S5TEAOTH	S5 Q41F % OTHER TEACHERS	These data suppressed for respondent confidentiality.
2036	S5HWLONG	S5 Q42E RESP (NOT PRINCIPAL) YR AT SCH	These data recoded for respondent confidentiality.
2056	S5BRTHYR	S5 Q46 YEAR PRINCIPAL WAS BORN	These data recoded for respondent confidentiality.
2057	S5ORIGIN	S5 Q47 PRINCIPAL IS HISPANIC/LATINO	These data suppressed for respondent confidentiality.
2058	S5RACE1	S5 Q48A PRINCIPAL IS AMERICAN INDIAN	These data suppressed for respondent confidentiality.
2059	S5RACE2	S5 Q48B PRINCIPAL IS ASIAN	These data suppressed for respondent confidentiality.
2060	S5RACE3	S5 Q48C PRINCIPAL IS BLACK	These data suppressed for respondent confidentiality.
2061	S5RACE4	S5 Q48D PRINCIPAL IS HAWAIIAN OR PAC IS	These data suppressed for respondent confidentiality.
2062	S5RACE5	S5 Q48E PRINCIPAL IS WHITE	These data suppressed for respondent confidentiality.
2063	S5YSTCH	S5 Q49A NUMBER OF YRS TEACHING	These data recoded for respondent confidentiality.
2064	S5TOTPRI	S5 Q49B NUMBER OF YRS AS PRINCIPAL	These data recoded for respondent confidentiality.
2065	S5PRINHR	S5 Q49C NUMBER YRS A PRINCIPAL HERE	These data recoded for respondent confidentiality.
2066	S5EDLVL	S5 Q50 HIGHEST LEVEL OF EDUCATION	These data recoded for respondent confidentiality.
2067	S5MAJOR	S5 Q51 MAJOR FIELD HIGHEST ED LEVEL	These data recoded for respondent confidentiality.
2071	K5PRTCLS	K5 Q1 NUMBER OF PORTABLE CLASSROOMS	These data recoded for respondent confidentiality.
2072	K5GUARDO	K5 Q2A1 OBSERVED SECURITY GUARD	These data suppressed for respondent confidentiality.
2073	K5GUARDS	K5 Q2A2 SCH CNFRM SECURITY GUARD	These data suppressed for respondent confidentiality.
2074	K5METDTO	K5 Q2B1 OBSERVED METAL DETECTORS	These data suppressed for respondent confidentiality.
2075	K5METDTS	K5 Q2B2 SCH CNFRM METAL DETECTORS	These data suppressed for respondent confidentiality.
2077	K5SCAMS	K5 Q2C2 SCH CNFRM SECURITY CAMERAS	These data suppressed for respondent confidentiality.
2079	K5BARSS	K5 Q2D2 SCH CNFRM WINDOW AND DOOR BARS	These data suppressed for respondent confidentiality.
2083	K5FENCES	K5 Q2F2 SCH CNFRM FENCING AROUND SCHOOL	These data suppressed for respondent confidentiality.
2088	K5INTCMO	K5 Q2I1 OBSERVED INTERCOMS	These data suppressed for respondent confidentiality.
2090	K5ALARMO	K5 Q2J1 OBSERVED FIRE ALARMS	These data suppressed for respondent confidentiality.
2091	K5ALARMS	K5 Q2J2 SCH CNFRM FIRE ALARMS	These data suppressed for respondent confidentiality.
2092	K5FREXTO	K5 Q2K1 OBSERVED FIRE EXTINGUISHERS	These data suppressed for respondent confidentiality.
2102	K5CLDOBS	K5 Q4OBS NUMBER OF CHILDREN OBSERVED	These data recoded for respondent confidentiality.

SOURCE: U.S. Department of Education, National Center for Education Statistics, Early Childhood Longitudinal Study, Kindergarten Class of 1998–99 third grade data collection, school year 2001–02.

REFERENCES

Nord, M. and Bickel, G. (2001). Estimating the Prevalence of Children's Hunger from the Current Population Survey Food Security Supplement. In M. Andrews and M. Prell (Eds.), *Second Food Security Measurement and Research Conference, Volume II: Papers—Food Assistance and Nutrition Research Report 11-2* (p.31-49). Washington, DC: Economic Research Service, U.S. Department of Agriculture.

U.S. Department of Agriculture, Food and Nutrition Service (2000). *Guide to Measuring Household Food Security, Revised 2000.* Alexandria, VA: U.S. Department of Agriculture.

U.S. Department of Agriculture, Economic Research Service (2002). *Measuring Children's Food Security in U.S. Households, 1995-99.* Washington, DC: U.S. Department of Agriculture.

9. CREATING A LONGITUDINAL FILE

Longitudinal analyses with the Early Childhood Longitudinal Study, Kindergarten Class of 1998–99 (ECLS-K) can be conducted both “within school year” and “across school years.” Examples of within-year analyses are those that look at children’s growth in cognitive scores between fall and spring of kindergarten or between fall and spring of first grade. Such analyses do not require the combined use of kindergarten and first grade data. They can be conducted using just the kindergarten base year files or just the first grade files. Therefore, within-school year analyses are not discussed in this chapter. Since data were only collected once for third grade, no within-third grade analyses are possible. Cross-year analyses, on the other hand, are those that combine information from two or more of the kindergarten, first grade, or third grade years and are the focus of this chapter.

This chapter describes how to combine (or merge) the kindergarten, first grade, and third grade files to create cross-year files for K–3 longitudinal analyses. The information contained in this chapter applies to users of the base year, first grade, and third grade files. Users of the public-use files can consider using the public-use longitudinal file briefly described in chapter 1, which combines data from the base year, first, and third grades. It contains longitudinal weights so that analysts can examine children’s growth and development between kindergarten and third grade. Although it is somewhat streamlined, it contains most of the variables in the restricted-use files.

This chapter begins with a discussion of K–3 longitudinal analyses and the types of research questions that can be addressed with cross-year files. It then describes the merging procedures and the K–3 longitudinal weights.

9.1 Conducting Longitudinal Analyses

As described in chapter 1, one of the primary goals of the ECLS-K is to understand how children’s early experiences influence their transition into kindergarten and their progression through the early elementary school years. A major strength of the ECLS-K design is that it captures important aspects of children’s experiences as they occur. Thus, information about children’s transition into kindergarten is measured in the fall of their kindergarten year and again in the spring. Capturing this information as it occurs means that the information is not distorted by faulty memory or by revisions to memory based on subsequent experiences. In addition, information from earlier points in time can be used as predictors of later events and experiences, thereby strengthening the ability of researchers to make causal inferences.

In conducting K–3 longitudinal analyses with the ECLS-K data, it is important to keep in mind the sample design described in chapter 4. Certain features of the design must be considered. First, because the first and third grade data are only released as child-based files, all analyses involving either first grade or third grade data will, of necessity, be child-based. Second, the first and third grade data are not representative of all first grade or third grade schools or classrooms or teachers in the United States. Since the sample was not freshened in third grade, the children are not representative of all children attending third grade in the 2001–2002 school year. Children who started their schooling in the U.S. in second or third grade are not represented in the sample. Researchers conducting K–3 analyses should not attempt to use the data to describe the population of all third grade children, their classrooms, teachers, or schools. However, information about the schools can be used in the child-based analyses to examine, for example, the influence of the school environment on children's learning or to describe the learning environments of the group of children who attended kindergarten 3 years earlier. Users may also examine the influence of the kindergarten year school characteristics on children's later school experiences.

9.2 Examples of Research Questions

A variety of research questions can be examined using the K–3 longitudinal files. The following are some examples:

1. How much do children's reading and math skills increase between the fall of kindergarten and the spring of third grade?

2. Do measures of school readiness at the beginning of kindergarten predict children's skill and knowledge levels at the end of third grade?

3. What family background characteristics (e.g., family poverty, parent education, maternal employment) affect children's later school outcomes?

4. Do children who easily adapted to a school setting in kindergarten do better in third grade than their peers who experienced more difficulty settling into school or are there any lingering effects of a slow adjustment to kindergarten?

5. Are there particular school or classroom characteristics that enhance growth rates in reading and math skills between first and third grade?

To study these and similar questions, researchers would combine information from two or more rounds of data collection, across the kindergarten, first, and third grade years. For the first question, the researcher would need to examine differences between fall-kindergarten and spring-third grade assessment scores. To do this, one would combine fall-kindergarten data with spring-third grade data. Similarly, questions 2 and 3 (regarding the relationship between readiness at kindergarten entry—or maternal employment in that time frame—and third grade outcomes) would be examined by combining data from the same two time points. Note that for question 3 one would need to include data from the parent interview in the base year.

Researchers who want to examine the influence of children's kindergarten adjustment on their later grade performance, as in question 4, might use data from several rounds (i.e., fall-kindergarten, spring-kindergarten, spring-first grade, and spring-third grade). For example, one could create variables from fall-kindergarten and spring-kindergarten to measure adjustment during kindergarten and then relate those variables to outcomes in the spring of the first and third grades.

To be assigned a longitudinal weight for the K–3 data, a case must have participated in at least one of the base year rounds, and in both spring-first grade and spring-third grade. Thus, the K–3 longitudinal weights should not be used to examine questions that only use data from the base year and the first grade years. For such analyses, it is advisable to use the K–1 longitudinal weights.

9.3 Merging Base Year Child-Level Data With the First Grade and Third Grade Child-Level Data

To create a K–3 data file, which combines data from the base year, first grade, and third grade data collections, an analyst should use the ECLS-K Base Year Restricted-Use or Public-Use Electronic Code Book (NCES 2000–097 or NCES 2001-029); the ECLS-K First Grade Restricted-Use or Public-Use Electronic Code Book (NCES 2002–127 or NCES 2002–134); and the ECLS-K Third Grade Restricted-Use or Public-Use Electronic Code Book (NCES 2003-002 or NCES 2004-002). The same procedures can be followed by users who wish to create a K–3 longitudinal file themselves using public-use data files. To create a longitudinal file, perform the following steps to merge the base year child-level variables needed for analysis with the first grade and third grade child-level variables needed:

1. Select the variables to be analyzed from the base year ECB child catalog and the variable CHILDID. This creates a "working taglist" (see section 8.4 in chapter 8 for more detail on how to create a working taglist).

2. Run the program generated after extraction to create a base year data set (DATA1).

3. Using the child catalog from the First Grade Restricted-Use ECB, select the variables to be analyzed and the variable CHILDID.

4. Run the program generated after extraction to create a first grade data set (DATA2).

5. Using the child catalog from the Third Grade Restricted-Use ECB, select the variables to be analyzed and the variable CHILDID.

6. Run the program generated after extraction to create a third grade data set (DATA3).

7. Sort DATA1 and DATA2 and DATA3 by CHILDID.

8. Merge DATA1 and DATA2 and DATA3 by CHILDID.

This merged file will contain 21,409 cases, some of which will not have K–3 longitudinal weights. For example, base year respondents who did not participate in either fall or spring of first grade or spring of third grade, and movers who were not included in the first grade and third grade sample, will not have any K–3 longitudinal weights. To select cases with K–3 longitudinal data, a user can use a K–3 longitudinal weight appropriate to the analysis.

9.4 K–3 Longitudinal Weights

9.4.1 Types of K–3 Longitudinal Weights

K–3 longitudinal weights are used to analyze data in a K–3 file created by merging base year, first grade, and third grade data, whereas cross-sectional weights are used for analyses within one round of data collection. There are several sets of K–3 longitudinal weights computed for children with complete data from different combinations of rounds. All K–3 longitudinal weights are child-level weights. There are no K–3 longitudinal weights at the school or teacher level since school- and teacher-level weights are not computed for the first grade or third grade year. The K-3 longitudinal weights are defined as follows:

- C45CW0 is nonzero if assessment data are present for both spring-first grade and spring-third grade, or if the child was excluded from direct assessment in both of these rounds of data collection due to a disability;

- C45PW0 is nonzero if parent interview data are present for both spring-first grade and spring-third grade;

- C245CW0 is nonzero if assessment data are present for spring-kindergarten and spring-first grade and spring-third grade, or if the child was excluded from direct assessment in all of these three rounds of data collection due to a disability;

- C245PW0 is nonzero if parent interview data are present for spring-kindergarten and spring-first grade and spring-third grade;

- C1_5FC0 is nonzero if assessment data are present for four rounds of data collections involving the full sample of children (fall-kindergarten, spring-kindergarten, spring-first grade, and spring-third grade), or if the child was excluded from direct assessment in all four of these rounds of data collection due to a disability;

- C1_5FP0 is nonzero if parent interview data are present for four rounds of data collections involving the full sample of children (fall-kindergarten, spring-kindergarten, spring-first grade, and spring-third grade);

- C1_5SC0 is nonzero if assessment data are present for all five rounds of data collection (fall-kindergarten, spring-kindergarten, fall-first grade, spring-first grade, and spring-third grade), or if the child was excluded from direct assessment in all five rounds of data collection due to a disability; and

- C1_5SP0 is nonzero if parent interview data are present for all five rounds of data collection (fall-kindergarten, spring-kindergarten, fall-first grade, spring-first grade, and spring-third grade).

The use of the K–3 longitudinal weights, available on the third grade restricted-use ECB, is described in exhibit 9-1. This exhibit is designed to help users choose appropriate weights for their analysis. First, decide which two or more points in time are the focus of the analysis. The analysis could pertain to two points in time (spring-first grade and spring-third grade), three points in time (spring-kindergarten, spring-first grade, and spring-third grade), four points in time (fall-kindergarten, spring-kindergarten, spring-first grade, and spring-third grade), or five points in time (all five rounds of data collection). For example, if the analysis uses spring-first grade and spring-third grade data, then the appropriate weights would be those beginning with C45 (denoting child-level data from round 4, spring-first grade AND round 5, spring-third grade). Second, consider the source of the data, which also affects the choice of the weight. In exhibit 9-1, details under "to be used for analysis of …" provide guidance based on whether the data were collected through the child assessments, parent interviews, or teacher questionnaires A or B. For the same example noted earlier, the two weights available are C45CW0 and

C45PW0. If parent data from spring-first grade and spring-third grade are needed for the analysis, then C45PW0 should be used.

Base year longitudinal weights for the analysis of the base year data (within the kindergarten year) alone are described in the base year user's manuals. First grade longitudinal weights for the analysis of the first grade data (within the first grade year) alone, and of the combined kindergarten/first grade data are described in the first grade user's manuals.

K–3 longitudinal weights are used to produce estimates of differences between two or more rounds of data collection spanning across kindergarten, first grade, and third grade. Simple examples involving two rounds of data collection are the differences in children's mean assessment scores between spring-first grade and spring-third grade using the C45CW0 weight and the difference in the total number of persons in the household size using C45PW0. K–3 longitudinal weights are also used to study the characteristics of children who were assessed in two or more rounds of data collection. For example, one can study how family background characteristics of children in kindergarten affect assessment scores in spring-third grade for children who were assessed in spring-kindergarten, spring-first grade, and spring-third grade. In this case, C245PW0 is used to study the characteristics of the children as reported by their parents, and C245CW0 is used to estimate the difference in assessment scores between spring-kindergarten and spring-third grade. As noted earlier, any longitudinal analysis that uses data from fall-first grade will be limited to a 27 percent subsample of children**.**

Exhibit 9-1. ECLS-K: K–3 longitudinal weights, spring-third grade: School year 2001–02

Weight	to be used for analysis of ...
C45CW0	child direct assessment data from BOTH spring-first grade and spring-third grade, alone or in conjunction with any combination of a limited set of child characteristics (e.g., age, sex, race/ethnicity).
C45PW0	parent interview data from BOTH spring-first grade and spring-third grade.
C245CW0	child direct assessment data from spring-kindergarten AND spring-first grade AND spring-third grade, alone or in conjunction with any combination of a limited set of child characteristics (e.g., age, sex, race/ethnicity).
C245PW0	parent interview data from spring-kindergarten AND spring-first grade AND spring-third grade.
C1_5FC0	child direct assessment data from FOUR rounds of data collections involving the FULL sample of children (fall-kindergarten, spring-kindergarten, spring-first grade, spring-third grade), alone or in conjunction with any combination of a limited set of child characteristics (e.g., age, sex, race/ethnicity).
C1_5FP0	parent interview data from FOUR rounds of data collections involving the FULL sample of children (fall-kindergarten, spring-kindergarten, spring-first grade, spring-third grade)
C1_5SC0	child direct assessment data from ALL FIVE rounds of data collection, alone or in conjunction with any combination of a limited set of child characteristics (e.g., age, sex, and race/ethnicity).
C1_5SP0	parent interview data from ALL FIVE rounds of data collection.

SOURCE: U.S. Department of Education, National Center for Education Statistics, Early Childhood Longitudinal Study, Kindergarten Class of 1998–99 third grade data collection, school year 2001–02.

There may be combinations of data for which no weights were developed. For further advice on which weights to use when analyzing a complex combination of data, contact NCES at *ECLS@ed.gov*.

9.4.2 Weighting Procedures

In this section we discuss the statistical procedures used to produce the K–3 longitudinal weights. These procedures are nearly identical to the procedures used for the cross-sectional weights (see chapter 4). The differences are primarily in how mover status and eligible respondents are defined, and in how adjustment cells are created. For example, in computing weight C45CW0, a child was identified as a mover if the child moved in either spring-first grade or spring-third grade; a respondent was defined as a child for whom both cross-sectional weights, C4CW0 and C5CW0, are nonzero. A child with a nonzero C45CW0 had both spring-first grade and spring-third grade scorable cognitive assessment data, or was excluded from the cognitive assessments because he or she was a child with disabilities. Longitudinal weights involving the fall-first grade collections were computed differently to adjust for the fact that only a subsample of children were included in fall-first grade.

9.4.2.1 Longitudinal Weights Not Involving the Fall-First Grade Data

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

- Subsampling of movers and
- Child-level nonresponse.

In the adjustment for subsampling of movers, mover status was created so that it was specific to each panel. For example, for the spring-first grade/spring-third grade panel (longitudinal weights C45CW0 and C45PW0), a child was a mover if he or she was a mover in spring-first grade *or* spring-third grade, i.e., in either round he attended a school that was not the school where he was sampled in kindergarten. The adjustment factor for subsampling movers was computed within cells created using the following characteristics: whether children were sampled in kindergarten or first grade, whether they were movers in spring-first grade, whether they were language minority children, the school type of their original sample school, and the region where their original sample school was located. Eight children with large weights had their weights trimmed by half. However, the weights were not redistributed because the total sum of weights was re-established in the raking procedure that came later. In both steps of the nonresponse 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 belonged to the language minority group (all weights), and the type of household collected from the parent interviews (C45PW0, C245PW0, C1_5FP0 only).

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

9.4.2.2 Longitudinal Weights Involving the Fall-First Grade Data

For the longitudinal weights involving the fall-first grade data collection where children were part of a subsample of the ECLS-K full sample, the initial weights were from fall-first grade. These were the base year child adjusted weights (as described in section 4.6.3.2 for base year respondents), incorporating the school subsampling factor appropriate for fall-first grade. These weights were also trimmed to reduce the weight of all the children in one private school that had a large school weight.

The adjustments for subsampling movers and for child nonresponse are identical to those for the other longitudinal weights. The adjustment factor for subsampling movers was computed within cells created using the following characteristics: whether children moved in fall-first grade or spring-first grade and whether they belonged in the language minority group. One child with large weights had his weight trimmed by half. However, the weights were not redistributed because the total sum of weights was re-established in the raking procedure that came later. In both steps of the nonresponse adjustment for the C1_5SC0 weight, separate nonresponse classes were created for movers and nonmovers using the parent interview response status from the base year as well as whether children belonged to the language minority group. For the C1_5SP0 weight, nonresponse classes were created using the type of household collected from the parent interviews and whether children belonged to the language minority group.

The raking dimensions are the same as those for the other longitudinal weights. After the first raking for the C1_5SC0 weight, four children had their weights trimmed, then all the weights were raked again. Only one raking was necessary for the C1_5SP0 weight.

9.4.3 Characteristics of Longitudinal Weights

The statistical characteristics of the longitudinal weights are presented in table 9-1. For each weight, the number of cases with nonzero values is presented together with the mean weight, the standard deviation, the coefficient of variation (i.e., the standard deviation as a percentage of the mean weight), the minimum value of the weight, the maximum value of the weight, the skewness, the kurtosis, and the sum of weights.

Table 9-1. Characteristics of child-level K–3 longitudinal weights, spring-third grade: School year 2001–02

Variable name	Number of cases	Mean	Standard deviation	CV[1] (× 100)	Minimum	Maximum	Skewness	Kurtosis	Sum
C45CW0	13,964	281.86	273.52	97.04	1.68	3,897.42	3.37	19.90	3,935,960
C45PW0	12,652	310.98	266.89	85.82	1.68	3,718.34	3.11	17.32	3,934,550
C245CW0	13,694	280.68	277.47	98.86	1.65	4,119.55	3.55	22.53	3,843,642
C245PW0	12,204	314.92	267.05	84.80	1.78	3,121.66	2.87	14.51	3,843,272
C1_5FC0	12,558	306.07	303.52	99.17	1.68	4,264.25	3.59	22.83	3,843,607
C1_5FP0	10,998	349.42	299.17	85.62	1.92	3,754.91	3.18	17.88	3,842,954
C1_5SC0	4,032	952.67	875.12	91.86	64.97	7,174.65	3.28	13.78	3,841,183
C1_5SP0	3,522	1,090.37	816.79	74.91	104.68	6,801.61	2.56	9.19	3,840,278

[1] Coefficient of variation.
SOURCE: U.S. Department of Education, National Center for Education Statistics, Early Childhood Longitudinal Study, Kindergarten Class of 1998–99 third grade data collection, school year 2001–02.

The difference in the estimate of the population of students (sum of weights) between the different panels of students and types of weights is due to a combination of factors, among them: (1) the number of base year respondents who became ineligible (due to death, leaving the country, or being a nonsampled mover) after the base year, (2) the adjustment of the weights for the children of unknown eligibility, and (3) the difference in the number of records used to construct sample-based control totals. Of the eight longitudinal weights computed, only the first two (C45CW0 and C45PW0) involve children sampled in first grade. For these two weights, the child records included in the file used for computing the control totals are records of base year respondents and records of eligible children sampled in first grade. For all other longitudinal weights, records of children sampled in first grade were not included in the file, causing the sum of weights to be smaller.

9.4.4 Variance Estimation

For each K–3 full sample weight listed in exhibit 9-1, a set of replicate weights was calculated. Replicate weights are used in the jackknife replication method to estimate the standard errors of survey estimates. Any adjustments done to the full sample weights were repeated for the replicate weights.

For longitudinal weights not involving the fall-first grade data, there are 90 replicate weights. For a description of how the replicates were formed, see chapter 4, section 4.7. For the two longitudinal weights involving fall-first grade (C1_5SC0 and C1_5SP0), there are 40 replicate weights. The smaller number of replicates was due to the fact that only a subsample of schools was included in the fall-first grade sample. The weights associated with the fall-first grade data do not account for the Durbin method of selecting primary sampling units (PSUs), since it no longer applied. Rather, they reflect the fact that only one of the two sampled PSUs in the non-self-representing (NSR) strata was kept in the subsample. To account for this feature, pairs of similar NSR PSUs were collapsed into 19 variance strata. The self-representing (SR) PSUs account for the remaining 21 variance strata.

Each replicate weight variable name has the same weight prefix as for the full sample weight variable name. For example, the replicate weights for C1_5FC0 are C1_5FC1 through C1_5FC90; the replicate weights for C1_5SC0 are C1_5SC1 through C1_5SC40.

Stratum and first-stage unit identifiers used with the Taylor Series method are provided for each of the K–3 longitudinal weights in the file. They are described in exhibit 9-2. For a description of the Taylor Series method, see chapter 4, section 4.7.2.

Specifications for computing standard errors are given in table 9-2. For each type of analysis described in table 9-2, users can choose between the replication method and the Taylor Series method for computing standard errors.

Exhibit 9-2. ECLS-K Taylor Series stratum and first-stage unit identifiers, spring-third grade: School year 2001–02

Variable name	Description
C45CSTR	Sampling stratum—spring-first grade/spring-third grade longitudinal C-weights
C45CPSU	First-stage primary sampling unit within stratum—spring-first grade/spring-third grade longitudinal C-weights
C45PSTR	Sampling stratum—spring-first grade/spring-third grade longitudinal P-weights
C45PPSU	First-stage primary sampling unit within stratum—spring-first grade/spring-third grade longitudinal P-weights
C245CSTR	Sampling stratum—spring-kindergarten/spring-first grade/spring-third grade longitudinal C-weights
C245CPSU	First-stage primary sampling unit within stratum—spring-kindergarten/spring-first grade/spring-third grade longitudinal C-weights
C245PSTR	Sampling stratum—spring-kindergarten/spring-first grade/spring-third grade longitudinal P-weights
C245PPSU	First-stage primary sampling unit within stratum—spring-kindergarten/spring-first grade/spring-third grade longitudinal P-weights
C15FCSTR	Sampling stratum—fall-kindergarten/spring-kindergarten/spring-first grade/spring-third grade longitudinal C-weights
C15FCPSU	First-stage primary sampling unit within stratum—fall-kindergarten/spring-kindergarten/spring-first grade/spring-third grade longitudinal C-weights
C15FPSTR	Sampling stratum—fall-kindergarten/spring-kindergarten/spring-first grade/spring-third grade longitudinal P-weights
C15FPPSU	First-stage primary sampling unit within stratum—fall-kindergarten/spring-kindergarten/spring-first grade/spring-third grade longitudinal P-weights
C15SCSTR	Sampling stratum—longitudinal C-weights covering all five rounds of data collection
C15SCPSU	First-stage primary sampling unit within stratum—longitudinal C-weights covering all five rounds of data collection
C15SPSTR	Sampling stratum—longitudinal P-weights covering all five rounds of data collection
C15SPPSU	First-stage primary sampling unit within stratum—longitudinal P-weights covering all five rounds of data collection

SOURCE: U.S. Department of Education, National Center for Education Statistics, Early Childhood Longitudinal Study, Kindergarten Class of 1998–99 third grade data collection, school year 2001–02.

Table 9-2. Specifications for computing standard errors, spring third-grade: School year 2001–02

Type of analysis	Full sample weight	Computing standard errors					Approximating sampling errors
		Replication method (WesVar, SUDAAN or AM)			Taylor Series method (SUDAAN, Stata, SAS or AM)		DEFT (Average root design effect)
		ID	Replicate weights	Jackknife method	Sample design	Nesting variables	
Spring-first grade/spring-third grade longitudinal							1.815
	C45CW0	CHILDID	C45CW1 – C45CW90	JK2	WR[1]	C45CSTR C45CPSU	
	C45PW0	PARENTID	C45PW1 – C45PW90	JK2	WR	C45PSTR C45PPSU	
Spring-kindergarten/ spring-first grade/ spring-third longitudinal							1.791
	C245CW0	CHILDID	C245CW1 – C245CW90	JK2	WR	C245CSTR C245CPSU	
	C245PW0	PARENTID	C245PW1 – C245PW90	JK2	WR	C245PSTR C245PPSU	
Fall-kindergarten/ spring-kindergarten/ spring- first grade/ spring- third grade longitudinal							1.749
	C1_5FC0	CHILDID	C1_5FC1 – C1_5FC90	JK2	WR	C15FCSTR C15FCPSU	
	C1_5FP0	PARENTID	C1_5FP1 – C1_5FP90	JK2	WR	C15FPSTR C15FPPSU	
All five round longitudinal							1.635
	C1_5SC0	CHILDID	C1_5SC1 – C1_5SC40	JK2	WR	C15SCSTR C15SCPSU	
	C1_5SP0	PARENTID	C1_5SP1 – C1_5SP40	JK2	WR	C15SPSTR C15SPPSU	

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

SOURCE: U.S. Department of Education, National Center for Education Statistics, Early Childhood Longitudinal Study, Kindergarten Class of 1998–99 third grade data collection, school year 2001–02.

For the replication method using WesVar, the full sample weight, the replicate weights, and the method of replication are required parameters. Variance estimation using the ECLS-K data should be done using the paired jackknife method (JK2). As an example, to compute the mean difference in reading scores between spring-kindergarten and spring-first grade and their standard errors, users need to specify C45CW0 as the full sample weight, C45CW1 to C45CW90 as the replicate weights, and JK2 as the method of replication.

For the Taylor Series method using SUDAAN, SAS, Stata, or AM the full sample weight, the sample design, the nesting stratum, and PSU variables are required. For the same example earlier, the full sample weight (C45CW0), the stratum variable (C45CSTR), and the PSU variable (C45CPSU) must be specified. The “with replacement” sample design option, WR, must also be specified if using SUDAAN.

9.4.5 Design Effects

An important analytic device is to compare the statistical efficiency of survey estimates with what would have been obtained in a hypothetical and usually impractical simple random sample (SRS) of the same size. For a discussion of design effects and their use, see chapter 4, section 4.8. In this section, design effects are presented for selected illustrative estimates produced using kindergarten-first grade longitudinal weights. The tables that follow show estimates, standard errors, and design effects for selected means and proportions based on the ECLS-K child and parent data. For each survey item, the tables present the number of cases, the estimate, the standard error taking into account the actual sample design (Design SE), the standard error assuming SRS (SRS SE), the root design effect (DEFT), and the design effect (DEFF). Standard errors (Design SE) were produced using JK2.

Standard errors and design effects are presented in tables 9-3 to 9-6. Data items are from the direct child assessment, the parent interview, and the teacher child-level questionnaire. Full sample weights were used to compute the estimates; then the corresponding replicate weights were used to compute standard errors and design effects.

Table 9-3. ECLS-K, spring-first grade/spring-third grade panel: standard errors and design effects using C45CW0-C45CW90 and C45PW0-C45PW90, by selected child and parent variables: School years 1999–2000 and 2001–02

Survey item	Variable name	Number of cases	Estimate	Design SE[1]	SRS SE[2]	DEFT[3]	DEFF[4]
Difference between spring-first grade and spring-third grade scores (mean)							
Reading scale score	C5R2RSCL-C4R2RSCL	13,504	39.41	0.285	0.130	2.195	4.819
Math scale score	C5R2MSCL- C4R2MSCL	13,842	29.36	0.229	0.097	2.366	5.600
Approaches to learning-Teacher	T5LEARN- T4LEARN	10,526	-0.02	0.010	0.006	1.546	2.390
Self-control-Teacher	T5CONTRO- T4CONTRO	10,351	0.01	0.010	0.007	1.490	2.219
Interpersonal-Teacher	T5INTERP- T4INTERP	10,311	-0.04	0.011	0.007	1.602	2.565
Externalizing problems-Teacher	T5EXTERN- T4EXTERN	10,451	0.07	0.010	0.006	1.661	2.759
Internalizing problems-Teacher	T5INTERN- T4INTERN	10,317	0.06	0.011	0.007	1.681	2.825
Other differences							
Child’s Body Mass Index (BMI)	C5BMI-C4BMI	12,588	1.77	0.032	0.018	1.791	3.209
Child’s height	C5HEIGHT-C4HEIGHT	12,950	4.67	0.022	0.013	1.751	3.067
Child’s weight	C5WEIGHT-C4WEIGHT	12,735	18.88	0.142	0.089	1.604	2.574
Household size	P5HTOTAL-P4HTOTAL	12,650	0.04	0.011	0.008	1.451	2.106
Number of hours watched TV after dinner	P5TVAFDH-P4TVAFDH	12,377	0.10	0.013	0.008	1.562	2.439
Child and parent characteristics from parent interview (percent)							
Lived in single parent family	P5HFAMIL	12,652	25.18	0.699	0.386	1.812	3.285
Lived in two-parent family	P5HFAMIL	12,652	72.48	0.784	0.397	1.975	3.901
Mom worked 35 hours+/week	P5HMEMP	9,265	67.07	0.826	0.488	1.691	2.859
Primary care is center-based	P5PRIMNW	4,462	36.46	1.357	0.721	1.883	3.545
Primary care is home-based	P5PRIMNW	4,462	63.54	1.357	0.721	1.883	3.545
Parents had high school or less	W3PARED	12,652	32.63	0.871	0.417	2.089	4.364
Household income category below median	W3INCCAT	12,652	41.82	1.038	0.439	2.366	5.598
Parent attended PTA	P5ATTENP	12,636	43.55	0.941	0.441	2.133	4.550
Practiced reading, writing, numbers daily	P5RDWRNM	12,557	49.44	0.602	0.446	1.350	1.823
Visited library	P5LIBRAR	12,558	54.35	0.798	0.445	1.795	3.222
Used computer 1-2 times per week	P5HOMECM, P5COMPWK	10,121	41.82	0.648	0.491	1.321	1.746
Had Internet access	P5HOMECM, P5INTACC	9,883	86.06	0.431	0.348	1.237	1.529
Used computer 1-2 times/week for homework	P5HOMECM, P5CMPEDU	9,862	53.06	0.639	0.503	1.271	1.615
Had family rule for TV	P5TVHOME, P5TVRULE	12,488	91.68	0.315	0.247	1.273	1.620
Did homework 3-4 times per week	P5OFTDHW	12,539	40.40	0.694	0.438	1.583	2.506
Have someone help with reading homework	P5HELPR	12,370	98.52	0.177	0.108	1.633	2.666
Completely true child and self have close time	P5WARMCL	12,263	68.49	0.586	0.419	1.397	1.953
Took away privilege when child angry	P5HITPRV	12,250	62.23	1.011	0.438	2.308	5.328
Self-reported in very good health	P5HEALTH	12,240	86.77	0.542	0.306	1.769	3.128
HH received food stamps in last 12 months	P5FSTAMP	12,472	14.03	0.736	0.311	2.367	5.604
Child characteristics from teacher questionnaire C (percent)							
Enrolled in third grade	T5GRADE	11,397	88.25	0.672	0.301	2.229	4.969
Average in language skills	T5RTLANG	11,262	70.83	0.784	0.428	1.831	3.353
Average in science/social studies	T5RTSCI	11,180	79.63	0.918	0.381	2.411	5.812
Average in math skills	T5RTMTH	11,227	77.37	0.805	0.395	2.040	4.160

See notes at end of table.

Table 9-3. ECLS-K, spring-first grade/spring-third grade panel: standard errors and design effects using C45CW0-C45CW90 and C45PW0-C45PW90, by selected child and parent variables: School years 1999–2000 and 2001–02—Continued

Survey item	Variable name	Number of cases	Estimate	Design SE[1]	SRS SE[2]	DEFT[3]	DEFF[4]
Child characteristics (mean)							
Age of child in months	R5AGE	13,901	111.21	0.093	0.040	2.299	5.285
Child's BMI	C5BMI	13,417	18.69	0.045	0.033	1.366	1.866
Child's household size	P5HTOTAL	12,652	4.58	0.025	0.012	2.015	4.059
Number of children <18 in child's HH	P5LESS18	12,652	2.52	0.021	0.010	2.099	4.404
Number of siblings in HH	P5NUMSIB	12,652	1.56	0.020	0.010	1.968	3.874
Number of hours watched TV after dinner	P5TVAFDH	12,470	0.88	0.012	0.007	1.656	2.742
Median						1.769	3.169
Mean						1.815	3.438
Standard deviation						0.366	1.396
Coefficient of variation						0.202	0.406
Minimum						1.237	1.529
Maximum						2.685	7.209

[1]Design SE is the standard error under the ECLS-K sample design. For an explanation of this statistic, see chapter 4, section 4.8.
[2] SRS SE is the standard error assuming simple random sample. For an explanation of this statistic, see chapter 4, section 4.8.
[3] DEFT is the root design effect. For an explanation of this statistic, see chapter 4, section 4.8.
[4] DEFF is the design effect. For an explanation of this statistic, see chapter 4, section 4.8.
SOURCE: U.S. Department of Education, National Center for Education Statistics, Early Childhood Longitudinal Study, Kindergarten Class of 1998–99 first and third grade data collections, school years 1999–2000 and 2001–02.

Table 9-4. ECLS-K, spring-kindergarten/spring-first grade/spring-third grade panel: standard errors and design effects using C245CW0-C245CW90 and C245PW0-C245PW90, by selected child and parent variables: School years 1998–99, 1999–2000, and 2001–02

Survey item	Variable name	Number of cases	Estimate	Design SE[1]	SRS SE[2]	DEFT[3]	DEFF[4]
Difference between spring-first grade and spring-third grade scores (mean)							
Reading scale score	C5R2RSCL-C4R2RSCL	13,260	39.39	0.278	0.131	2.126	4.518
Math scale score	C5R2MSCL- C4R2MSCL	13,579	29.39	0.238	0.097	2.445	5.976
Approaches to learning-Teacher	T5LEARN- T4LEARN	10,376	-0.02	0.011	0.007	1.670	2.790
Self-control-Teacher	T5CONTRO- T4CONTRO	10,206	0.01	0.010	0.006	1.553	2.411
Interpersonal-Teacher	T5INTERP- T4INTERP	10,165	-0.04	0.012	0.007	1.737	3.016
Externalizing problems-Teacher	T5EXTERN- T4EXTERN	10,300	0.06	0.009	0.006	1.594	2.541
Internalizing problems-Teacher	T5INTERN- T4INTERN	10,175	0.06	0.010	0.007	1.534	2.353
Difference between spring-kindergarten and spring-third grade scores (mean)							
Reading scale score	C5R2RSCL-C2R2RSCL	12,867	69.03	0.356	0.147	2.423	5.873
Math scale score	C5R2MSCL- C2R2MSCL	13,451	52.60	0.295	0.110	2.692	7.248
Approaches to learning-Teacher	T5LEARN- T2LEARN	10,836	-0.10	0.010	0.007	1.536	2.360
Self-control-Teacher	T5CONTRO- T2CONTRO	10,674	-0.01	0.010	0.006	1.553	2.411
Interpersonal-Teacher	T5INTERP- T2INTERP	10,609	-0.06	0.010	0.007	1.427	2.037
Externalizing problems-Teacher	T5EXTERN- T2EXTERN	10,770	0.07	0.010	0.006	1.592	2.535
Internalizing problems-Teacher	T5INTERN- T2INTERN	10,648	0.10	0.013	0.006	2.025	4.099
Other differences							
Child's BMI	C5BMI-C4BMI	12,370	1.76	0.032	0.018	1.764	3.111
Child's height	C5HEIGHT-C4HEIGHT	12,727	4.66	0.022	0.013	1.738	3.019
Child's weight	C5WEIGHT-C4WEIGHT	12,507	18.77	0.137	0.089	1.536	2.360
Household size	P5HTOTAL-P4HTOTAL	12,202	0.04	0.012	0.008	1.442	2.080
Number of hours watched TV after dinner	P5TVAFDH-P4TVAFDH	11,962	0.10	0.015	0.009	1.664	2.768
Child and parent characteristics from parent interview (percent)							
Lived in single parent family	P5HFAMIL	12,204	25.08	0.706	0.392	1.800	3.240
Lived in two-parent family	P5HFAMIL	12,204	72.53	0.806	0.404	1.994	3.975
Mom worked 35 hours+/week	P5HMEMP	8,996	66.99	0.782	0.496	1.578	2.490
Primary care is center-based	P5PRIMNW	4,308	36.83	1.405	0.735	1.912	3.657
Primary care is home-based	P5PRIMNW	4,308	63.17	1.405	0.735	1.912	3.657
Parents had high school or less	W3PARED	12,204	32.11	0.840	0.423	1.987	3.948
Household income category below median	W3INCCAT	12,204	41.48	1.053	0.446	2.361	5.575
Parent attended PTA	P5ATTENP	12,188	43.81	0.907	0.450	2.017	4.070
Practiced reading, writing, numbers daily	P5RDWRNM	12,124	49.48	0.630	0.455	1.386	1.922
Visited library	P5LIBRAR	12,124	54.86	0.810	0.452	1.793	3.215
Used computer 1-2 times per week	P5HOMECM, P5COMPWK	9,848	41.99	0.690	0.497	1.387	1.924
Had Internet access	P5HOMECM, P5INTACC	9,615	86.08	0.458	0.353	1.297	1.683
Used computer 1-2 times/week for homework	P5HOMECM, P5CMPEDU	9,594	53.05	0.606	0.510	1.189	1.414
Had family rule for TV	P5TVHOME, P5TVRULE	12,061	91.72	0.335	0.251	1.335	1.781
Did homework 3-4 times per week	P5OFTDHW	12,107	40.73	0.701	0.447	1.569	2.462
Have someone help with reading homework	P5HELPR	11,941	98.52	0.175	0.111	1.580	2.497
Completely true child and self have close time	P5WARMCL	11,854	68.53	0.595	0.426	1.396	1.948
Took away privilege when child angry	P5HITPRV	11,840	62.23	1.010	0.446	2.267	5.140
Self-reported in very good health	P5HEALTH	11,835	87.07	0.515	0.308	1.670	2.790
HH received food stamps in last 12 months	P5FSTAMP	12,050	13.93	0.740	0.315	2.346	5.504

See notes at end of table.

Table 9-4. ECLS-K, spring-kindergarten/spring-first grade/spring-third grade panel: standard errors and design effects using C245CW0-C245CW90 and C245PW0-C245PW90, by selected child and parent variables: School years 1998–99, 1999–2000, and 2001–02—Continued

Survey item	Variable name	Number of cases	Estimate	Design SE[1]	SRS SE[2]	DEFT[3]	DEFF[4]
Child characteristics from teacher questionnaire C (percent)							
Enrolled in third grade	T5GRADE	11,209	88.30	0.690	0.304	2.271	5.159
Average in language skills	T5RTLANG	11,080	71.36	0.751	0.429	1.750	3.061
Average in science/social studies	T5RTSCI	10,997	80.16	0.790	0.381	2.076	4.311
Average in math skills	T5RTMTH	11,046	77.66	0.723	0.396	1.825	3.331
Child characteristics (mean)							
Age of child in months	R5AGE	13,636	111.03	0.080	0.038	2.094	4.383
Child's BMI	C5BMI	13,167	18.66	0.045	0.033	1.362	1.855
Child's household size	P5HTOTAL	12,204	4.57	0.025	0.013	1.976	3.903
Number of children <18 in child's HH	P5LESS18	12,204	2.52	0.021	0.010	2.046	4.188
Number of siblings in HH	P5NUMSIB	12,204	1.55	0.020	0.010	1.927	3.714
Number of hours watched TV after dinner	P5TVAFDH	12,045	0.88	0.012	0.007	1.603	2.569
Median						1.738	3.040
Mean						1.791	3.340
Standard deviation						0.345	1.315
Coefficient of variation						0.192	0.394
Minimum						1.189	1.414
Maximum						2.692	7.248

[1] Design SE is the standard error under the ECLS-K sample design. For an explanation of this statistic, see chapter 4, section 4.8.
[2] SRS SE is the standard error assuming simple random sample. For an explanation of this statistic, see chapter 4, section 4.8.
[3] DEFT is the root design effect. For an explanation of this statistic, see chapter 4, section 4.8.
[4] DEFF is the design effect. For an explanation of this statistic, see chapter 4, section 4.8.
SOURCE: U.S. Department of Education, National Center for Education Statistics, Early Childhood Longitudinal Study, Kindergarten Class of 1998–99 kindergarten, first grade, and third grade data collections, school years 1998–99, 1999–2000, and 2001–02.

Table 9-5. ECLS-K, fall-kindergarten/spring-kindergarten/spring-first grade/spring-third grade panel: standard errors and design effects using C1_5FC0-C1_5FC90 and C1_5FP0-C1_5FP90, by selected child and parent variables: School years 1998–99, 1999–2000, and 2001–02

Survey item	Variable name	Number of cases	Estimate	Design SE[1]	SRS SE[2]	DEFT[3]	DEFF[4]
Difference between spring-first grade and spring-third grade scores (mean)							
Reading scale score	C5R2RSCL-C4R2RSCL	12,170	39.43	0.290	0.136	2.125	4.515
Math scale score	C5R2MSCL- C4R2MSCL	12,454	29.42	0.224	0.102	2.206	4.867
Approaches to learning-Teacher	T5LEARN- T4LEARN	9,464	-0.02	0.010	0.007	1.464	2.144
Self-control-Teacher	T5CONTRO- T4CONTRO	9,314	0.02	0.010	0.007	1.496	2.239
Interpersonal-Teacher	T5INTERP- T4INTERP	9,273	-0.04	0.012	0.007	1.616	2.612
Externalizing problems-Teacher	T5EXTERN- T4EXTERN	9,395	0.07	0.009	0.006	1.494	2.233
Internalizing problems-Teacher	T5INTERN- T4INTERN	9,277	0.05	0.011	0.007	1.670	2.788
Difference between spring-kindergarten and spring-third grade scores (mean)							
Reading scale score	C5R2RSCL-C2R2RSCL	11,816	69.04	0.358	0.153	2.345	5.499
Math scale score	C5R2MSCL- C2R2MSCL	12,347	52.57	0.271	0.114	2.377	5.651
Approaches to learning-Teacher	T5LEARN- T2LEARN	9,918	-0.10	0.011	0.007	1.533	2.350
Self-control-Teacher	T5CONTRO- T2CONTRO	9,767	0.00	0.011	0.007	1.496	2.239
Interpersonal-Teacher	T5INTERP- T2INTERP	9,710	-0.06	0.011	0.008	1.433	2.053
Externalizing problems-Teacher	T5EXTERN- T2EXTERN	9,851	0.07	0.010	0.007	1.523	2.320
Internalizing problems-Teacher	T5INTERN- T2INTERN	9,741	0.09	0.014	0.007	2.095	4.387
Other differences							
Child's BMI	C5BMI-C4BMI	11,312	1.76	0.032	0.019	1.678	2.815
Child's height	C5HEIGHT-C4HEIGHT	11,659	4.64	0.022	0.013	1.721	2.963
Child's weight	C5WEIGHT-C4WEIGHT	11,440	18.72	0.135	0.093	1.444	2.086
Household size	P5HTOTAL-P4HTOTAL	10,996	0.04	0.011	0.008	1.318	1.736
Number of hours watched TV after dinner	P5TVAFDH-P4TVAFDH	10,791	0.11	0.017	0.009	1.823	3.322
Child and parent characteristics from parent interview (percent)							
Lived in single parent family	P5HFAMIL	10,998	25.07	0.691	0.413	1.672	2.794
Lived in two-parent family	P5HFAMIL	10,998	72.61	0.800	0.426	1.880	3.536
Mom worked 35 hours+/week	P5HMEMP	8,115	67.28	0.836	0.521	1.604	2.573
Primary care is center-based	P5PRIMNW	3,883	36.96	1.490	0.775	1.923	3.699
Primary care is home-based	P5PRIMNW	3,883	63.04	1.490	0.775	1.923	3.699
Parents had high school or less	W3PARED	10,998	32.01	0.828	0.445	1.861	3.463
Household income category below median	W3INCCAT	10,998	41.11	1.059	0.469	2.257	5.092
Parent attended PTA	P5ATTENP	10,983	43.97	0.928	0.474	1.959	3.837
Practiced reading, writing, numbers daily	P5RDWRNM	10,925	49.16	0.644	0.478	1.346	1.813
Visited library	P5LIBRAR	10,926	55.23	0.855	0.476	1.798	3.234
Used computer 1-2 times per week	P5HOMECM, P5COMPWK	8,934	42.18	0.727	0.523	1.391	1.935
Had Internet access	P5HOMECM, P5INTACC	8,729	86.36	0.432	0.368	1.175	1.381
Used computer 1-2 times/week for homework	P5HOMECM, P5CMPEDU	8,709	53.40	0.648	0.534	1.213	1.471
Had family rule for TV	P5TVHOME, P5TVRULE	10,878	91.60	0.369	0.266	1.389	1.928
Did homework 3-4 times per week	P5OFTDHW	10,909	40.85	0.770	0.471	1.636	2.675
Have someone help with reading homework	P5HELPR	10,764	98.45	0.211	0.119	1.770	3.132
Completely true child and self have close time	P5WARMCL	10,687	68.54	0.652	0.449	1.451	2.106
Took away privilege when child angry	P5HITPRV	10,679	62.19	1.019	0.469	2.172	4.718
Self-reported in very good health	P5HEALTH	10,668	87.13	0.525	0.324	1.620	2.624
HH received food stamps in last 12 months	P5FSTAMP	10,857	13.68	0.755	0.330	2.288	5.234

See notes at end of table.

Table 9-5. ECLS-K, fall-kindergarten/spring-kindergarten/spring-first grade/spring-third grade panel: standard errors and design effects using C1_5FC0-C1_5FC90 and C1_5FP0-C1_5FP90, by selected child and parent variables: School years 1998–99, 1999–2000, and 2001–02—Continued

Survey item	Variable name	Number of cases	Estimate	Design SE[1]	SRS SE[2]	DEFT[3]	DEFF[4]
Child characteristics from teacher questionnaire C (percent)							
Enrolled in third grade	T5GRADE	10,225	88.65	0.703	0.314	2.242	5.027
Average in language skills	T5RTLANG	10,098	71.61	0.760	0.449	1.694	2.870
Average in science/social studies	T5RTSCI	10,018	80.51	0.803	0.396	2.029	4.116
Average in math skills	T5RTMTH	10,070	77.92	0.763	0.413	1.847	3.410
Child characteristics (mean)							
Age of child in months	R5AGE	12,507	111.02	0.079	0.040	1.989	3.957
Child's BMI	C5BMI	12,074	18.64	0.047	0.035	1.343	1.804
Child's household size	P5HTOTAL	10,998	4.58	0.025	0.013	1.934	3.740
Number of children <18 in child's HH	P5LESS18	10,998	2.53	0.022	0.011	2.013	4.051
Number of siblings in HH	P5NUMSIB	10,998	1.56	0.020	0.011	1.871	3.502
Number of hours watched TV after dinner	P5TVAFDH	10,863	0.88	0.012	0.008	1.550	2.404
Median						1.694	2.917
Mean						1.749	3.172
Standard deviation						0.315	1.138
Coefficient of variation						0.180	0.359
Minimum						1.175	1.381
Maximum						2.377	5.651

[1]Design SE is the standard error under the ECLS-K sample design. For an explanation of this statistic, see chapter 4, section 4.8.
[2] SRS SE is the standard error assuming simple random sample. For an explanation of this statistic, see chapter 4, section 4.8.
[3] DEFT is the root design effect. For an explanation of this statistic, see chapter 4, section 4.8.
[4] DEFF is the design effect. For an explanation of this statistic, see chapter 4, section 4.8.
SOURCE: U.S. Department of Education, National Center for Education Statistics, Early Childhood Longitudinal Study, Kindergarten Class of 1998–99 kindergarten, first grade, and third grade data collections, school years 1998–99, 1999–2000, and 2001–02.

Table 9-6. ECLS-K, panel of all five rounds: standard errors and design effects for the full sample using C1_5SC0-C1_5SC40 and C1_5SP0-C1_5SP40, by selected child and parent variables: School years 1998–99, 1999-2000, and 2001–02

Survey item	Variable name	Number of cases	Estimate	Design SE[1]	SRS SE[2]	DEFT[3]	DEFF[4]
Difference between spring-first grade and spring-third grade scores (mean)							
Reading scale score	C5R2RSCL-C4R2RSCL	3,903	40.01	0.387	0.247	1.567	2.455
Math scale score	C5R2MSCL- C4R2MSCL	4,005	29.22	0.384	0.180	2.131	4.543
Approaches to learning-Teacher	T5LEARN- T4LEARN	2,803	-0.03	0.024	0.012	1.957	3.830
Self-control-Teacher	T5CONTRO- T4CONTRO	2,748	0.01	0.022	0.012	1.823	3.323
Interpersonal-Teacher	T5INTERP- T4INTERP	2,736	-0.06	0.021	0.013	1.634	2.670
Externalizing problems-Teacher	T5EXTERN- T4EXTERN	2,787	0.06	0.020	0.011	1.783	3.180
Internalizing problems-Teacher	T5INTERN- T4INTERN	2,733	0.06	0.017	0.013	1.351	1.825
Difference between spring-kindergarten and spring-third grade scores (mean)							
Reading scale score	C5R2RSCL-C2R2RSCL	3,800	68.97	0.531	0.267	1.987	3.947
Math scale score	C5R2MSCL- C2R2MSCL	3,976	52.66	0.433	0.202	2.144	4.596
Approaches to learning-Teacher	T5LEARN- T2LEARN	2,989	-0.10	0.021	0.013	1.602	2.567
Self-control-Teacher	T5CONTRO- T2CONTRO	2,942	-0.01	0.017	0.009	1.823	3.323
Interpersonal-Teacher	T5INTERP- T2INTERP	2,903	-0.05	0.019	0.013	1.426	2.033
Externalizing problems-Teacher	T5EXTERN- T2EXTERN	2,976	0.06	0.021	0.011	1.842	3.392
Internalizing problems-Teacher	T5INTERN- T2INTERN	2,925	0.09	0.020	0.012	1.616	2.612
Other differences							
Child's BMI	C5BMI-C4BMI	3,704	1.72	0.061	0.032	1.911	3.651
Child's height	C5HEIGHT-C4HEIGHT	3,760	4.59	0.039	0.022	1.806	3.263
Child's weight	C5WEIGHT-C4WEIGHT	3,723	18.36	0.255	0.157	1.622	2.632
Household size	P5HTOTAL-P4HTOTAL	3,521	0.02	0.022	0.015	1.467	2.153
Number of hours watched TV after dinner	P5TVAFDH-P4TVAFDH	3,451	0.11	0.023	0.016	1.420	2.017
Child and parent characteristics from parent interview (percent)							
Lived in single parent family	P5HFAMIL	3,522	26.27	0.893	0.742	1.204	1.449
Lived in two-parent family	P5HFAMIL	3,522	71.34	1.020	0.762	1.339	1.792
Mom worked 35 hours+/week	P5HMEMP	2,581	66.45	1.195	0.929	1.286	1.653
Primary care is center-based	P5PRIMNW	1,184	34.49	1.947	1.382	1.409	1.986
Primary care is home-based	P5PRIMNW	1,184	65.51	1.947	1.382	1.409	1.986
Parents had high school or less	W3PARED	3,522	31.46	0.864	0.783	1.104	1.219
Household income category below median	W3INCCAT	3,522	41.22	1.672	0.829	2.016	4.064
Parent attended PTA	P5ATTENP	3,519	45.22	1.692	0.839	2.017	4.067
Practiced reading, writing, numbers daily	P5RDWRNM	3,499	48.29	1.276	0.845	1.510	2.280
Visited library	P5LIBRAR	3,498	55.58	1.549	0.840	1.844	3.400
Used computer 1-2 times per week	P5HOMECM, P5COMPWK	2,836	43.98	1.011	0.932	1.085	1.177
Had Internet access	P5HOMECM, P5INTACC	2,768	86.19	0.980	0.656	1.494	2.231
Used computer 1-2 times/week for homework	P5HOMECM, P5CMPEDU	2,765	53.00	1.385	0.949	1.460	2.131
Had family rule for TV	P5TVHOME, P5TVRULE	3,470	91.63	0.459	0.471	0.975	0.951
Did homework 3-4 times per week	P5OFTDHW	3,492	41.66	1.148	0.835	1.375	1.892
Have someone help with reading homework	P5HELPR	3,436	98.22	0.418	0.226	1.852	3.430
Completely true child and self have close time	P5WARMCL	3,414	69.17	1.011	0.790	1.279	1.636
Took away privilege when child angry	P5HITPRV	3,408	59.70	1.512	0.840	1.800	3.240
Self-reported in very good health	P5HEALTH	3,410	87.02	0.864	0.576	1.501	2.253
HH received food stamps in last 12 months	P5FSTAMP	3,474	14.44	1.231	0.596	2.064	4.261

See notes at end of table.

Table 9-6. ECLS-K, panel of all five rounds: standard errors and design effects for the full sample using C1_5SC0-C1_5SC40 and C1_5SP0-C1_5SP40, by selected child and parent variables: School years 1998–99, 1999-2000, and 2001–02—Continued

Survey item	Variable name	Number of cases	Estimate	Design SE[1]	SRS SE[2]	DEFT[3]	DEFF[4]
Child characteristics from teacher questionnaire C (percent)							
Enrolled in third grade	T5GRADE	3,065	88.41	1.500	0.578	2.594	6.727
Average in language skills	T5RTLANG	3,031	71.98	1.070	0.816	1.312	1.721
Average in science/social studies	T5RTSCI	3,009	81.23	1.157	0.712	1.626	2.643
Average in math skills	T5RTMTH	3,029	77.43	1.045	0.759	1.376	1.894
Child characteristics (mean)							
Age of child in months	R5AGE	4,022	111.09	0.130	0.071	1.832	3.357
Child's BMI	C5BMI	3,861	18.53	0.089	0.062	1.437	2.066
Child's household size	P5HTOTAL	3,522	4.61	0.039	0.023	1.669	2.786
Number of children <18 in child's HH	P5LESS18	3,522	2.55	0.033	0.019	1.709	2.919
Number of siblings in HH	P5NUMSIB	3,522	1.58	0.032	0.019	1.675	2.807
Number of hours watched TV after dinner	P5TVAFDH	3,467	0.88	0.027	0.014	1.941	3.768
Median						1.622	2.622
Mean						1.635	2.751
Standard deviation						0.317	1.086
Coefficient of variation						0.194	0.395
Minimum						0.975	0.951
Maximum						2.594	6.727

[1]Design SE is the standard error under the ECLS-K sample design. For an explanation of this statistic, see chapter 4, section 4.8.
[2] SRS SE is the standard error assuming simple random sample. For an explanation of this statistic, see chapter 4, section 4.8.
[3] DEFT is the root design effect. For an explanation of this statistic, see chapter 4, section 4.8.
[4] DEFF is the design effect. For an explanation of this statistic, see chapter 4, section 4.8.
SOURCE: U.S. Department of Education, National Center for Education Statistics, Early Childhood Longitudinal Study, Kindergarten Class of 1998–99 kindergarten, first grade, and third grade data collections, school years 1998–99, 1999–2000, and 2001–02.

The median design effect is 3.2 for the spring-first grade/spring-third grade panel, 3.0 for spring-kindergarten/spring-first grade/spring-third grade panel, 2.9 for the panel of children in all four rounds of data collection involving the full sample of children, and 2.6 for the panel of children in all five rounds of data collection that included children sampled for fall-first grade only.

Table 9-7 presents the median design effects for subgroups based on school type, child's sex and race/ethnicity, geographic region, level of urbanicity, and the socioeconomic status scales of the parents. For the panels that include the full sample of children, the median design effect is lowest for Pacific Islanders (hovering around 1.2 or 1.3) and highest for children in Catholic schools (around 3.4 or 3.5). For the panel involving all five rounds, the range of variability of the median design effects is very different from all other panels. The all-five-round panel has a much reduced sample size, as

Table 9-7. ECLS-K panel: median design effects for subgroups, kindergarten through third grade: School years 1998–99, 1999–2000, and 2001–02

Subgroups	Spring-first grade/ spring-third grade		Spring-kindergarten/ spring-first grade/ spring-third grade		Fall-kindergarten/ spring-kindergarten/ spring-first grade/ spring-third grade		All five round panel	
	DEFT[1]	DEFF[2]	DEFT[1]	DEFF[2]	DEFT[1]	DEFF[2]	DEFT[1]	DEFF[2]
All students	1.769	3.169	1.738	3.040	1.694	2.917	1.622	2.622
School type								
Public	1.640	2.801	1.645	2.757	1.632	2.693	1.501	2.242
Private	1.880	3.521	1.876	3.559	1.849	3.444	1.756	3.113
Catholic private	1.846	3.425	1.850	3.447	1.867	3.494	1.670	2.823
Other private	1.704	2.925	1.644	2.733	1.683	2.832	1.644	2.710
Sex								
Male	1.594	2.564	1.598	2.558	1.547	2.390	1.349	1.818
Female	1.606	2.592	1.598	2.563	1.591	2.573	1.544	2.364
Race/ethnicity								
White	1.742	3.003	1.699	2.847	1.657	2.744	1.604	2.599
Black	1.474	2.175	1.451	2.114	1.433	2.081	1.230	1.509
Hispanic	1.378	1.901	1.358	1.848	1.341	1.810	1.311	1.717
Asian	1.501	2.237	1.488	2.201	1.499	2.241	1.582	2.511
Pacific Islander	1.145	1.310	1.138	1.293	1.106	1.216	1.942	3.374
American Indian	1.369	1.861	1.398	1.964	1.282	1.626	1.892	2.743
Other	1.394	1.955	1.421	2.048	1.418	2.044	1.327	1.754
Region								
Northeast	1.716	2.974	1.704	2.920	1.674	2.804	1.541	2.364
Midwest	1.724	3.008	1.792	3.213	1.714	2.950	1.610	2.529
South	1.738	2.981	1.709	2.947	1.637	2.691	1.458	2.125
West	1.621	2.620	1.644	2.719	1.605	2.585	1.621	2.635
Urbanicity								
Central city	1.659	2.770	1.663	2.778	1.645	2.734	1.559	2.354
Urban fringe and large town	1.616	2.598	1.640	2.685	1.597	2.549	1.541	2.330
Small town and rural area	1.764	3.048	1.708	2.918	1.802	3.237	1.720	3.070
Socioeconomic status quintiles								
First	1.465	2.147	1.408	1.979	1.389	1.924	1.219	1.489
Second	1.479	2.266	1.438	2.073	1.387	1.959	1.325	1.749
Third	1.423	2.022	1.429	2.021	1.404	1.971	1.339	1.793
Fourth	1.495	2.234	1.506	2.281	1.457	2.136	1.362	1.848
Fifth	1.512	2.305	1.503	2.263	1.477	2.209	1.317	1.767

NOTE: Each median is based on 49 items.
[1]DEFT is the root design effect. For an explanation of this statistic, see section 4.8.
[2]DEFF is the design effect. For an explanation of this statistic, see section 4.8.
SOURCE: U.S. Department of Education, National Center for Education Statistics, Early Childhood Longitudinal Study, Kindergarten Class of 1998–99 kindergarten, first grade, and third grade data collections, school years 1998–99, 1999–2000, and 2001–02.

it includes the fall-first grade subsample from the full base year sample. For this panel, the median design effects range from 1.5 for children of lowest socioeconomic level to a high of 4.1 for Pacific Islanders and 4.5 for American Indians. In this reduced panel, the samples of Pacific Islanders and American Indians are highly clustered, resulting in the higher design effects.

Standard errors and design effects were not computed for items from the teacher and school administrator questionnaires since there are no teacher or school weights computed for spring-third grade year. Although standard errors and design effects may also be calculated for the teacher and school administrator questionnaires at the child level, they are quite large compared to those typically found for the ECLS-K data. Design effects for teacher and school items are large because the intraclass correlation is 100 percent for children in the same school and very high for children in the same class; children attending the same school have the same school data, and children in the same class have the same teacher data. The correlation is not 100 percent for children in the same class because teacher data include not only items about the teacher and the class but also items about the individual students as completed by their teachers.