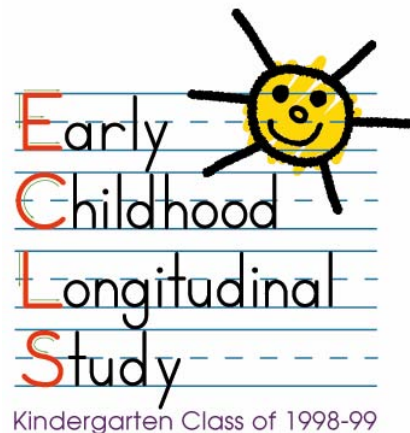

NATIONAL CENTER FOR EDUCATION STATISTICS



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

<u>Chapter</u>		<u>Page</u>
	GETTING STARTED	xxv
1	INTRODUCTION	1-1
	1.1 Background.....	1-4
	1.2 Conceptual Model.....	1-5
	1.3 Study Components	1-6
	1.4 ECLS-K Data Files	1-7
	1.4.1 Differences Between ECLS-K Restricted-Use and Public-Use Files	1-7
	1.4.2 Overview of Available Data Files.....	1-9
	1.5 Contents of Manual.....	1-12
2	DESCRIPTION OF DATA COLLECTION INSTRUMENTS	2-1
	2.1 Direct Child Assessments	2-3
	2.1.1 Socioemotional Development	2-4
	2.1.2 Cognitive Components.....	2-5
	2.1.2.1 Reading.....	2-6
	2.1.2.2 Mathematical Thinking	2-7
	2.1.2.3 Science.....	2-9
	2.1.3 Physical Components	2-10
	2.2 Parent Interview	2-10
	2.3 Teacher Questionnaires.....	2-13
	2.3.1 Academic Rating Scale	2-16
	2.3.2 Teacher Social Rating Scale.....	2-18
	2.4 Special Education Teacher Questionnaires.....	2-19
	2.5 School Administrator Questionnaire.....	2-21
	2.6 School Fact Sheet.....	2-23
	2.7 School Facilities Checklist.....	2-23
	2.8 Student Records Abstract Form	2-23
	REFERENCES	2-25

TABLE OF CONTENTS (continued)

<u>Chapter</u>		<u>Page</u>
3	ASSESSMENT AND RATING SCALE SCORES USED IN THE ECLS-K.....	3-1
	3.1 Direct Cognitive Assessment.....	3-1
	3.1.1 Number-Right Scores.....	3-4
	3.1.2 Item Response Theory (IRT) Scale Scores; Standardized Scores (T-Scores)	3-5
	3.1.3 Item Cluster Scores	3-8
	3.1.3.1 Reading.....	3-9
	3.1.3.2 Science.....	3-10
	3.1.4 Proficiency Levels.....	3-11
	3.1.4.1 Highest Proficiency Level Mastered	3-13
	3.1.4.2 Proficiency Probability Scores	3-15
	3.1.5 Choosing the Appropriate Score for Analysis.....	3-18
	3.1.5.1 Item Response Theory-Based Scores	3-19
	3.1.5.2 Scores Based on Number Right for Subsets of Items (Non-IRT Based Scores)	3-20
	3.1.6 Measuring Gains	3-21
	3.1.7 Reliability.....	3-23
	3.1.8 Validity.....	3-26
	3.2 Indirect Cognitive Assessment	3-28
	3.2.1 Comparison to Direct Cognitive Assessment.....	3-29
	3.2.2 Rasch Scores Available for the Academic Rating Scale.....	3-29
	3.3 Teacher Social Rating Scale	3-36
	3.4 Self-Description Questionnaire.....	3-38
	REFERENCES	3-40

TABLE OF CONTENTS (continued)

<u>Chapter</u>		<u>Page</u>
4	SAMPLE DESIGN AND IMPLEMENTATION	4-1
	4.1 Base Year Sample	4-1
	4.2 Fall-First Grade Subsample	4-4
	4.3 Spring-First Grade Sample	4-7
	4.3.1 Subsampling Movers	4-7
	4.3.2 Student Freshening	4-11
	4.4 Spring-Third Grade Sample	4-12
	4.4.1 Subsampling Movers	4-12
	4.4.2 Language Minority Children	4-13
	4.5 Sample Attrition	4-13
	4.6 Calculation and Use of Sample Weights	4-17
	4.6.1 Types of Sample Weights	4-18
	4.6.2 Weighting Procedures	4-21
	4.6.3 Computation of Spring-First Grade Initial Child Weights ..	4-22
	4.6.3.1 Base Year Nonresponse-Adjusted School Weights	4-22
	4.6.3.2 Base Year Child Weights	4-23
	4.6.3.2.1 Base Year Child Weights for Base Year Respondents	4-23
	4.6.3.2.2 Base Year Child Weights for Eligible Children Sampled in First Grade	4-24
	4.6.4 Computation of Spring-Third Grade Child Weights	4-27
	4.6.4.1 Adjustment for Movers	4-27
	4.6.4.2 Adjustment for Nonresponse	4-28
	4.6.4.3 Raking to Sample-Based Control Totals	4-28

TABLE OF CONTENTS (continued)

<u>Chapter</u>		<u>Page</u>
	4.6.5	Types of Weights and Their Use..... 4-29
	4.6.5.1	Weights To Be Used With Direct Child Assessment Data (C5CW0)..... 4-29
	4.6.5.2	Weights To Be Used With Parent Data (C5PW0)..... 4-30
	4.6.5.3	Weights To Be Used With a Combination of Child Direct Assessment Data and Parent Interview Data and Teacher Data (C5CPTW0)... 4-31
	4.6.6	Replicate Weights 4-31
	4.6.7	Characteristics of Sample Weights 4-31
	4.7	Variance Estimation..... 4-32
	4.7.1	Paired Jackknife Replication Method 4-32
	4.7.2	Taylor Series Method..... 4-34
	4.7.3	Specifications for Computing Standard Errors 4-35
	4.8	Design Effects..... 4-37
	4.8.1	Use of Design Effects..... 4-38
	4.8.2	Average Design Effects for the ECLS-K 4-39
		REFERENCES 4-44
5		DATA COLLECTION METHODS AND RESPONSE RATES..... 5-1
	5.1	Overview of Data Collection Methods 5-1
	5.2	Field Staff Training..... 5-1
	5.2.1	Advance Contact and Recruitment Training 5-2
	5.2.2	Spring-Third Grade Training 5-3
	5.2.2.1	Parent Interviewer-Only Training 5-3
	5.2.2.2	Field Supervisor Training..... 5-4
	5.2.2.3	Assessor Training 5-4
	5.2.2.4	Certification of the Child Assessors 5-5

TABLE OF CONTENTS (continued)

<u>Chapter</u>		<u>Page</u>
5.3	Fall Preassessment School Contact.....	5-6
5.3.1	Advance Mailings	5-7
5.3.2	Preassessment Contact	5-7
	5.3.2.1 Identifying ECLS-K Sampled Children Who Withdrew From the School	5-8
	5.3.2.2 Reviewing Information About ECLS-K Sampled Children	5-8
	5.3.2.3 Reviewing Parent Consent	5-8
	5.3.2.4 Contacting Families of Home-Schooled Children	5-9
5.3.3	Preparing for Spring-Third Grade Data Collection.....	5-9
5.4	Spring-Third Grade Data Collection.....	5-9
5.4.1	Preassessment School Contact	5-10
5.4.2	Timeline of the Direct Child Assessments.....	5-11
	5.4.2.1 Conducting the Direct Child Assessments	5-11
	5.4.2.2 Accommodations and Exclusions.....	5-13
5.4.3	Conducting the Parent Interview.....	5-13
5.4.4	Conducting Data Collection on Children Who Withdrew From Their Previous Round School.....	5-14
5.4.5	Teacher and School Data Collection.....	5-17
5.5	Data Collection Quality Control	5-18
5.5.1	Child Assessments Observations	5-19
5.5.2	Validation of Parent Interviews.....	5-20
5.5.3	Validations of School Visits.....	5-20
5.5.4	Assessor Interrater Reliability	5-21
5.6	Spring-Third Grade Completion Rates	5-23
5.6.1	Students Sampled in Kindergarten.....	5-23
5.6.2	Students Sampled in First Grade.....	5-39
5.6.3	Spring-Third Grade Completion Rates—All Children	5-41
5.6.4	Spring-Third Grade Completion Rates Conditioned on Child Assessment	5-43
	REFERENCES	5-45

TABLE OF CONTENTS (continued)

<u>Chapter</u>		<u>Page</u>
6	DATA PREPARATION	6-1
6.1	Coding and Editing Specifications for Computer-Assisted Interviews (CAI)	6-1
6.1.1	Range Specifications	6-2
6.1.2	Consistency Checks (Logical Edits)	6-2
6.1.3	Coding	6-2
	6.1.3.1 Review of “Other, specify” Items.....	6-3
	6.1.3.2 Parent Occupation Coding.....	6-3
	6.1.3.3 Race/Ethnicity Coding.....	6-4
	6.1.3.4 Partially Complete Parent Initiatives	6-5
	6.1.3.5 Household Roster in the Parent Interview	6-5
6.2	Coding and Editing Specifications for Hard-Copy Questionnaires	6-6
6.2.1	Receipt Control	6-6
6.2.2	Coding	6-7
	6.2.2.1 Review of “Other, specify” Items.....	6-7
	6.2.2.2 Coding Teacher/Race Ethnicity.....	6-7
	6.2.2.3 Coding Teacher Language.....	6-8
6.2.3	Data Entry	6-8
6.2.4	Data Editing	6-8
	6.2.4.1 Range Specifications	6-8
	6.2.4.2 Consistency Checks (Logical Edits).....	6-9
	6.2.4.3 Frequency and Cross-Tabulation Review.....	6-9
	REFERENCES	6-10
7	DATA FILE CONTENT AND COMPOSITE VARIABLES.....	7-1
7.1	Identification Variables.....	7-2
7.2	Missing Values	7-3
7.3	Variable Naming Conventions.....	7-6
7.4	Composite Variables.....	7-6

TABLE OF CONTENTS (continued)

<u>Chapter</u>		<u>Page</u>
7.4.1	Child Composite Variables	7-8
7.4.1.1	Child’s Age at Assessment (R5AGE)	7-8
7.4.1.2	Gender (R5GENDER).....	7-8
7.4.1.3	Child’s Date of Birth (R5DOBY, R5DOBMM, and R5DOBDD).....	7-9
7.4.1.4	Race/Ethnicity (W3AMERIN, W3ASIAN, W3PACISL, W3BLACK, W3WHITE, W3HISP, W3MT1RAC, W3RACETH, and R5RACE).....	7-10
7.4.1.5	Child’s Height (C5HEIGHT)	7-11
7.4.1.6	Child’s Weight (C5WEIGHT).....	7-11
7.4.1.7	Child’s Body Mass Index (C5BMI)	7-12
7.4.1.8	Child’s Disability Status (P5DISABL).....	7-12
7.4.1.9	Nonparental Care (P5CARNOW)	7-13
7.4.1.10	Hours Per Week in Child Care (P5HRSNOW) ...	7-13
7.4.1.11	Number of Child Care Arrangements (P5NUMNOW)	7-15
7.4.1.12	Primary Nonparental Child Care Arrangement (P5PRIMNW).....	7-15
7.4.2	Family and Household Composite Variables.....	7-16
7.4.2.1	Number of Siblings (P5NUMSIB)	7-16
7.4.2.2	Parent and Household Members’ Age (P5LESS18, P5OVER18, P5HDAGE, and P5HMAGE).....	7-17
7.4.2.3	Food Security Status.....	7-18
7.4.2.4	Food Security Status: Continuous Measures (P5FSSCAL and P5FSCHSC).....	7-19
7.4.2.5	Food Security Status: Categorical Measures (P5FSSTAT and P5FSCHST)	7-19
7.4.2.6	Food Security Status: Raw Scores (P5FSRAW and P5FSCHRA).....	7-20
7.4.2.7	Socioeconomic Status (SES) and Poverty (W3DADSCR, W3MOMSCR, W3SESL, W3SESQ5, W3INCCAT, W3POVRTY).....	7-20
7.4.2.8	Parent Education (W3PARED, W3DADED, and W3MOMED).....	7-27
7.4.2.9	Parent Race/Ethnicity (P5HDRACE and P5HMRACE)	7-28

TABLE OF CONTENTS (continued)

<u>Chapter</u>		<u>Page</u>
7.4.3	Teacher Composite Variables	7-29
7.4.3.1	Grade-Level Composite (T5GLVL).....	7-29
7.4.3.2	Class Size (A5CLSZ)	7-30
7.4.4	School and Class Composite Variables.....	7-31
7.4.4.1	School Type (S5SCTYP)	7-31
7.4.4.2	Public or Private School (S5PUPRI).....	7-32
7.4.4.3	School and Grade-Level Enrollment (S5ENRLS, S5ENRLT)	7-32
7.4.4.4	Percent Minority Students in the School (S5MINOR).....	7-33
7.4.4.5	School Instructional Level (S5SCLVL).....	7-34
7.4.4.6	School Year Start and End Dates (L5SCHBDD, L5SCHBMM, L5SCHBY Y, L5SCHEDD, L5SCHEMM, L5SCHEYY).....	7-35
7.4.5	Student Records Abstract and Field Management System Composite Variables	7-35
7.4.5.1	Year-Round Schools (F5YRRND).....	7-36
7.4.5.2	Indicator of Whether Child Received Special Education Services (F5SPECS).....	7-36
7.4.5.3	Indicator of Whether Child Has an Individualized Education Plan (IEP) on Record at School (U5RIEP)	7-36
7.4.6	Parent Identifiers and Household Composition (P5DADID, P5MOMID, P5HPARNT, P5HDAD, P5HMOM, P5HFAMIL, P5MOMTYP, P5DADTYP).....	7-36
7.4.7	Industry and Occupation Codes Used in the ECLS-K	7-39
7.5	Methodological Variables	7-43
7.6	Children Who Changed Schools.....	7-44
7.6.1	Children Who Changed Schools During Third Grade Data Collection.....	7-44
7.6.2	Children Who Changed Schools Between Rounds (R5DEST, R5R4SCHG)	7-47
7.7	Composite Table	7-48
7.8	Masked Variables	7-49

TABLE OF CONTENTS (continued)

<u>Chapter</u>		<u>Page</u>
	REFERENCES	7-80
8	ELECTRONIC CODE BOOK	8-1
	8.1 Introduction.....	8-1
	8.1.1 Hardware/Software Requirements	8-2
	8.1.2 ECB Features	8-2
	8.2 Installing, Starting, and Exiting the ECB.....	8-3
	8.2.1 Installing the ECB Program on Your Personal Computer ..	8-3
	8.2.2 Starting the ECB	8-7
	8.2.3 Exiting the ECB	8-10
	8.2.4 Removing the ECB Program From Your Personal Computer.....	8-11
	8.2.5 Title Bar	8-12
	8.2.6 Menu Bar.....	8-12
	8.2.7 Using Shortcut Keys to Navigate.....	8-13
	8.3 Variable List	8-13
	8.3.1 Searching the Code Book for Variables	8-16
	8.3.1.1 Using the Go Button.....	8-16
	8.3.1.2 Narrowing Your Variable Search.....	8-17
	8.3.1.3 Expanding Your Variable Search.....	8-20
	8.3.1.4 Resetting Your Variable List.....	8-22
	8.4 Working Taglist	8-22
	8.4.1 Opening a Taglist.....	8-23
	8.4.2 Adding Variables to the Working Taglist	8-24
	8.4.3 Removing Variables From the Working Taglist	8-26
	8.4.4 Saving Taglists.....	8-27
	8.4.5 Exporting Taglists	8-29
	8.4.6 Importing Taglists	8-30
	8.4.7 Using Predefined Taglists	8-31
	8.4.8 Deleting Taglists	8-32
	8.4.9 Viewing Code Book and Variable Information	8-33

TABLE OF CONTENTS (continued)

<u>Chapter</u>		<u>Page</u>
8.5	Extracting Data from the ECB	8-41
	8.5.1 Reviewing the Extract Specifications	8-48
	8.5.2 Repairing and Compacting the Database	8-49
8.6	Menu Bar Descriptions	8-50
8.7	Child Catalog	8-51
	8.7.1 Third Grade Child Catalog Topical Variable Groupings	8-55
	8.7.2 Child Catalog Predefined Taglists.....	8-61
	8.7.3 Child Catalog Limiting Fields.....	8-61
9	CREATING A LONGITUDINAL FILE	9-1
	9.1 Conducting Longitudinal Analyses.....	9-1
	9.2 Examples of Research Questions.....	9-2
	9.3 Merging Base Year Child-Level Data with the First Grade and Third Grade Child-Level Data	9-3
	9.4 K-3 Longitudinal Weights.....	9-4
	9.4.1 Type of K-3 Longitudinal Weights	9-4
	9.4.2 Weighting Procedures	9-7
	9.4.2.1 Longitudinal Weights Not Involving the Fall-First Grade Data	9-8
	9.4.2.2 Longitudinal Weights Involving the Fall-First Grade Data	9-9
	9.4.3 Characteristics of Longitudinal Weights.....	9-10
	9.4.4 Variance Estimation	9-11
	9.4.5 Design Effects	9-14

List of Tables

<u>Table</u>		
3-1	Direct cognitive assessment: Routing test number-right, kindergarten/ first grade (K-1) assessments: School years 1998-99 and 1999-2000.....	3-5
3-2	Direct cognitive assessment: Routing test number-right, third grade assessment: School year 2001-02.....	3-5

TABLE OF CONTENTS (continued)

List of Tables (continued)

<u>Table</u>		<u>Page</u>
3-3	Direct cognitive assessment: Item Response Theory scale scores: School year 2001–02	3-7
3-4	Direct cognitive assessment: standardized scores: School year 2001–02.....	3-8
3-5	Direct cognitive assessment: Reading cluster scores: School year 2001–02...	3-10
3-6	Direct cognitive assessment: Science cluster scores: School year 2001–02....	3-11
3-7	Third grade direct cognitive assessment: highest proficiency level mastered, in percent: School year 2001–02	3-14
3-8	Third grade direct cognitive assessment: proficiency probability scores— reading: School year 2001–02	3-16
3-9	Third grade direct cognitive assessment: proficiency probability scores— mathematics: School year 2001–02	3-17
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.....	3-25
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.....	3-25
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	3-26
3-13	Percent agreement of highest proficiency level-mastered score, by round of data collection: School years 1998–99, 1999–2000, and 2001–02 ...	3-26
3-14	Validity coefficients for reading and mathematics field test item pools: School year 2001–02	3-28
3-15	Academic Rating Scale response scale: School year 2001–02	3-28
3-16	Person separation reliability statistics for the third-grade Rasch-based score, by category: School year 2001–02.....	3-30

TABLE OF CONTENTS (continued)

List of Tables (continued)

<u>Table</u>		<u>Page</u>
3-17	Third grade Academic Rating Scale: variable names, descriptions, ranges, weighted means, and standard deviations: School year 2001–02.....	3-31
3-18	Spring-third grade Academic Rating Scale language and literacy item difficulties: School year 2001–02	3-31
3-19	Spring-third grade Academic Rating Scale mathematical thinking item difficulties: School year 2001–02	3-32
3-20	Spring-third grade Academic Rating Scale science item difficulties: School year 2001–02	3-32
3-21	Spring-third grade Academic Rating Scale social studies item difficulties: School year 2001–02	3-33
3-22	Spring-third grade Academic Rating Scale language and literacy standard errors: School year 2001–02.....	3-34
3-23	Spring-third grade Academic Rating Scale mathematical thinking standard errors: School year 2001–02.....	3-34
3-24	Spring-third grade Academic Rating Scale science standard errors: School year 2001–02	3-35
3-25	Spring-third grade Academic Rating Scale social studies standard errors: School year 2001–02	3-35
3-26	Social Rating Scale response scale: School year 2001–02	3-36
3-27	Teacher Social Rating Scale scores: variable names, descriptions, ranges, weighted means, and standard deviations: School year 2001–02.....	3-37
3-28	Split-half reliability for the teacher Social Rating Scale scores: School year 2001–02	3-38
3-29	Self-Description Questionnaire scale reliabilities (alpha coefficient): School year 2001–02	3-39
3-30	Self-Description Questionnaire scale range, mean, and standard deviation (weighted): School year 2001–02	3-39

TABLE OF CONTENTS (continued)

List of Tables (continued)

<u>Table</u>		<u>Page</u>
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.....	4-2
4-2	Number of schools in the ECLS-K base year school sample, by selected school characteristics: School year 1998–99	4-3
4-3	Number (unweighted) of children in the ECLS-K base year student sample, by selected characteristics: School year 1998–99	4-5
4-4	Number of base year cooperating schools selected for fall-first grade, by selected school characteristics: School year 1999–2000	4-6
4-5	Number (unweighted) of children subsampled for fall-first grade, by selected characteristics: School year 1999–2000	4-8
4-6	Number (unweighted) of children in spring-first grade sample excluding freshened students, by selected characteristics: School year 1999–2000	4-10
4-7	Number (unweighted) of children in spring-third grade sample excluding freshened students, by selected characteristics: School year 2001–02	4-14
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.....	4-16
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.....	4-16
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.....	4-17
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	4-21

TABLE OF CONTENTS (continued)

List of Tables (continued)

<u>Table</u>		<u>Page</u>
4-12	Number of children who were not assessed in spring-third grade, by special situations: School year 2001–02	4-30
4-13	Characteristics of the first grade child-level weights: School year 1999–2000	4-32
4-14	ECLS-K Taylor Series stratum and first-stage unit identifiers	4-35
4-15	Specifications for computing standard errors, spring-third grade: School year 2001–02	4-36
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	4-41
4-17	ECLS-K median design effects for subgroups—child, parent, and teacher questionnaire part C data: School year 2001–02	4-43
5-1	Number and percent of trainees, by scores on training certification form: School year 2001–02	5-6
5-2	Results of the Telephone Research Center’s locating efforts, spring–third grade: School year 2001–02	5-10
5-3	Completed child assessments by round of data collection and selected characteristics: School years 1998–99, 1999–2000, and 2001–02	5-12
5-4	Number of children excluded from and requiring an accommodation in the spring-third grade assessments: School year 2001–02.....	5-13
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.....	5-15
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	5-16
5-7	Results of the child assessments observations, spring-third grade: School year 2001-02	5-20

TABLE OF CONTENTS (continued)

List of Tables (continued)

<u>Table</u>		<u>Page</u>
5-8	Interrater reliability on child assessment validation item, by subject area and level: School year 2001–02	5-22
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	5-24
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	5-26
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	5-28
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.....	5-32
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.....	5-33
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.....	5-34
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	5-36
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	5-37
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	5-38

TABLE OF CONTENTS (continued)

List of Tables (continued)

<u>Table</u>		<u>Page</u>
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	5-39
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.....	5-40
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	5-42
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	5-44
7-1	Incremented ages of previous household members based on round household member entered study, spring-third grade: School year 2001–02	7-18
7-2	Levels of detailed income range, spring-third grade: School year 2001–02	7-21
7-3	Households asked to report income to the nearest \$1,000, spring-third grade: School year 2001–02.....	7-22
7-4	Missing data for SES source variables, spring-third grade: School year 2001–02	7-22
7-5	Selected statistics on imputed education variables, spring-third grade: School year 2001–02	7-24
7-6	Selected statistics on imputed labor force status, spring-third grade: School year 2001–02	7-25
7-7	Selected statistics on imputed occupation variables, spring-third grade: School year 2001–02	7-25
7-8	Selected statistics on imputed detailed income range, spring-third grade: School year 2001–02	7-25

TABLE OF CONTENTS (continued)

List of Tables (continued)

<u>Table</u>		<u>Page</u>
7-9	ECLS-K spring-third grade and Census poverty thresholds for 2001: School year 2001–02	7-27
7-10	Pointers to parent figure questions, spring-third grade: School year 2001–02	7-40
7-11	Case status and school ID numbers for children not followed or located, spring-third grade: School year 2001–02.....	7-45
7-12	Spring-third grade composite variables: School year 2001–02	7-50
7-13	Recoded and suppressed data on the ECLS-K Third-Grade Public-Use Data Fill: School year 2001–02	7-74
9-1	Characteristics of child-level K–3 longitudinal weights, spring-third grade: School year 2001–02.....	9-10
9-2	Specifications for computing standard errors, spring third-grade: School year 2001–02.....	9-13
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 ...	9-15
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	9-17
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.....	9-19
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.....	9-21
9-7	ECLS-K panel: median design effects for subgroups, kindergarten through third grade: School years 1998–99, 1999–2000, and 2001–02	9-23

TABLE OF CONTENTS (continued)

List of Exhibits

<u>Exhibit</u>		<u>Page</u>
1-1	ECLS-K waves of data collection: Years 1998–2004	1-2
1-2	ECLS-K conceptual model	1-6
2-1	Instruments used in the ECLS-K, by round of data collection: School years 1998–99, 1999–2000, and 2001–02	2-2
2-2	ECLS-K direct child assessments, by domain and round of data collection: School years 1998–99, 1999–2000, and 2001–02	2-3
2-3	ECLS-K parent interview, by major content topics and round of data collection: School years 1998–99, 1999–2000, and 2001–02.....	2-11
2-4	Teacher questionnaires, by major content topics and round of data collection: School years 1998–99, 1999–2000, and 2001–02	2-14
2-5	Special education teacher questionnaires, by major content topics and round of data collection: School years 1998–99, 1999–2000, and 2001–02 ...	2-20
2-6	School administrator questionnaire, by major content topics and round of data collection: School years 1998–99, 1999–2000, and 2001–02.....	2-22
3-1	Reading proficiency levels, kindergarten through third grade: School years 1998–99, 1999–2000, and 2000–02	3-12
3-2	Mathematics proficiency levels, kindergarten through third grade: School years 1998–99, 1999–2000, and 2000–02	3-12
4-1	ECLS-K third grade cross-sectional weights: School year 2001–02.....	4-19
7-1	Missing values codes, ECLS-K data: School years 1998–99, 1999–2000, and 2001–02.....	7-3
7-2	Prefixes for ECLS-K third grade variables and cross-sectional and cross-round panel weights: School year 2001–02.....	7-7
8-1	Windows Run screen	8-4
8-2	InstallShield Wizard	8-4
8-3	Welcome window	8-5

TABLE OF CONTENTS (continued)

List of Exhibits (continued)

<u>Exhibit</u>		<u>Page</u>
8-4	Choose Destination Location.....	8-5
8-5	Setup Status.....	8-6
8-6	InstallShield Wizard Complete.....	8-6
8-7a	Desktop icon.....	8-7
8-7b	Desktop screen—click start.....	8-8
8-8	First-time user dialog box.....	8-8
8-9	ECB splash screen.....	8-9
8-10	Select Catalog screen.....	8-9
8-11	Main ECB screen.....	8-10
8-12	Exit screen.....	8-11
8-13	Save working taglist dialog box.....	8-11
8-14	Title Bar.....	8-12
8-15	Menu Bar.....	8-12
8-16	ECB main screen.....	8-14
8-17	Variable List.....	8-15
8-18	Go button.....	8-17
8-19	Narrow Search Text dialog box.....	8-18
8-20	No Matches Found message.....	8-19
8-21	Example of narrowing a search.....	8-20

TABLE OF CONTENTS (continued)

List of Exhibits (continued)

<u>Exhibit</u>		<u>Page</u>
8-22	Expand Search Text dialog box	8-21
8-23	No Matches Found message	8-22
8-24	ECB Working Taglist	8-23
8-25	Open Taglist dialog box.....	8-24
8-26	Add variables buttons	8-25
8-27	Add Taglist dialog box	8-26
8-28	Remove variables buttons	8-26
8-29	Save Taglist As dialog box	8-28
8-30	Save Taglist As dialog box (#2).....	8-29
8-31	Pulldown menu to select Taglist Export	8-30
8-32	Export Taglist dialog box	8-30
8-33	Pulldown menu to select Taglist Import	8-31
8-34	Import Taglist dialog box	8-32
8-35	Delete Taglist selection.....	8-33
8-36	Delete Taglist confirmation window	8-33
8-37	Code book information	8-34
8-38	Code book view	8-35
8-39	Navigation buttons.....	8-35
8-40	Printing status screen	8-36

TABLE OF CONTENTS (continued)

List of Exhibits (continued)

<u>Exhibit</u>		<u>Page</u>
8-41	Export code book selection screen.....	8-37
8-42	Variable Quick View	8-38
8-43	Add All Variables dialog box	8-39
8-44	View of the entire code book	8-40
8-45	Printer dialog box.....	8-40
8-46	Limiting fields dialog box.....	8-42
8-47	Save SAS program file dialog box.....	8-43
8-48	Save SAS program file location browse screen	8-44
8-49	Save SAS data file dialog box	8-44
8-50	Save SPSS program file dialog box	8-45
8-51	Save SPSS program file location browse screen	8-46
8-52	Save SPSS data file dialog box	8-46
8-53	Save Stata program file dialog box	8-47
8-54	Save Stata program file location browse screen	8-48
8-55	Save Stata data file dialog box.....	8-48
8-56	Repair database completed screen	8-49
8-57	Menu Bar Descriptions	8-50
8-58	Child catalog required variables	8-51
8-59	Child catalog blocks.....	8-53

TABLE OF CONTENTS (continued)

List of Exhibits (continued)

<u>Exhibit</u>		<u>Page</u>
8-60	Selected child catalog topical variable groupings.....	8-56
8-61	Third grade predefined taglist—child catalog	8-62
8-62	Child catalog Extract Specifications window	8-63
8-63	Third Grade Public-Use ECB main screen	8-64
9-1	ECLS-K: K–3 longitudinal weights, spring-third grade: School year 2001–02.....	9-7
9-2	ECLS-K Taylor Series stratum and first-stage unit identifiers, spring-third grade: School year 2001–02.....	9-12

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, SPSS 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 data¹ 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 schools³ (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

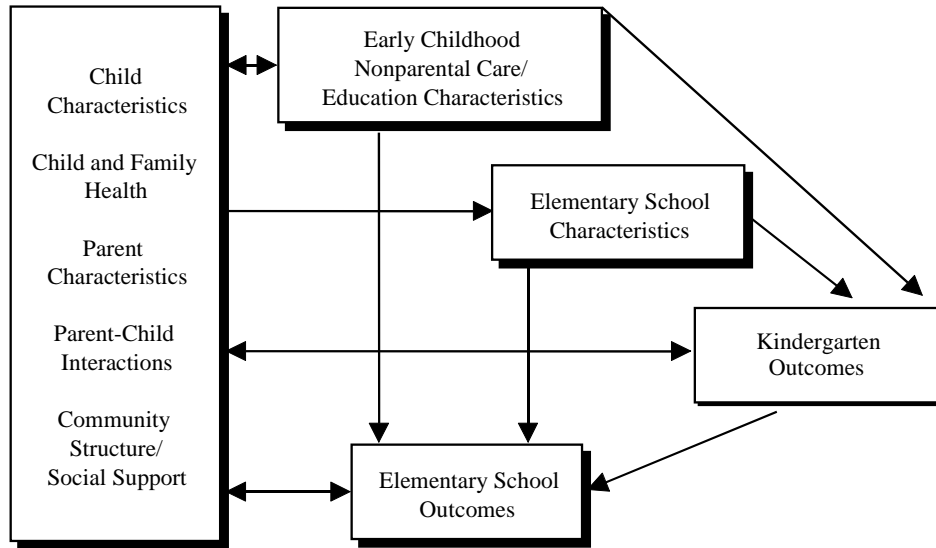
enables researchers to study children's cognitive, social, and emotional growth and to relate trajectories of change to variations in children's experiences in kindergarten and the early grades.

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 outlier 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 than 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 regroupes 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- kindergarten	Spring-first grade 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 statistics¹ 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

¹Person separation reliability is a measure of internal consistency and is analogous to the KR-20 and Cronbach’s alpha. Person separation is the ratio of the adjusted standard deviation to the root mean standard error: $((S. D. of Measure)^2 - (RMSE)^2) / RMSE$. Person separation reliability is the square of this separation statistic divided by one plus the separation squared (Linacre and Wright, 2000). SOURCE: U.S. Department of Education, National Center for Education Statistics, Early Childhood Longitudinal Study, Kindergarten Class of 1998–99 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

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

SOURCE: U.S. Department of Education, National Center for Education Statistics, Early Childhood Longitudinal Study, Kindergarten Class of 1998–99 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	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 nonrespondents 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
Midsized city	1,109	874	235
Urban fringe of large city	1,558	1,205	353
Urban fringe of midsized 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	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	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. <i>Exception:</i> 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 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 freshening³ 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
- $\frac{\text{The sum of the mover adjusted weights of all children (any group) over the sum of the mover adjusted weights of children who were eligible respondents, eligible nonrespondents or ineligible (group a, b or c), if the child was not of unknown eligibility.}}{\text{The sum of the mover adjusted weights of all children (any group) over the sum of the mover adjusted weights of children who were eligible respondents, eligible nonrespondents or ineligible (group a, b or c), if the child was not of unknown eligibility.}}$

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

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

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	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 ($\times 100$)	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:

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

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

Type of analysis	Full sample weight	Computing standard errors					Approximating
		Replication method			Taylor Series method		sampling errors
		(WesVar, SUDAAN or AM)			(SUDAAN, STATA, SAS or AM)		DEFT
		ID	Replicate weights	Jackknife method	Sample design	Nesting variables	(Average root design effect)
Spring-third grade cross-sectional							1.841
	C5CW0	CHILDID	C5CW1 – C5CW90	JK2	WR ¹	C5TCWSTR C5TCWPSU	
	C5PW0	PARENTID	C5PW1 – C5PW90	JK2	WR ¹	C5TPWSTR C5TPWPSU	
	C5CPTW0	CHILDID	C5CPTW1 – C5CPTW90	JK2	WR ¹	C5CPTSTR C5CPTPSU	

¹WR = 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:
$$Var_{SRS} = \frac{1}{n} \frac{\sum_1^n w_i (x_i - \bar{x}_w)^2}{\sum_1^n w_i} = SE_{SRS}^2,$$

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	10,671	41.57	0.625	0.477	1.311	1.718
	P5COMPWK						
Had internet access	P5HOMECM	10,417	85.61	0.466	0.344	1.356	1.840
	P5INTACC						
Used computer 1-2 times per week for homework	P5HOMECM	10,394	52.58	0.593	0.490	1.210	1.465
	P5CMPEDU						
Had family rule for TV	P5TVHOME	13,285	91.58	0.323	0.241	1.342	1.802
	P5TVRULE						
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

¹Design SE is the standard error under the ECLS-K sample design. For an explanation of this statistic, see section 4.8.

²SRS SE is the standard error assuming simple random sample. For an explanation of this statistic, see section 4.8.

³DEFT is the root design effect. For an explanation of this statistic, see section 4.8.

⁴DEFF 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

¹DEFT is the root design effect. For an explanation of this statistic, see section 4.8.

²DEFF 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

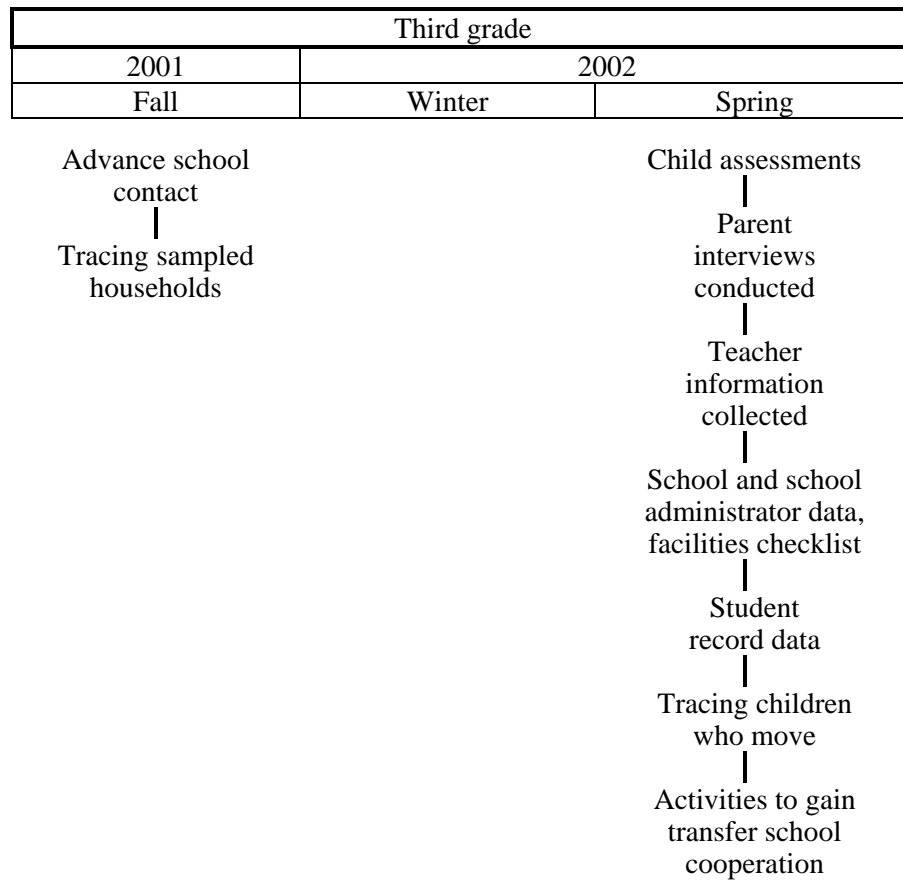
The ECLS-K third grade data collection was conducted in the fall and spring of the 2001–02 school year. Fall data collection included contacting sampled schools to set appointments to conduct the child assessments in the spring of the school year, verifying the parent consent procedures, linking children 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



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 accommodations¹ 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.

¹The 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

¹The 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		
	Completes ²	Completion rate		Completes ³	Completion rate	
		Weighted	Unweighted		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		
	Completes ²	Completion rate		Completes ³	Completion rate	
		Weighted	Unweighted		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		
	Completes ²	Completion rate		Completes ²	Completion rate	
		Weighted	Unweighted		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		
	Completes ²	Completion rate		Completes ²	Completion rate	
		Weighted	Unweighted		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		
	Completes ²	Completion rate		Completes ²	Completion rate		Completes ²	Completion rate	
		Weighted	Unweighted		Weighted	Unweighted		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		
	Completes ²	Completion rate		Completes ³	Completion rate	
		Weighted	Unweighted		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		
	Completes ²	Completion rate		Completes ²	Completion rate		Completes ²	Completion rate	
		Weighted	Unweighted		Weighted	Unweighted		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		
	Completes ²	Completion rate		Completes ³	Completion rate	
		Weighted	Unweighted		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		
	Completes ²	Completion rate		Completes ²	Completion rate	
		Weighted	Unweighted		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		
	Completes ²	Completion rate		Completes ²	Completion rate		Completes ²	Completion rate	
		Weighted	Unweighted		Weighted	Unweighted		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		
	Completes	Completion rate		Completes	Completion rate		Completes	Completion rate	
		Weighted	Unweighted		Weighted	Unweighted		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		
	Completes	Completion rate		Completes	Completion rate		Completes	Completion rate	
		Weighted	Unweighted		Weighted	Unweighted		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 Commerce, 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 CHILDDID, 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 DOBY) 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		Number of cases resolved
			Broad income range	Known	
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 missing⁶ 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, S5HISPCT, 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 S5ADAFGL (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, CHILDDGEN. 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 CHILDDGEN 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, R5DOBY (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	CHILGEN (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
R5DOBY	Child	Child’s date of birth year	DOBY, 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 (CHQ.520), 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 (CCQ.403), 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	R5DOBY	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	L5SCHBY	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	A57YRSL	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 DISABILITIES	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	T5CHRDP	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	S5ASNPC	S5 Q3A PERCENT OF ASIAN STUDENTS	These data suppressed for respondent confidentiality.
1855	S5HSPCT	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.

8. ELECTRONIC CODE BOOK

8.1 Introduction

The purpose of this chapter is to provide users of the Early Childhood Longitudinal Study, both Kindergarten Class of 1998–99 (ECLS-K) and Birth Cohort (ECLS-B), with specific directions for using the Electronic Code Book (ECB) CD-ROM. The information in this chapter provides a comprehensive tour through the ECB that addresses all of the functions and capabilities of the program. These functions allow users to access the accompanying catalog and “view” the data in various ways by performing customized searches, queries, and extractions. The organization of this document provides a “start to finish” approach through the system, beginning with the installation of the ECB, utilizing the ECB’s functions, navigating through the catalog, and performing user-specified data extractions.

Sections 8.1 through 8.6 contain general instructions on using the ECB and apply to both the ECLS-K ECB and the ECLS-B ECB, including descriptions of the menu bars (exhibit 8-57). The exhibits and examples given in these sections are generic and will not exactly match what the users see on their own screens. The last section, 8.7, is file-specific and reflects the actual data on the file.

The ECB CD-ROM contains an ECB that allows users to easily examine the variables in the ECB data set. The data user can create SAS, SPSS for Windows, and Stata programs that will generate an extract data file from the text (ASCII) data file on the CD-ROM. For more information about the data file, see section 8.7.

Additionally, the CD-ROM contains Portable Document Format (PDF) files of the associated questionnaires in appendix A and the record layout for the data file in appendix B. When needed, additional user’s guides and additional supplementary files may also be included in additional appendices.

8.1.1 Hardware/Software Requirements

The ECB program is designed to run under Windows 95[®], Windows 98[®], Windows 2000[®], Windows XP[®], or Windows NT[®] 4.0 on a Pentium-class or higher PC. The PC should also have a minimum of 20 megabytes (MB) of available disk space. The program will visually fit best on screens set to a desktop area of 800 x 600 pixels. It will still work on other screen settings, but it may not make the best use of the available screen space. You can check/set your desktop area as follows:

1. Click on the Windows Start button.
2. Select the Settings menu and then the Control Panel folder icon.
3. In the Control Panel window, click on the Display icon.
4. Select the Settings tab.
5. Set the Desktop Area to 800 x 600 pixels with the Desktop Area slidebar.

The ECB requires approximately 20 MB of available disk space on your hard drive. If 20 MB of space is not available, you may wish to delete unnecessary files from the drive to make space for the ECB.

8.1.2 ECB Features

The ECB allows a user to do the following:

- Search the names and labels of variables in the database (called the catalog) to select variables for analysis (see section 8.3, Variable List).
- Examine the question wording, response categories, and response frequencies for variables the user selects (see section 8.4.9, Viewing Code Book and Variable Information).
- Create a list of variables to be extracted from the catalog, save the list for later use, print the list as a code book, or use a predefined list on the ECB (see section 8.4, Working Taglist).
- Automatically generate SAS, SPSS for Windows, or Stata programs to extract selected variables from the whole data set or for a subset of the cases that are defined by the user (see section 8.5, Extracting Data from the ECB).

The ECB does not create a SAS, SPSS for Windows, or Stata data file. It will prepare the statements that you can use with your own SAS, SPSS for Windows, or Stata software to create your file. As noted earlier, the CD-ROM contains an ASCII data set that the ECB uses to extract specific subdata files. The CD-ROM must be in the drive for the data to be extracted.

8.2 Installing, Starting, and Exiting the ECB

The ECB is provided on a CD-ROM and is intended to be installed and run from within the Windows 95 (or any later version of Windows) environment. The sections in this chapter provide you with step-by-step instructions for installing the program on your personal computer (PC), starting the program, and exiting the program once you have completed your tasks.

If you installed the ECB on your PC before, you have to uninstall the old program before installing this version. For details on how to remove the ECB program from your PC, please see section 8.2.4.

8.2.1 Installing the ECB Program on Your Personal Computer

Program installation is initiated by running the Setup.exe file found within the CD-ROM's root directory.

How To Install the Program:

1. Close all applications on your computer.
2. Insert the installation CD-ROM into your PC's CD-ROM drive.
3. From the desktop Start menu, select Run.
4. Type "D:\Setup.exe" into the "Open" field of the Run screen, shown in exhibit 8-1. If your CD-ROM drive is assigned a different drive letter, substitute it for the "D."

Exhibit 8-1. Windows Run screen



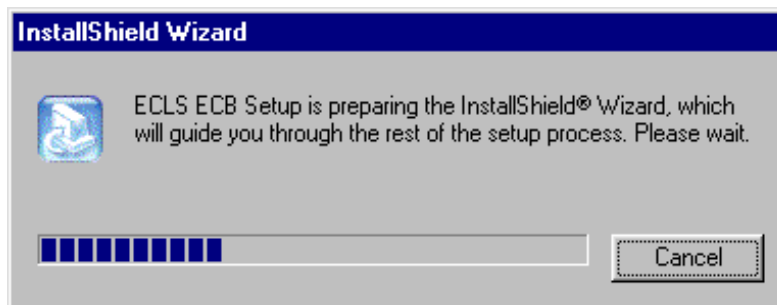
5. Click on the OK button to start the installation. You will now see several installation screens, some of which will prompt you for a response.

Depending on your PC's configuration, you may encounter warning messages during installation. To respond, always keep the newer version of a file being copied and ignore any access violations that occur during file copying.

If you are installing multiple ECBs (not different versions of the same ECB) on your PC, you may receive a message warning that setup is about to replace pre-existing files. To respond, always opt to continue the installation although the default is to cancel the setup. When you get a followup message to confirm whether the installation should be continued, press "Yes" to continue although the default is "No."

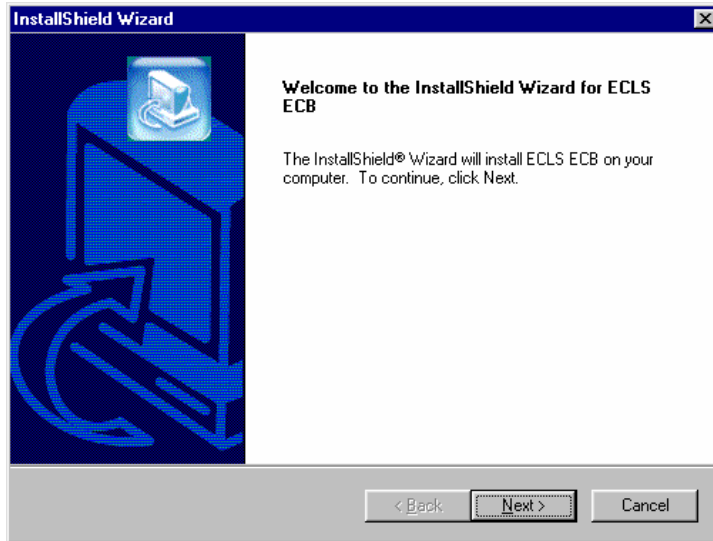
6. The screen shown in exhibit 8-2 indicates that the setup is being prepared.

Exhibit 8-2. InstallShield Wizard



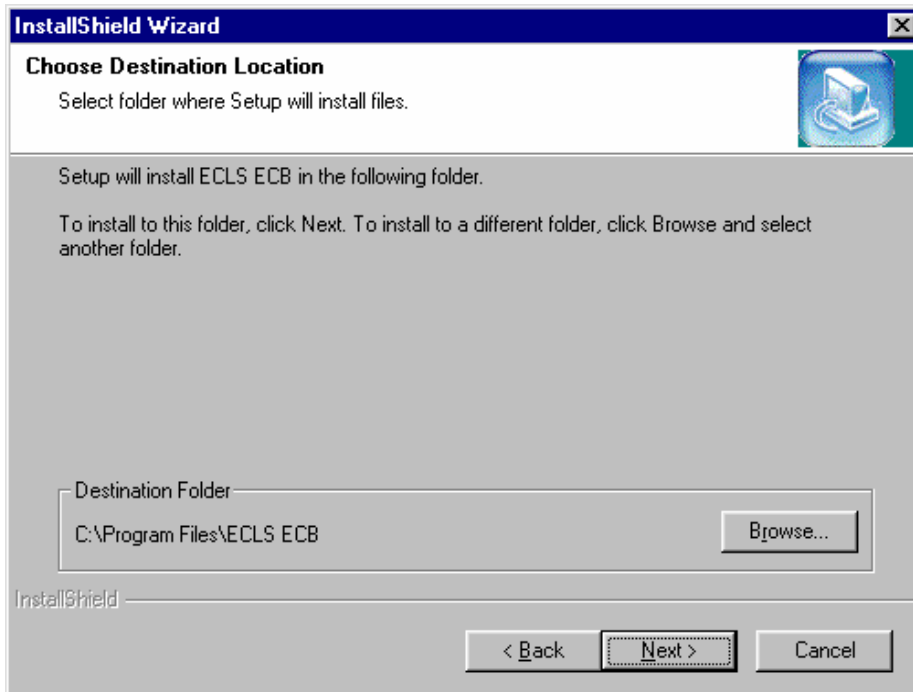
7. You will be prompted to continue with the installation in the Welcome window shown in exhibit 8-3. Click on the Next button to continue.

Exhibit 8-3. Welcome window



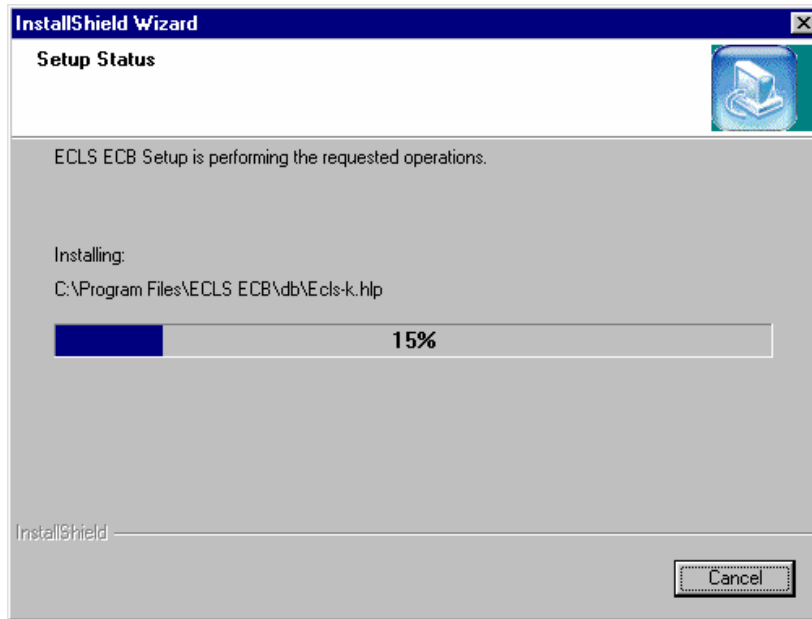
8. When you continue, you will be prompted to choose a destination location for the installation in the window shown in exhibit 8-4. If you wish to change the destination location, click on the Browse button to change the directory. Click on the Next button when the desirable destination folder is shown.

Exhibit 8-4. Choose Destination Location



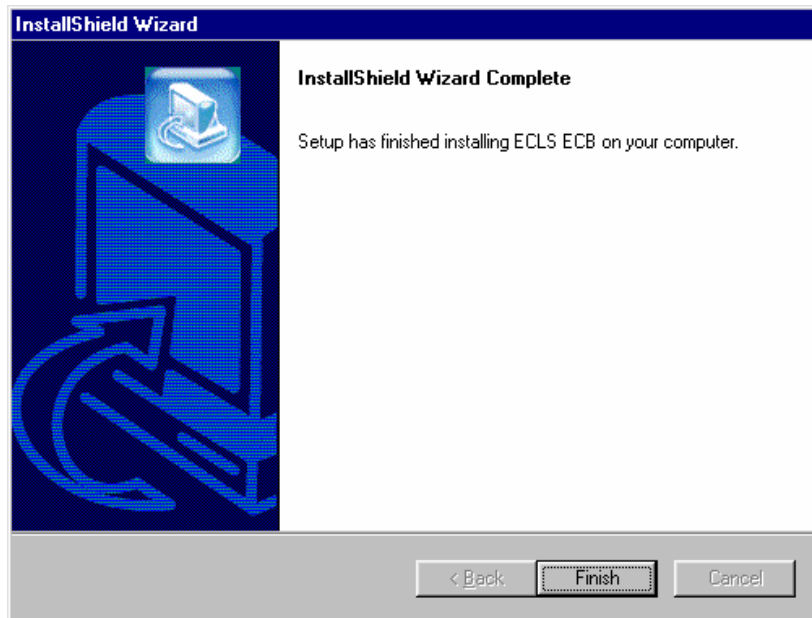
9. Setup will then start installing files. Exhibit 8-5 shows the setup status.

Exhibit 8-5. Setup Status



10. Once the installation is completed, the InstallShield Wizard Complete window shown in exhibit 8-6 will appear. Click on the Finish button to finish the process and return to your PC's desktop.

Exhibit 8-6. InstallShield Wizard Complete



11. The installation process should take about a minute, depending on the speed of the computer on which the ECB is being installed.

Another option for installing the ECB software is to go to the Start menu and go to Settings. Select Control Panel and select Add/Remove Programs from the options. Click on the Install button and follow the directions. Make sure the ECB CD-ROM is in the CD-ROM drive before starting. The program will automatically find the file Setup.exe in the CD-ROM and begin installation. The process will begin at point 5 in the section above.

8.2.2 Starting the ECB

Now that you have installed the ECB on your PC, you can start the program by simply selecting it from the Windows Start, Programs Menu, ECB.

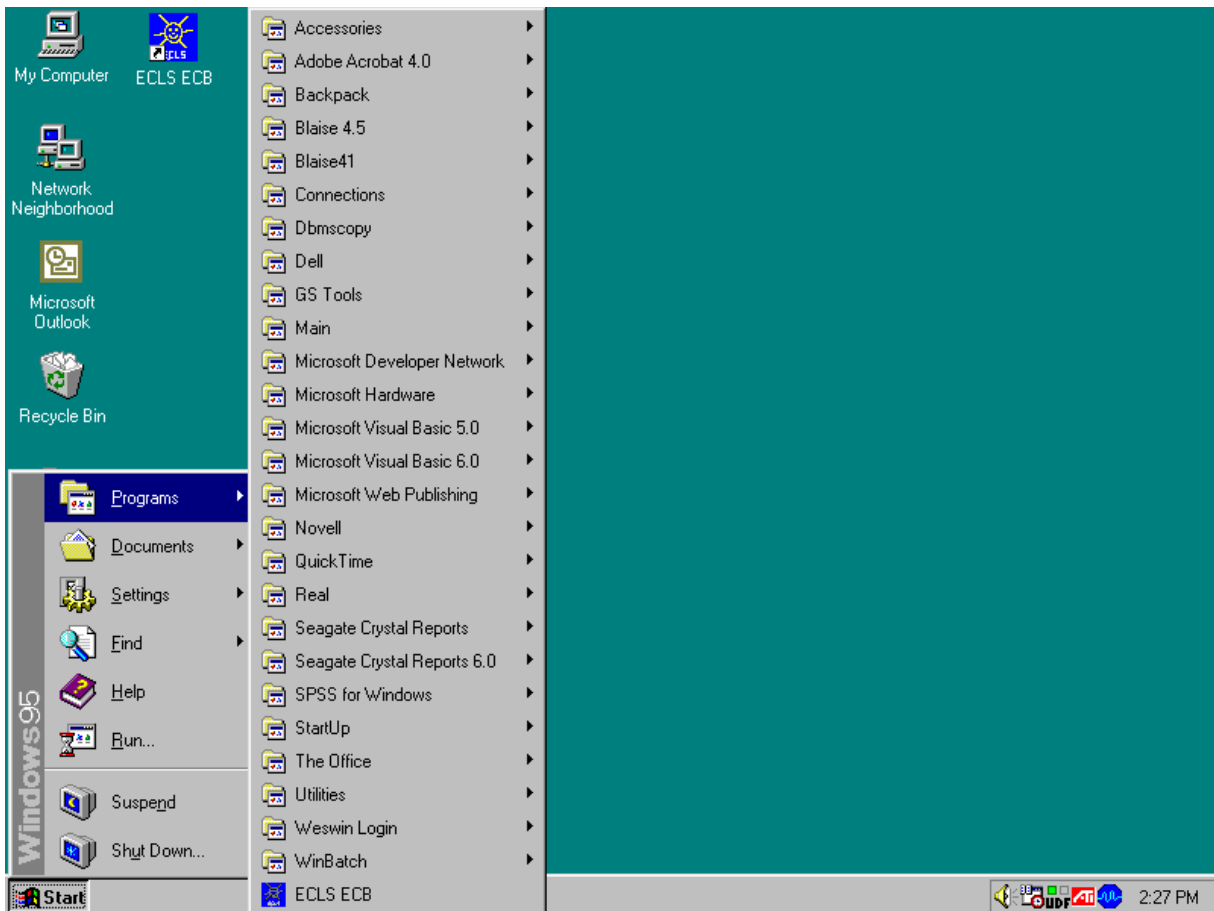
How to Start the ECB:

1. On the desktop screen (exhibit 8-7b), click on the Start button and then point to Programs. Click on the ECB title to invoke the program. Alternatively, click on the ECB desktop icon (exhibit 8-7a) shown below to invoke the program.

Exhibit 8-7a. Desktop icon

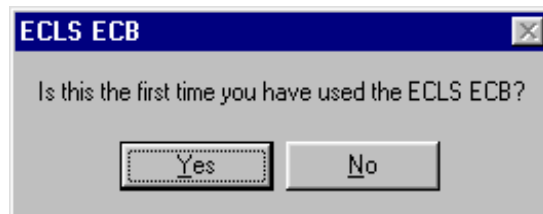


Exhibit 8-7b. Desktop screen—click start



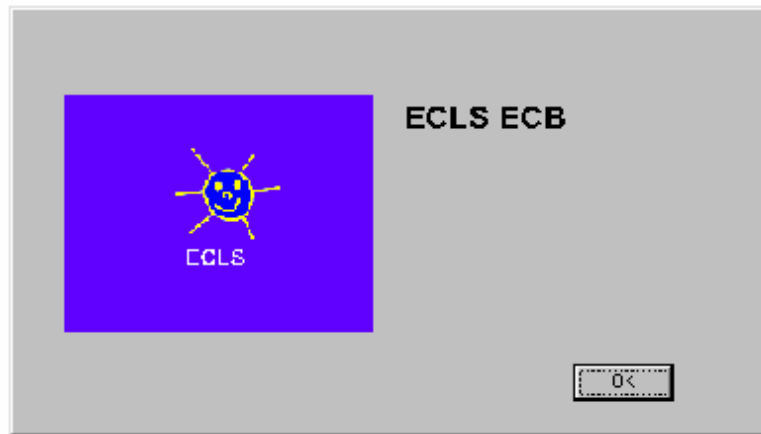
2. If you are a first-time user of the ECB, exhibit 8-8 will display and ask if you are a new ECB user.

Exhibit 8-8. First-time user dialog box



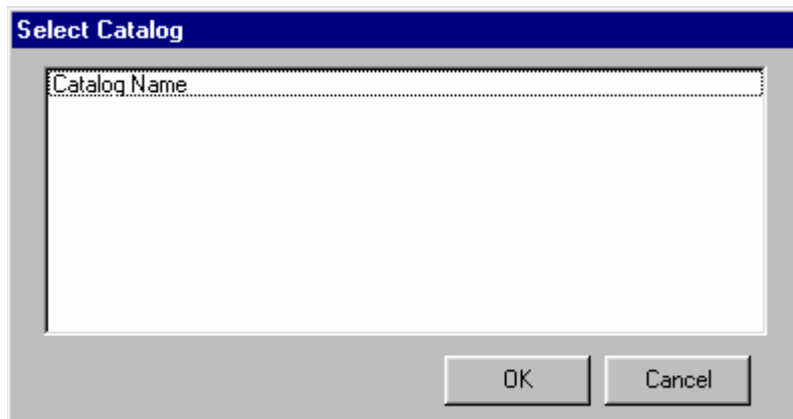
3. Click “Yes” if you are a first-time user. The ECB splash-screen shown in exhibit 8-9 will display.

Exhibit 8-9. ECB splash screen



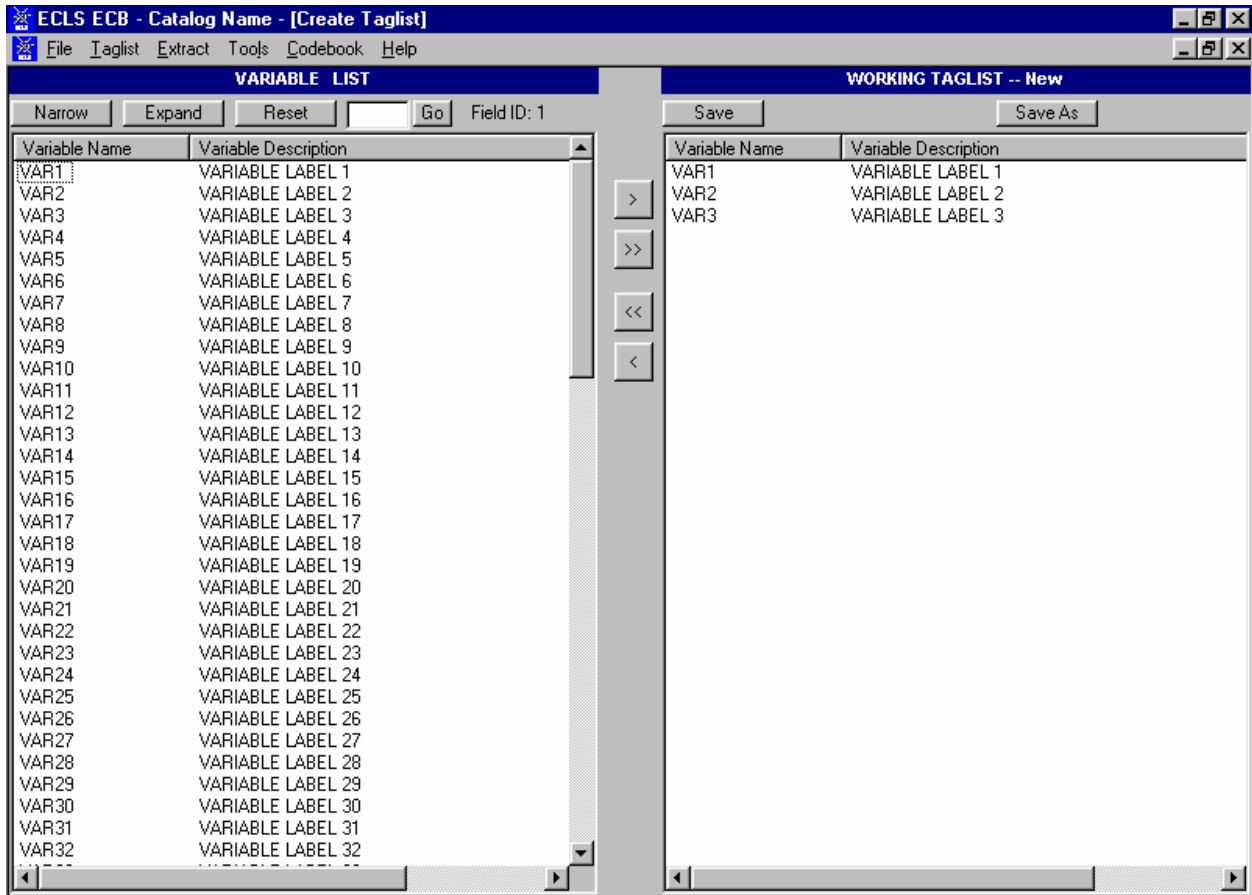
4. On the Select Catalog screen (exhibit 8-10), highlight the name of the catalog. (The ECB has only one catalog.)

Exhibit 8-10. Select Catalog screen



5. Click OK to open the Main ECB screen, shown in exhibit 8-11.

Exhibit 8-11. Main ECB screen



6. You are now ready to use the functions of the ECB as described in the following sections.

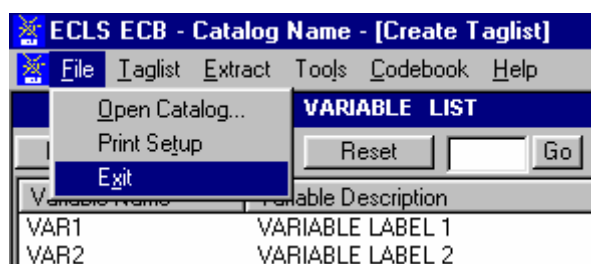
8.2.3 Exiting the ECB

The ECB can be shut down at any time; however, you will be prompted to save any unsaved information.

How To Shut Down the ECB:

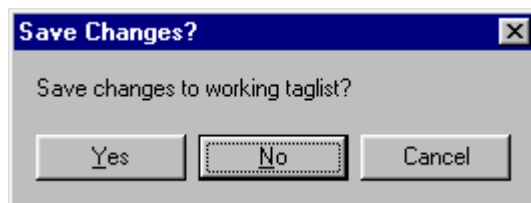
1. From the File menu, click on the Exit option as shown in exhibit 8-12.

Exhibit 8-12. Exit screen



2. If you have not saved your Working Taglist, you will be prompted with the dialog box shown in exhibit 8-13.

Exhibit 8-13. Save working taglist dialog box



3. If you DO NOT wish to save your Working Taglist, click on the “No” button. If you DO wish to save your Working Taglist, click the “Yes” button. For more information, refer to section 8.4.4, Saving Taglists.

8.2.4 Removing the ECB Program From Your Personal Computer

How to Uninstall the ECB:

1. Click on the Windows Start button.
2. Select the Settings menu.
3. In the Control Panel window, click on the Add/Remove Programs.
4. Select “ECB” and click on the Add/Remove button.

5. Follow any prompts. You will be prompted by the InstallShield Wizard to confirm the uninstallation and finish the process.
6. The program is designed so that the uninstallation will keep the taglists when the ECB program is uninstalled in order that all the saved taglists will be retained when the ECB is reinstalled. As a result, the uninstallation will not remove the directory where the ECB was located.

8.2.5 Title Bar

The Title Bar, shown below in exhibit 8-14, is the horizontal bar located at the top of the main screen. It will list the name of the program and the catalog that you have opened, and it will indicate that you are in the “Create Taglist” mode.

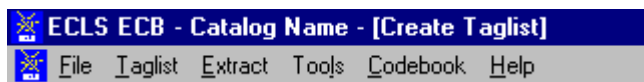
Exhibit 8-14. Title Bar



8.2.6 Menu Bar

Selecting items from the pulldown menus listed on the Menu Bar (exhibit 8-15) provides access to the available action commands. Section 8.6 shows the choices and functions available within each menu.

Exhibit 8-15. Menu Bar



How to Access the Menu Bar Items:

1. Point to an item on the Menu Bar and click.
2. Click on a command from the dropdown list.

The Menu Bar may also be activated and its options selected using the shortcut keys described in section 8.2.7.

8.2.7 Using Shortcut Keys to Navigate

The shortcut keys provide a means for selecting menu options and screen buttons without the use of a mouse. These shortcut keys are identified by an underscore under the shortcut letter within the option or button label. The menus that appear on the windows are activated by simultaneously selecting the <ALT> key and the underscored letter. An example of this is the activation of the Taglist Menu by selecting the key combination of <ALT>-<T>. Once the menu is activated and all options are displayed, the options can be selected by then pressing the underscored letter for the desired option or by pressing the arrow keys to move between the options.

Not all screens have shortcut keys. They may, however, be used without mouse capability by pressing the <TAB> key. The <TAB> key moves the cursor or highlight through the options and buttons within the windows. When the desired option or button is highlighted, it can be selected by pressing the <ENTER> key.

8.3 Variable List

The ECB main screen, shown in exhibit 8-16, comprises two primary lists that each provide functions for reviewing, grouping, and extracting variable data from the opened catalog. These lists include the Variable List and the Working Taglist.

The Variable List, shown in exhibit 8-17, is a list of all variables associated with the current catalog. When you first open a catalog, all variables contained in the catalog are displayed in the Variable List. Once the catalog is open and the Variable List is displayed, you can scroll through the list using the scrollbar controls at the right side of the Variable List screen. Additionally, you can press <PgUp> and <PgDn> to scroll the list one screen at a time. <Ctrl><Home> and <Ctrl><End> will move to the first and last variable in the list, respectively. Also, the arrow keys can be used to move through the list of variable names.

The “Field ID” at the upper right corner of the Variable List shows the field ID of the selected variable on the Variable List.

Exhibit 8-16. ECB main screen

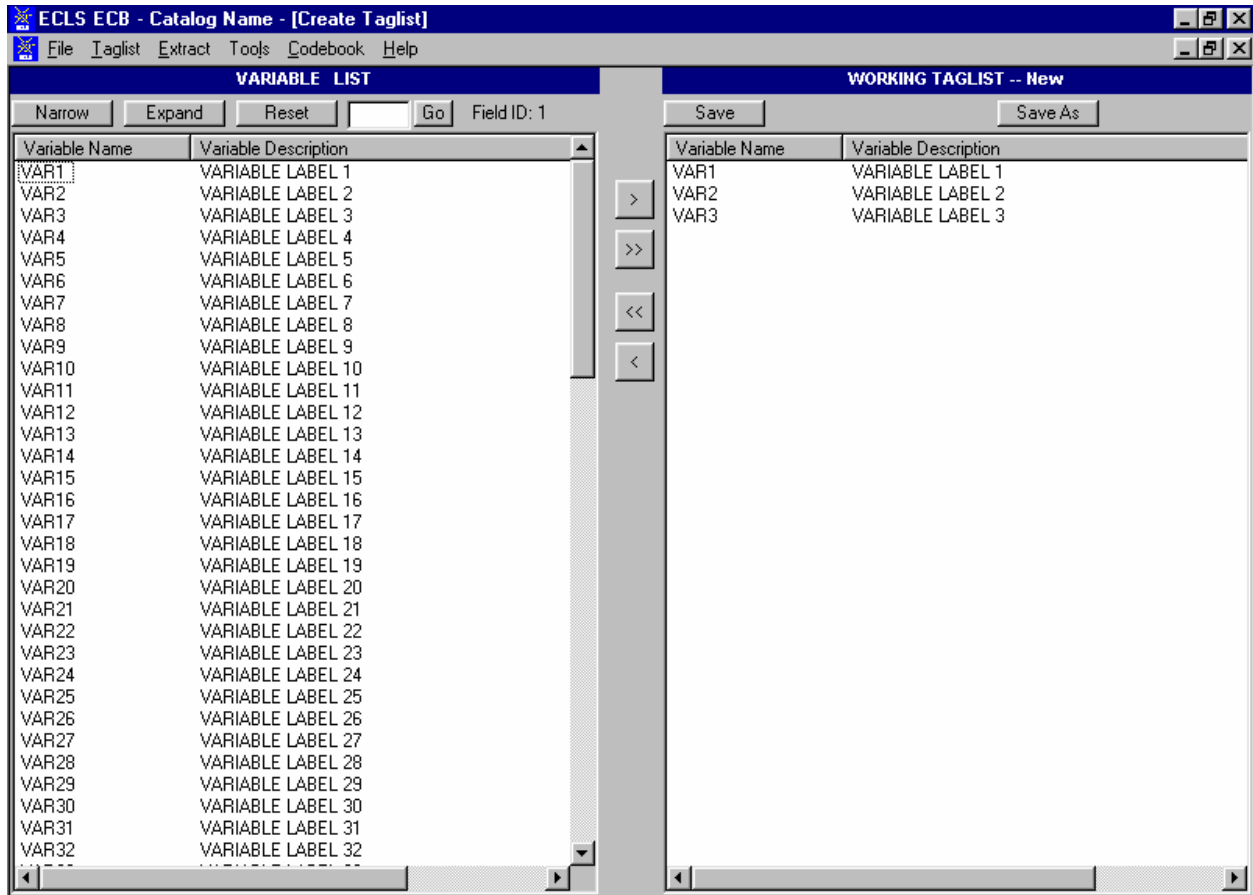
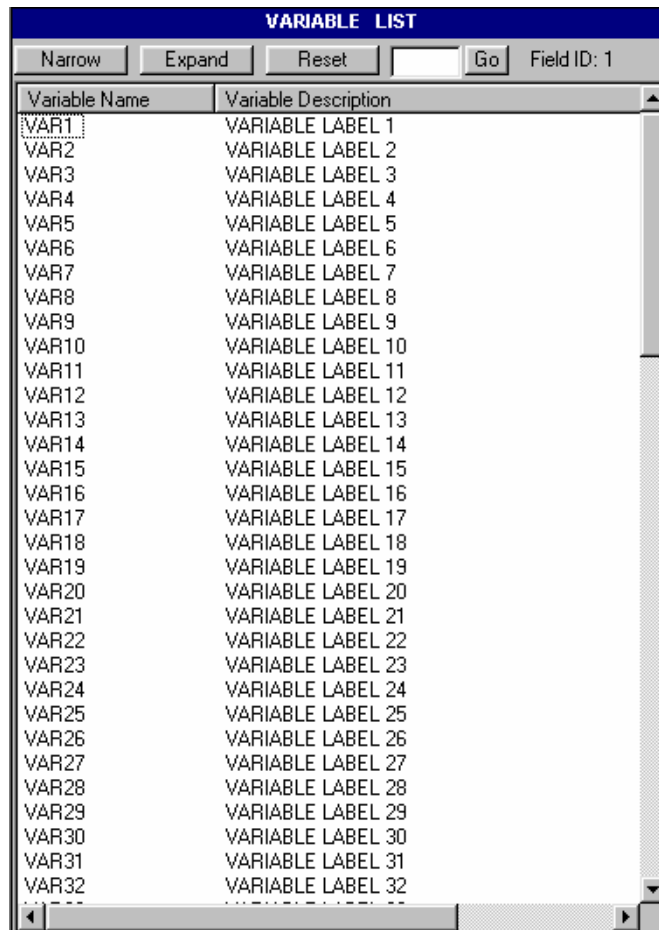


Exhibit 8-17. Variable List



The screenshot shows a window titled "VARIABLE LIST". At the top, there are buttons for "Narrow", "Expand", "Reset", and "Go", along with a text field containing "Field ID: 1". Below this is a table with two columns: "Variable Name" and "Variable Description". The table lists 32 variables, each with a name (VAR1 through VAR32) and a corresponding description (VARIABLE LABEL 1 through VARIABLE LABEL 32). The first row, VAR1, is highlighted with a dashed border. The table has a scrollbar on the right side.

Variable Name	Variable Description
VAR1	VARIABLE LABEL 1
VAR2	VARIABLE LABEL 2
VAR3	VARIABLE LABEL 3
VAR4	VARIABLE LABEL 4
VAR5	VARIABLE LABEL 5
VAR6	VARIABLE LABEL 6
VAR7	VARIABLE LABEL 7
VAR8	VARIABLE LABEL 8
VAR9	VARIABLE LABEL 9
VAR10	VARIABLE LABEL 10
VAR11	VARIABLE LABEL 11
VAR12	VARIABLE LABEL 12
VAR13	VARIABLE LABEL 13
VAR14	VARIABLE LABEL 14
VAR15	VARIABLE LABEL 15
VAR16	VARIABLE LABEL 16
VAR17	VARIABLE LABEL 17
VAR18	VARIABLE LABEL 18
VAR19	VARIABLE LABEL 19
VAR20	VARIABLE LABEL 20
VAR21	VARIABLE LABEL 21
VAR22	VARIABLE LABEL 22
VAR23	VARIABLE LABEL 23
VAR24	VARIABLE LABEL 24
VAR25	VARIABLE LABEL 25
VAR26	VARIABLE LABEL 26
VAR27	VARIABLE LABEL 27
VAR28	VARIABLE LABEL 28
VAR29	VARIABLE LABEL 29
VAR30	VARIABLE LABEL 30
VAR31	VARIABLE LABEL 31
VAR32	VARIABLE LABEL 32

The Variable List provides you with a comprehensive means of reviewing and identifying the variables that you want to use. To help you select the desired variables, the ECB provides you with the following capabilities:

- Perform searches of variable names and descriptions (see section 8.3.1);
- View code book information for each variable (see section 8.4.9); and
- Move selected variables to a Working Taglist (see section 8.4.2).

8.3.1 Searching the Code Book for Variables

The ECB allows you to search a catalog's Variable List for variables meeting criteria you specify. The Narrow Search and Expand Search functions are used to develop and refine the variables listed in your Variable List before adding them to your Working Taglist. Help screens with topical variable groupings were designed for each catalog to expedite searching. The catalog-specific topical variable groupings can be found in exhibit 8-60 at the end of this chapter.

8.3.1.1 Using the Go Button

Using the Go button, located at the top of the Variable List column, allows you to quickly move to a particular variable in the Variable List. You use the field ID presented in the help screens described earlier.

How To Use the Go Button:

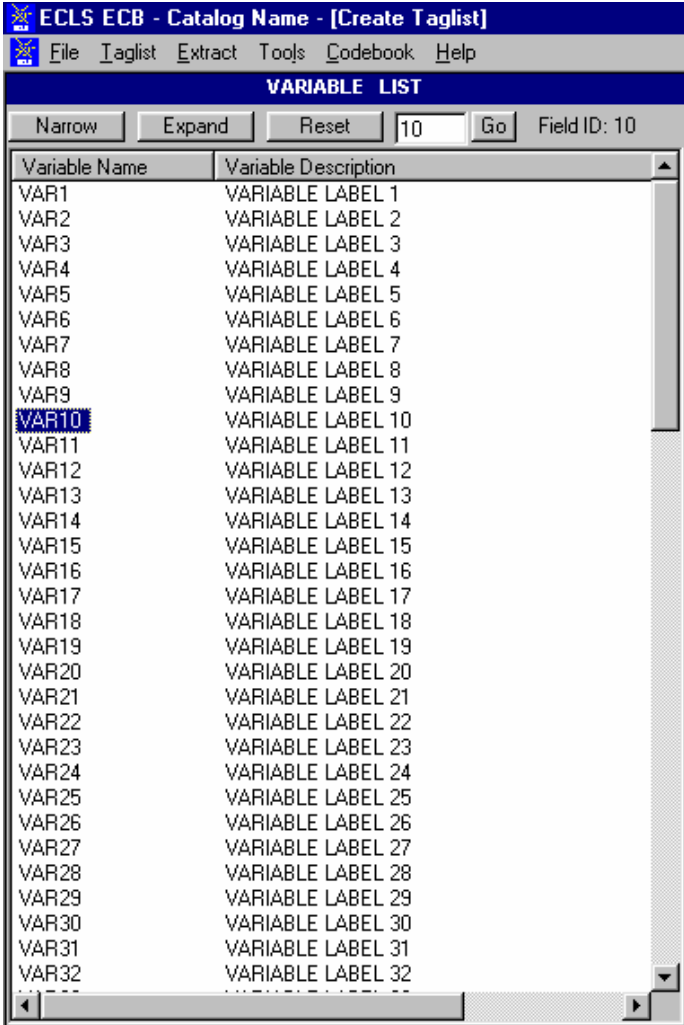
1. Type the field ID in the input box on the left of the Go button.
2. Click on the Go button.
3. The Variable List will then scroll down automatically to show the selected variable.
4. The selected variable is highlighted.
5. The field ID of the current variable selected is shown on the right of the Go button (exhibit 8-18).
6. Click the Reset button to return to the top of the original Variable List (Field ID 1) or enter another field ID to scroll to another variable.

For field IDs that identify different groups of variables, please refer to exhibit 8-60 for the catalog-specific topical variable groupings.

The Go button will not be available in a narrowed or expanded list. After a Narrow Search or an Expand Search, you must reset the Variable List (see section 8.3.1.4) before you can use the Go button.

The “Field ID” remains active in a narrowed or expanded list. However, the field IDs indicate the order of the variables in the catalog rather than that in the Variable List. As a result, the field IDs would not change in a narrowed or expanded list.

Exhibit 8-18. Go button



8.3.1.2 Narrowing Your Variable Search

The Narrow Search function can be used to narrow the list of variables displayed in the Variable List. Since some catalogs have several thousand variables, this feature helps eliminate the variables that do not apply to your analysis. In performing the Narrow Search, you can enter key

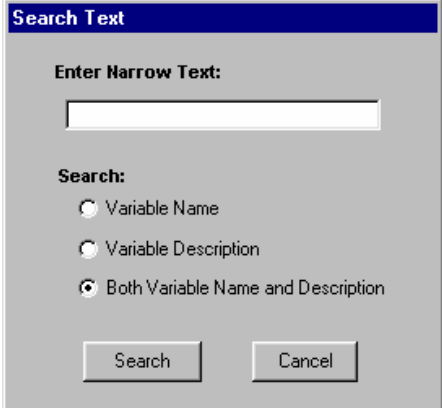
characters, words, or phrases as your criteria for searching the variable names, variable descriptions, or both. Use the keywords in the “Variable Identifier” column in exhibit 8-60 to search the catalog variables by variable description. If “Variable Name” appears under the variable in the Variable Identifier column, you must search by variable name. Also, the Narrow Search can be performed multiple times allowing you to repeatedly refine the list of variables displayed in the Variable List column.

Performing the Narrow Search function will only narrow down the variables listed in the Variable List window and will not affect those in the Working Taglist window.

How To Conduct a Narrow Search:

1. Click on the Narrow button located above the Variable List window.
2. The Narrow Search dialog box appears as shown in exhibit 8-19.

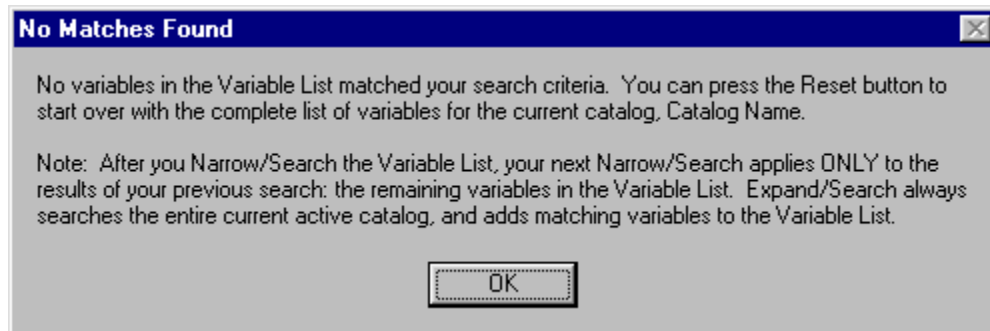
Exhibit 8-19. Narrow Search Text dialog box



3. Enter a key character string, word, or phrase in the Enter Narrow Text field. Character strings can include a single alphanumeric character or a sequence of several characters. The search is not case sensitive. The results returned will be all entries that contain that exact sequence of letters, numbers, spaces, and words.
4. Click in the Variable Name, Variable Description, or Both Variable Name and Description radio button to specify where to search.
5. Click on the Search button to initiate the search.
6. The variables meeting the specified criteria will be displayed in the Variables List column.

If no variable names or descriptions in the catalog contain the specified search text, then the message shown in exhibit 8-20 will appear.

Exhibit 8-20. No Matches Found message



7. Repeat the Narrow Search procedure if necessary.

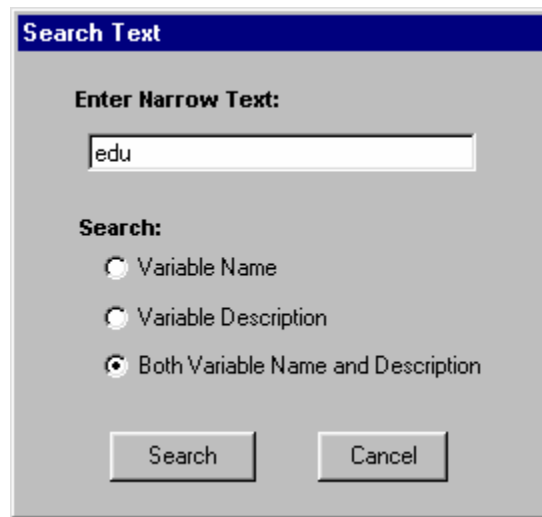
Please note that the field ID at the upper right corner of the Variable List reflects the order of the variables in the catalog rather than that in the narrowed Variable List.

Example of Narrowing a Search

The following example shows you how to narrow the Variable List. In this example, you want to include all the variables from the catalog that measure education. Do the following:

1. In the Variable List, click on the Narrow button.
2. In the Search Text Box (shown in exhibit 8-21), type in “edu” and then click on the Search button.

Exhibit 8-21. Example of narrowing a search



The image shows a dialog box titled "Search Text". It has a blue header bar with the text "Search Text". Below the header, there is a section labeled "Enter Narrow Text:" with a text input field containing the word "edu". Below that is a section labeled "Search:" with three radio button options: "Variable Name", "Variable Description", and "Both Variable Name and Description". The "Both Variable Name and Description" option is selected. At the bottom of the dialog are two buttons: "Search" and "Cancel".

3. The new Variable List will include only the variables that have the text “edu” in the variable name or the variable description.

Exhibit 8-60 (at the end of this chapter) displays the appropriate keywords from the associated catalog. Simply find the topic of interest in the Topic column first. And then enter in the Search Text Box the matching keywords in the Variable Identifier to narrow the search.

8.3.1.3 Expanding Your Variable Search

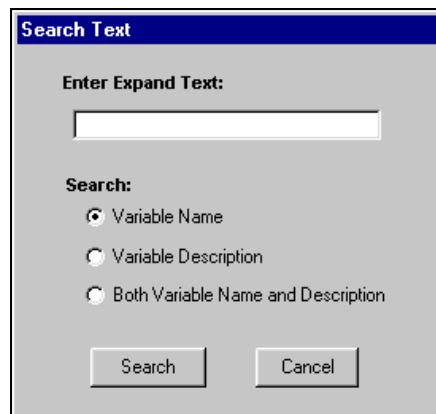
The Expand Search function can be used to expand a previously narrowed list of variables displayed in the Variable List. After performing a Narrow Search operation, you can add variables to your current Variable List that meet your specified criteria. In performing the Expand Search, you can enter key characters, words, or phrases as your criteria for searching the variable names, variable descriptions, or both. Also, the Expand Search can be performed multiple times, allowing you to repeatedly expand the list of variables displayed in the Variable List column.

Performing the Expand Search function will only expand the variables listed in the Variable List window and will not affect those in the Working Taglist window.

How To Conduct an Expand Search:

1. Click on the Expand button located above the Variable List window.
2. The Expand Search dialog box will appear as shown in exhibit 8-22.

Exhibit 8-22. Expand Search Text dialog box



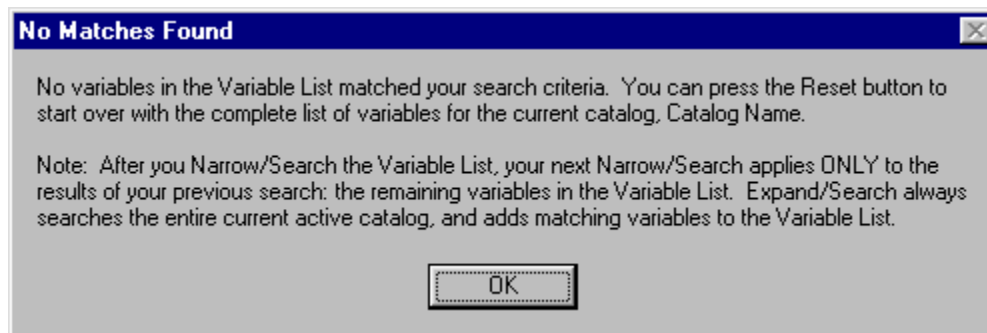
The image shows a dialog box titled "Search Text". It contains a text input field labeled "Enter Expand Text:". Below this field is a "Search:" section with three radio buttons: "Variable Name" (which is selected), "Variable Description", and "Both Variable Name and Description". At the bottom of the dialog are two buttons: "Search" and "Cancel".

3. Enter a key character string, word, or phrase in the Enter Expand Text field. Character strings can include a single alphanumeric character or a sequence of several characters. The search is not case sensitive. The results returned will be all entries that contain that exact sequence of letters, numbers, spaces, and words.
4. Click in the Variable Name, Variable Description, or Both Variable Name and Description radio button to specify where to search.
5. Click on the Search button to initiate the search.
6. The variables meeting the specified criteria will be added to the variables already displayed in the Variables List column.
7. Repeat the Expand Search procedure if necessary.

If no variable names or descriptions in the catalog contain the specified search text, then the message shown in exhibit 8-23 will appear.

Please note that the field ID at the upper right corner of the Variable List reflects the order of the variables in the catalog rather than that in the expanded Variables List.

Exhibit 8-23. No Matches Found message



8.3.1.4 Resetting Your Variable List

Following a narrowing or expanding of the Variable List as described earlier, it is possible to reset the list to display ALL of the variables available in the catalog. The Variable List is reset by clicking on the Reset button located at the top of the Variable List column. Resetting the Variable List does not affect the variables listed in the Working Taglist.

8.4 Working Taglist

The Working Taglist, shown in exhibit 8-24, displays a list of variables that are currently selected or tagged for extraction. All Working Taglists contain a set of variables, called required variables that will be automatically included in all data files that the user creates. The required variables provide a foundational data set upon which other variables rely. These required variables cannot be untagged or deleted from the Working Taglist by the user. When a catalog is first opened, the default Working Taglist consists of only the required variables for that catalog. (See exhibit 8-58 for the catalog-specific required variables.) To create a taglist, add the variables you have selected to the required variables.

Exhibit 8-24. ECB Working Taglist

Variable Name	Variable Description
VAR1	VARIABLE LABEL 1
VAR2	VARIABLE LABEL 2
VAR3	VARIABLE LABEL 3

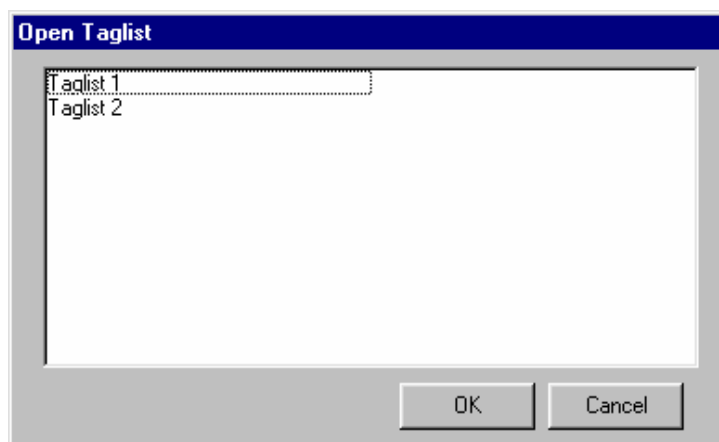
8.4.1 Opening a Taglist

The ECB allows you to open a predefined or previously saved taglist and display it in the Working Taglist column. Taglists, however, are saved as part of a particular catalog and can only be opened as part of the associated catalog.

How To Open a Taglist:

1. Open a catalog.
2. Select Open from the Taglist pulldown menu.
3. The Open Taglist dialog box, shown in exhibit 8-25, appears.

Exhibit 8-25. Open Taglist dialog box





4. Highlight the taglist that you wish to open.
5. Click on the OK button.

If you have made modifications to the taglist currently open in the Working Taglist column, you will be prompted to save your changes.

8.4.2 Adding Variables to the Working Taglist

Variables can be added to your Working Taglist after you have identified the variables in the ECB's catalog that you want to extract. The user-selected variables can be added to the Working Taglist by selecting one of the two command buttons described in exhibit 8-26. The Working Taglist may also have variables added to it from a previously saved taglist. When moving or adding variables to the Working Taglist, the ECB will not permit variables to be listed multiple times. This is an automatic feature of the ECB.

Exhibit 8-26. Add variables buttons

Command Button	Description
	The Tag button moves variables that are selected in the Variable List to the Working Taglist for extraction.
	The Tag All button moves all variables in the Variable List to the Working Taglist for extraction.

Multiple variables can be selected by using the following Microsoft Windows[®] techniques:

- Simultaneously pressing the <SHIFT> + Up/Down arrow keys or
- Pressing <CTRL> + left-mouse clicking on the items to be selected (or deselected). Also, <SHIFT> + left-mouse clicking extends the selection to include all list items between the current selection and the location of the click.

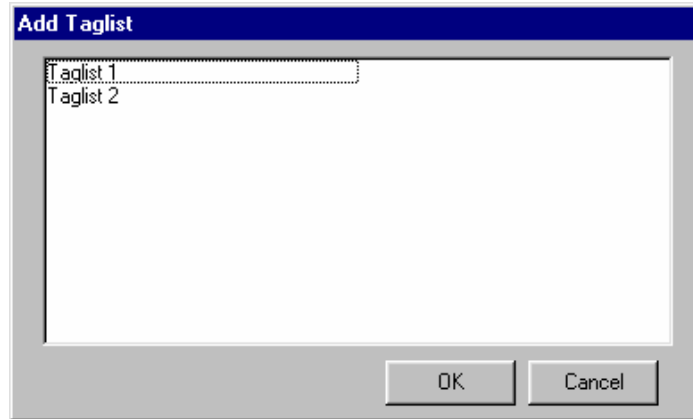
How To Add Variables to a Working Taglist:

1. Highlight the variable(s) in the Variables List that you wish to add. (See Microsoft Windows[®] techniques discussed earlier.)
2. Click on the Tag button, and the selected variables are added to your Working Taglist. To add all variables from the catalog displayed in the Variable List window to your Working Taglist, click on the Tag All button.

How To Add Variables From Another Taglist:

1. Click on the Taglist pulldown menu to display the menu options.
2. Select the Add option to display a list of previously saved taglists, shown in exhibit 8-27.
3. Highlight the saved taglist whose variables you wish to add to your Working Taglist.
4. Click on the OK button.
5. The new variables are added to your Working Taglist.



Exhibit 8-27. Add Taglist dialog box



8.4.3 Removing Variables From the Working Taglist

Variables are removed from your Working Taglist by selecting one or more of the nonrequired variables and clicking one of the two command buttons described in exhibit 8-28. All variables can be removed by clicking on the Untag All button. All but the required variables will be deleted from your Working Taglist. Required variables are variables that are automatically extracted for all user-created files and cannot be removed from the taglist by the user.

Exhibit 8-28. Remove variables buttons

Command Button	Description
	The Untag button removes variables that are selected from the Working Taglist.
	The Untag All button removes all non-required variables from the Working Taglist.

Attempting to remove or untag required variables from the Working Taglist is not permitted by the ECB. A message will be displayed indicating that the required variable cannot be untagged.

How To Untag Variables From the Working Taglist:

1. Highlight the variable(s) in the Working Taglist that you wish to remove. (See Microsoft Windows[®] techniques discussed in previous page.)

2. Click on the Untag button, and the selected variables are removed from your Working Taglist. To remove all nonrequired variables from the Working Taglist, click on the Untag All button.

8.4.4 Saving Taglists

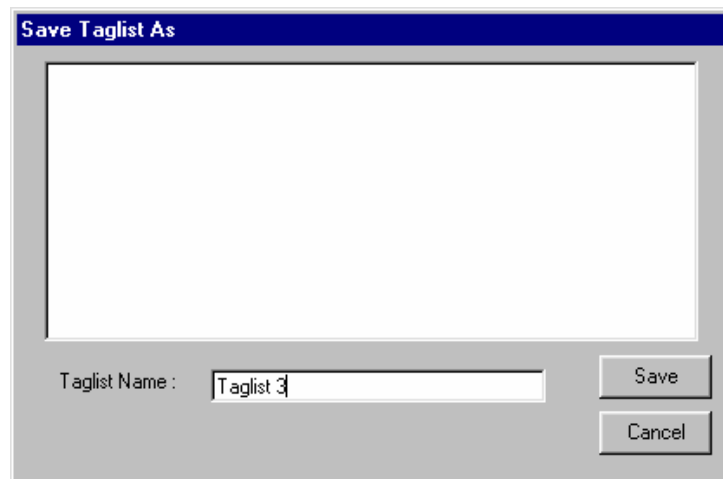
The ECB has the ability to save the newly created or modified taglist displayed in the Working Taglist column. Taglists can be saved either under the name already assigned or under a new name. If you have opened a new taglist and have not yet assigned it a name, you will be presented with the Save As dialog box. If you have opened a predefined taglist and have made modifications to it, you must save the modified taglist to a new name. You will also be prompted to save your Working Taglist changes if you attempt to close the catalog or if you open or import another taglist.

How To Save a New Taglist:

1. Complete any changes you wish to make to the new taglist.
2. Click on the Save or Save As button above the Working Taglist column. You can also select the Save or Save As options from the Taglist pulldown menu.
3. The Save Taglist As dialog box appears as shown in exhibit 8-29.
4. Enter the new name for the taglist in the Taglist Name field.
5. Click on the Save button.
6. The newly assigned taglist name now appears in the Working Taglist header bar.

If a name that already exists is entered, you will be prompted to replace the old taglist with the new taglist. Click “Yes” only if you wish to replace the old taglist with the new taglist.

Exhibit 8-29. Save Taglist As dialog box

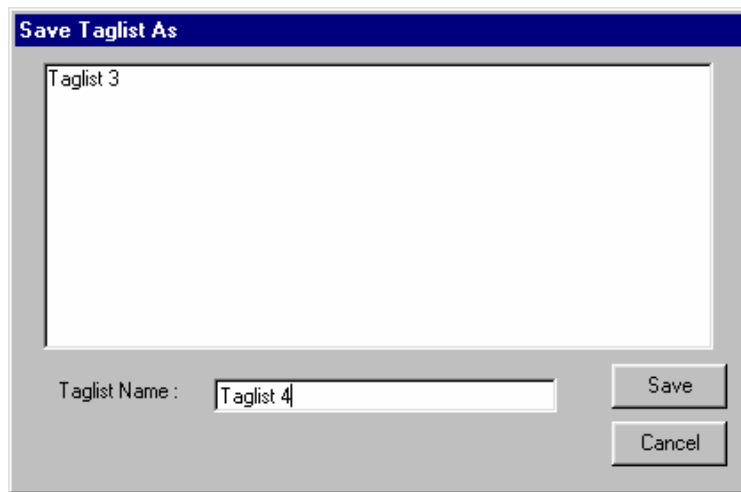


How To Save an Existing Taglist Under a New Name:

1. Complete any changes you wish to make to the existing taglist.
2. Click on the Save As button above the Working Taglist column. You can also click on the Taglist pulldown menu and select the Save As option.
3. The Save Taglist As dialog box appears, shown in exhibit 8-30, with the current taglist name in the Taglist Name field.
4. Enter the new name of the taglist in the Taglist Name field.
5. Click on the Save button.
6. The newly assigned taglist name now appears in the Working Taglist header bar.

If a name that already exists is entered, you will be prompted to replace the old taglist with the new taglist. Click “Yes” only if you wish to replace the old taglist with the new taglist or enter a unique name.

Exhibit 8-30. Save Taglist As dialog box (#2)



8.4.5 Exporting Taglists

Taglists can be saved as external files (*.tlt) for distribution. However, the exported files should be accessed only through the ECBs. Manually modifying the files outside of the ECB software is not recommended.

How To Export a Taglist:

1. Add to the Working Taglist all the variables that you would like to export.
2. Click on the Taglist pulldown menu (exhibit 8-31) and select the Export option.
3. The Export Working Taglist To dialog box appears.
4. Enter the file name for your taglist.
5. Click on the Save button.
6. You will be prompted to replace the file if the file name you entered already exists. Do so or click on “No” to enter a new file name.

The Working Taglist will be saved under the filename you enter.

Exhibit 8-31. Pulldown menu to select Taglist Export

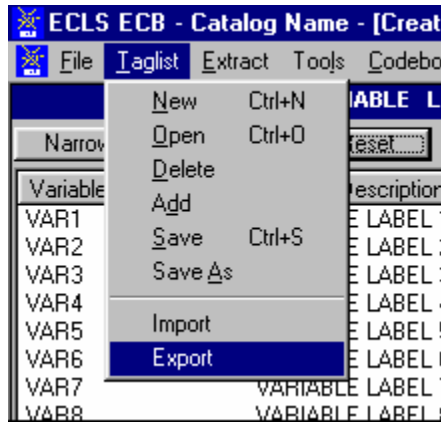
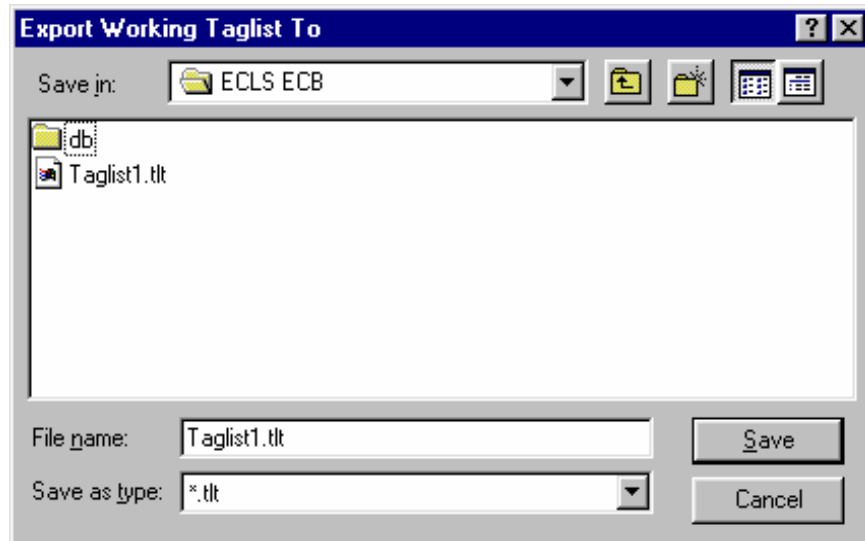


Exhibit 8-32. Export Taglist dialog box



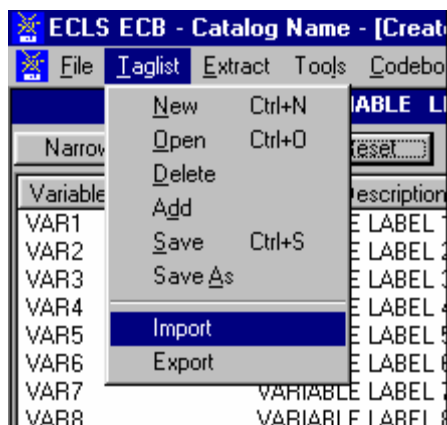
8.4.6 Importing Taglists

Taglists can be imported to the Working Taglist from external *.tlt files that are created by the ECB Taglist/Export function. Please note that only taglists exported from the same catalog of the same version ECB should be imported.

How To Import a Taglist:

1. Save the current Working Taglist before importing new taglist if desired.
2. Click on the Taglist pulldown menu (exhibit 8-33) and select the Import option.

Exhibit 8-33. Pulldown menu to select Taglist Import



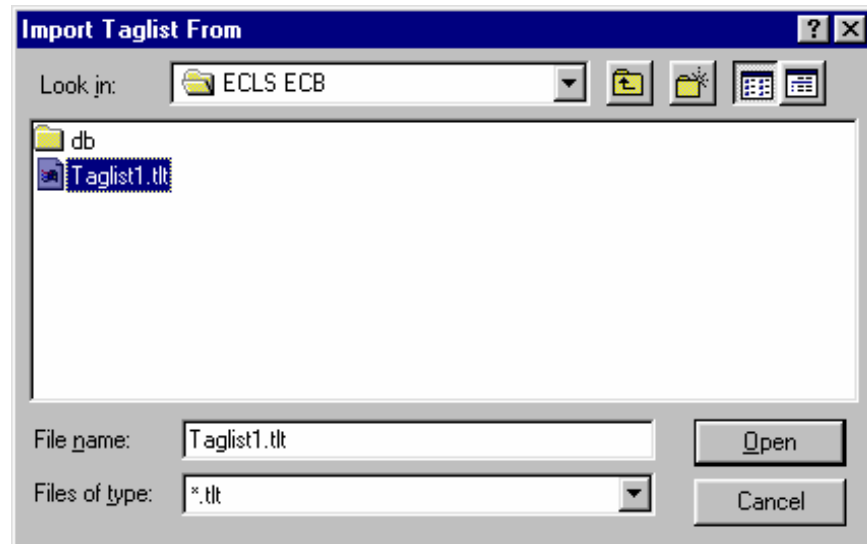
3. You will be prompted to save the current Working Taglist if unsaved changes have been made. Save the taglist if desired.
4. The Import Taglist From dialog box appears.
5. Enter the file name for the taglist you want to import.
6. Click on the Open button.

The Working Taglist will be replaced by the new imported taglist.

8.4.7 Using Predefined Taglists

The ECB provides predefined taglists that address specific topics. These predefined taglists can be added to your Working Taglist or can be opened as a new Working Taglist. Opening these predefined taglists is performed using the same steps as opening a user-saved taglist presented in section 8.4.1. Users can add as many of the predefined taglists as desired to the open Working Taglist. See section 8.7.2 for listings and descriptions of the catalog-specific predefined taglists.

Exhibit 8-34. Import Taglist dialog box



8.4.8 Deleting Taglists

The ECB provides the capability to permanently delete previously saved taglists. Predefined taglists provided with the ECB, however, cannot be deleted through this function.

How To Delete a Taglist:

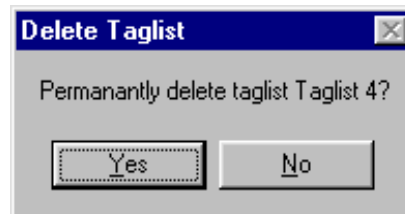
1. Close the taglist currently displayed in the Working Taglist column by selecting the New option from the Taglist pulldown menu.
2. The Working Taglist will be replaced by a New taglist.
3. Click on the Taglist pulldown menu and select the Delete option.
4. The Delete Taglist selection screen, shown in exhibit 8-35, appears with the taglists listed that may be deleted.
5. Highlight the taglist that is to be deleted and click on the OK button.
6. A confirmation screen, shown in exhibit 8-36, verifies your intention to delete the taglist.
7. Click on the "Yes" button to permanently delete the saved taglist.

Please note that you cannot delete the taglist that is currently open as the Working Taglist.

Exhibit 8-35. Delete Taglist selection



Exhibit 8-36. Delete Taglist confirmation window



8.4.9 Viewing Code Book and Variable Information

The code book for a taglist displayed in the Working Taglist column can be created, viewed, and printed from the ECB main screen. The code book displays several pieces of information about each variable that are described in exhibit 8-37.

Exhibit 8-37. Code book information

Field	Description
Question Text	The question that was asked of the respondent by the interviewer or that was on the self-administered instruments.
Variable Name/ Description	The name of the variable as it appears in the catalog and a brief description of its content.
Record Number	The row number of the variable within the catalog data file.
Format	The format of the variable. The first character is either “A” or “N” for alphabetical or numeric. Most variables are numeric except the identifiers—which begin with an “A.” The number following the “A” or “N” is the length of the variable. For numeric variables, the number after the decimal point is the number of decimal places.
Comment	Information to clarify specific information about a variable.
Position	The column number (position) of the variable within the catalog data file.
Response	A brief statement of each response code’s meaning.
Codes	The numeric codes specifying each response.
Frequency	The numeric count of respondents providing the corresponding response code. The frequency counts are unweighted.
Percent	The percentage of respondents providing the corresponding response code. The percents are unweighted.

How To View the Code Book for Tagged Variables:

1. Complete any changes you wish to make to the displayed taglist.
2. Click on the Code Book pulldown menu and select the View option.
3. The code book for the current taglist opens in a new window as shown in exhibit 8-38.
4. Use the buttons described in exhibit 8-39 to navigate through the displayed code book.

Exhibit 8-38. Code book view

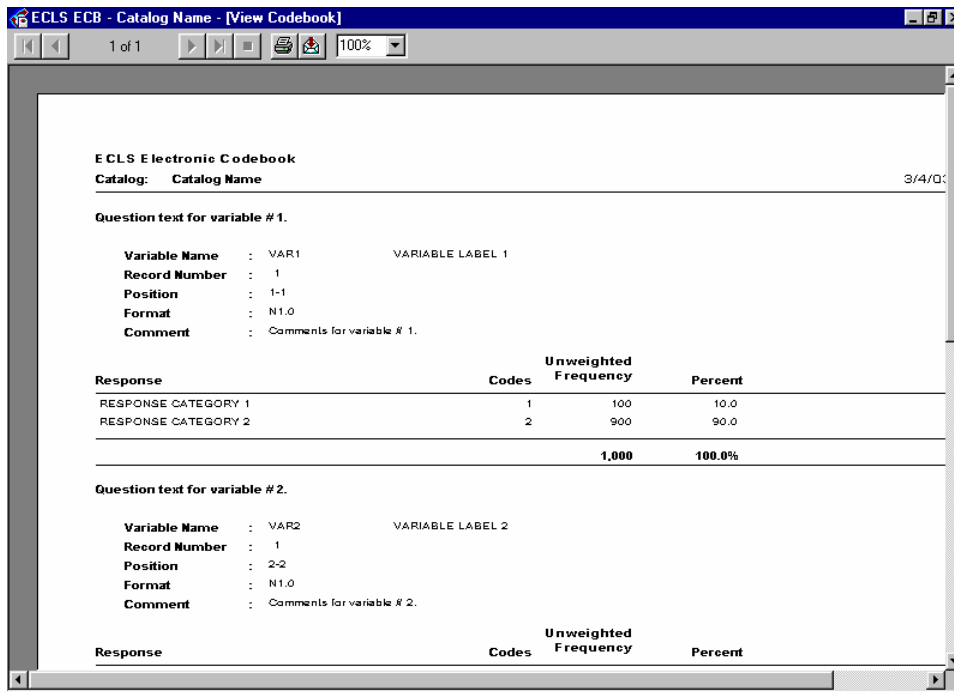







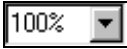


Exhibit 8-39. Navigation buttons

Command button	Description
	Click this button to change the displayed page to the first page.
	Click this button to change to the previous page.
	Click this button to advance to the next page.
	Click this button to change the displayed page to the last page.
	Click this button to discontinue a page change.
	Click this button to print the code book. Refer to the procedure below for steps on printing the code book.
	Click this button to export the code book to a different destination and save it as a different file format. Refer to the procedure below for steps on exporting the code book.
	Click the dropdown arrow to select a display magnification of the code book.

NOTE: The counter “1 of 1+” on the tool bar on top of the screen indicates the current page number and the last page number of the report. Users must navigate to the last page of the report to load the entire report. Once the user has viewed the last page of the report, the “+” sign will disappear and the correct last page number will show.

5. Once you have finished viewing the code book, close the screen by clicking on the Windows “X” control located in the top right corner of the window. You may also close the window using the other standard Windows defaults: by clicking on the windows icon in the upper left corner and selecting Close, or by pressing Alt-F4.

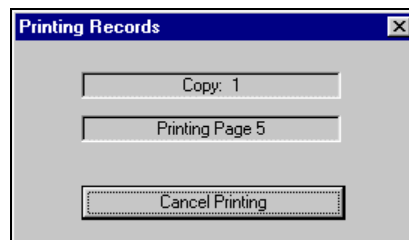
How To Print the Code Book:

1. Complete any changes you wish to make to the displayed taglist.
2. Click on the Code Book pulldown menu and select the Print option.
3. The Printing Status screen, shown in exhibit 8-40, appears, and the code book prints on your PC’s default printer.

How To Export the Code Book:

1. Complete any changes you wish to make to the displayed taglist.
2. Click on the Code Book pulldown menu and select the View option.

Exhibit 8-40. Printing status screen




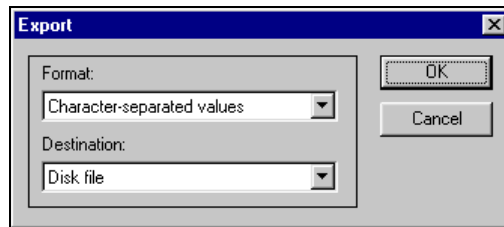
3. The code book for the current taglist opens in a new window, similar to the one shown in exhibit 8-38.
4. Click on the Export code book button: 
5. The Export code book selection screen, shown in exhibit 8-41, appears.

Exhibit 8-41. Export code book selection screen



6. Select the desired options from the “Format” pulldown menu and the “Destination” pulldown menu.
7. Click on the OK button and complete any subsequent screens required for exporting the file.

Please note that exporting a catalog in its entirety will take a long time due to the large size.

The code book and its variables can be selected to display their information from either the Variable List or the Working Taglist. The information that can be displayed for a variable includes the variable name and label, the question wording associated with the variable, the position and format of the variable on the data file, each response value and its label, unweighted frequencies, and the unweighted percentage distributions as listed on exhibit 8-37. The entire code book can also be viewed after moving all of the catalog’s variables to the Working Taglist. The following procedures describe how to view some or all code book variables:

How To Display Information for a Single Code Book Variable:

1. Locate the desired variable from either the Variable List or the Working Taglist.
2. Click on the variable name to highlight it and press <ENTER> -or- double-click on the variable name to view the variable information as shown in exhibit 8-42.

Exhibit 8-42. Variable Quick View

Question text for variable # 1.

Variable Name : VAR1 (VARIABLE LABEL 1)
Record Number : 1
Position : 1-1
Format : N1.0
Comment : Comments for variable # 1.

Response	Code	Unweighted Freq.	Percent
RESPONSE CATEGORY 1	1	100	10.0%
RESPONSE CATEGORY 2	2	900	90.0%
Total		1,000	100.0%

The Variable Name is the only field that can be highlighted for displaying the variable’s code book information. Clicking on the variable description field will not activate the Variable Quick View.

3. When you are done reviewing the variable information, close the window by clicking on the Windows control “X” in the upper right corner of the screen. You’ll return to the main screen.

How to Print Information for a Single Code Book Variable:

The ECB currently does not support printing the information for a single variable directly to the printer. If you must print the information for a single variable, follow these steps:

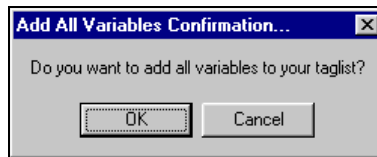
1. Double-click on the variable to activate the Variable Quick View (see the previous “How To” section for details).
2. With the Variable Quick View being the active window on top, press <Alt> + <Print Screen> to save the image of the Variable Quick View window.

3. In any application that supports bitmap images (e.g., Microsoft Paint, Microsoft Word, etc.), paste the saved image.
4. Print the image to the printer using the print function of the application that you are using.

How to Display and Print the Entire Code Book or Selected Pages:

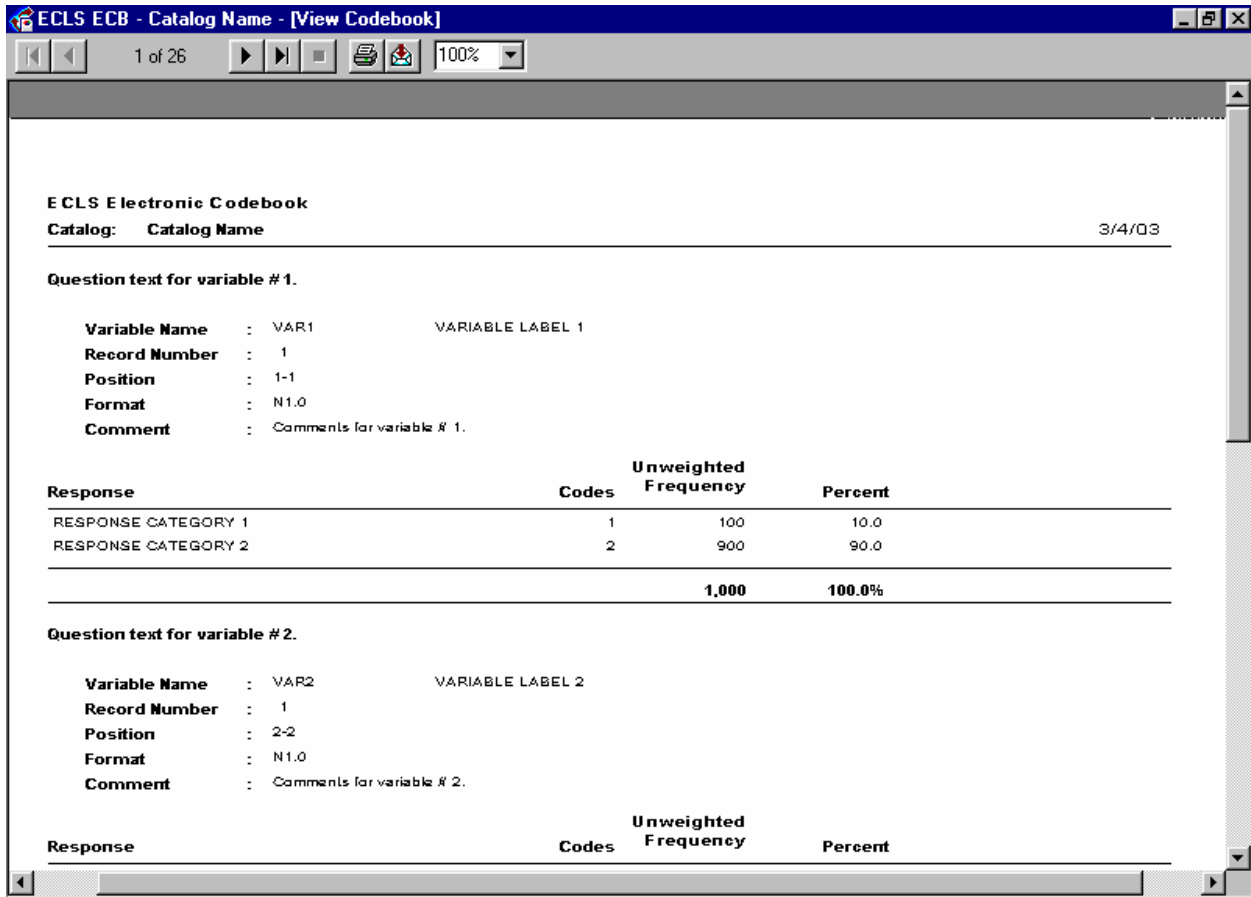
1. Move all of the catalog's variables displayed in the Variable List to the Working Taglist by clicking on the Tag All button.
2. Click on the OK button of the Add All Variables Confirmation dialog box, shown in exhibit 8-43.

Exhibit 8-43. Add All Variables dialog box



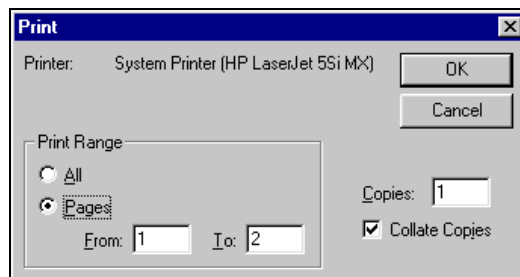
3. All of the variables listed in the Variable List are now displayed in the Working Taglist.
4. Select View from the Code Book pulldown menu.
5. The entire code book displays as shown in exhibit 8-44. Note that this view includes ALL variables in the catalog and can span more than 1000 pages depending on the size of the ECB. The page number is in the upper left corner of the window.

Exhibit 8-44. View of the entire code book



- To print the entire code book, click on the printer icon displayed at the top of the code book screen. Select ALL from the Printer Dialog box (exhibit 8-45). Enter the number of copies you want and click on the OK button.

Exhibit 8-45. Printer dialog box



7. To print selected pages of the code book, select Pages from the Printer Dialog box. Enter the pages you want to print and the number of copies you want. Click on the OK button.
8. When you are done viewing the entire code book, close the window by clicking on the Windows control “X” in the upper right corner of the screen. You will return to the main screen.

8.5 Extracting Data from the ECB

Once the variables have been selected (tagged) for extraction and reside in the Working Taglist, the next step is to generate the code through which the statistical analysis software can retrieve and display the results. The ECB provides options for generating the code for analyzing data with the SAS, SPSS for Windows, or Stata statistical analysis programs.

To run these programs, you will need the appropriate statistical software and the ECB CD-ROM from which the program can extract data.

SPSS users should note that an entire catalog can produce a Frequencies command statement with more than 500 variables. This may produce a warning of “too many variables,” and the Frequencies command will not execute. Users may work around this limitation by dividing the Variable List into two or more Frequencies commands.

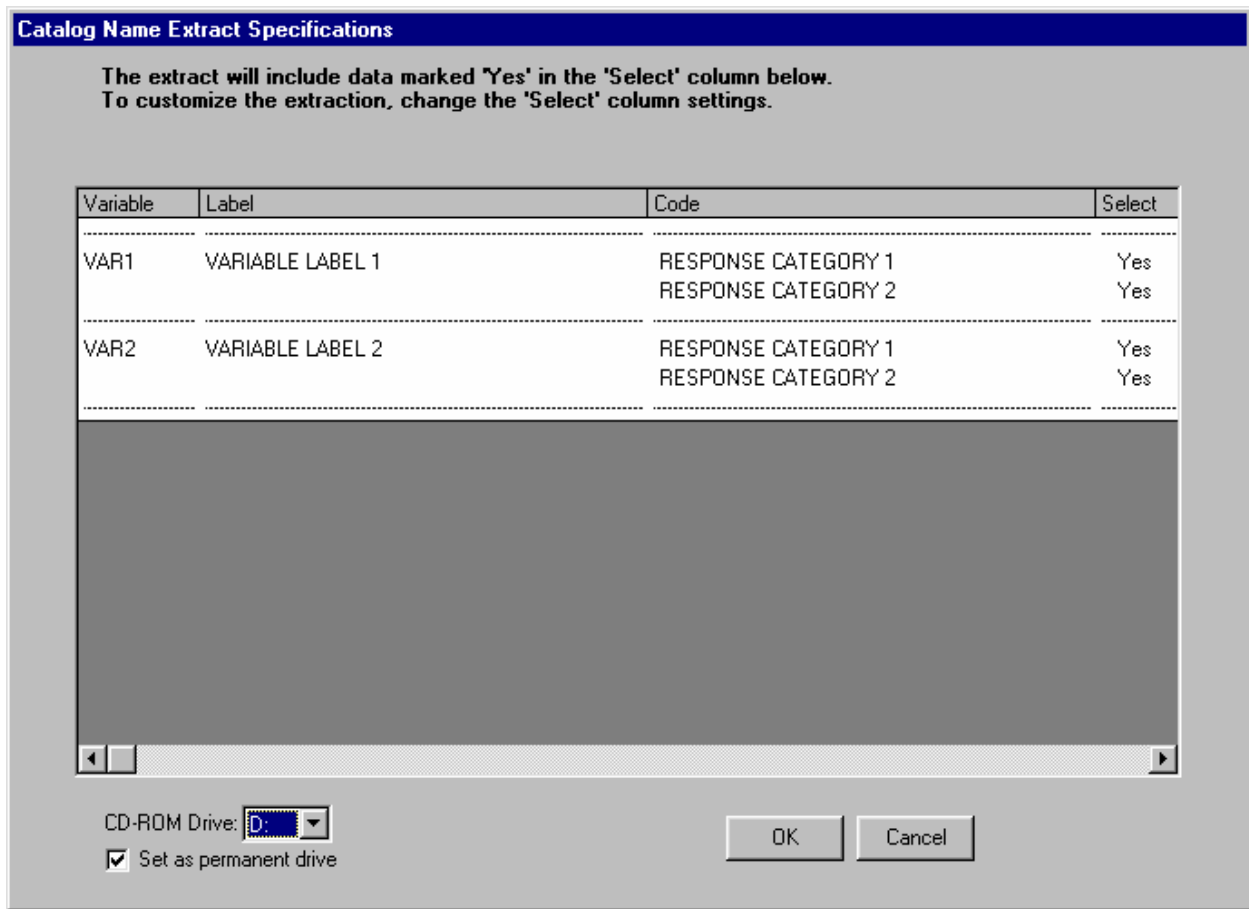
When extracting data to be used with either the SAS, SPSS for Windows, or Stata programs, a dialog box will be presented that allows the user to define the extract population through the Limiting Fields. See exhibit 8-46. The Limiting Fields include various subgroups of respondents that are typically of interest to analysts. These subgroups can be selected or deselected to narrow the data field that is extracted.

Also, please note that the ECB extract function allows the user to specify the drive letter of the CD-ROM drive. If you attempt to run the resulting SAS, SPSS, and Stata programs on a workstation with a different CD-ROM drive letter, you must alter the program code accordingly or regenerate the program code using the ECB.

The SAS, SPSS, or Stata source code generated by the ECB to read in the data may contain code statements that are “commented” out (e.g., with * in SAS). These code statements either run descriptive statistics (e.g., frequencies, means, etc.), or associate formats with variables. They are commented out because not all analysts will want them included in the source code.

SAS users (prior to SAS, Version 8) should note that, although the ECB will allow data set names larger than eight characters, the SAS system will reject these names at run-time.

Exhibit 8-46. Limiting fields dialog box

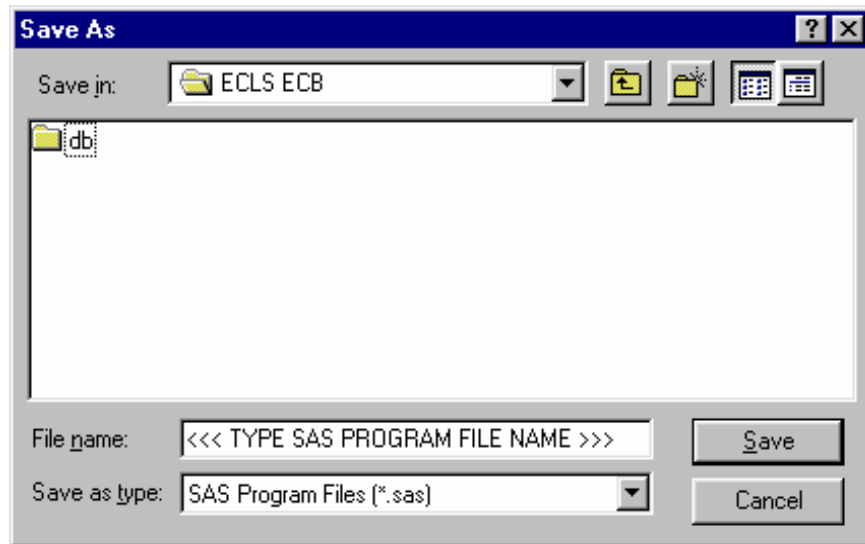


Refer to section 8.7.3 for instructions on using and modifying the catalog-specific limiting variables.

How To Extract a File to SAS Format:

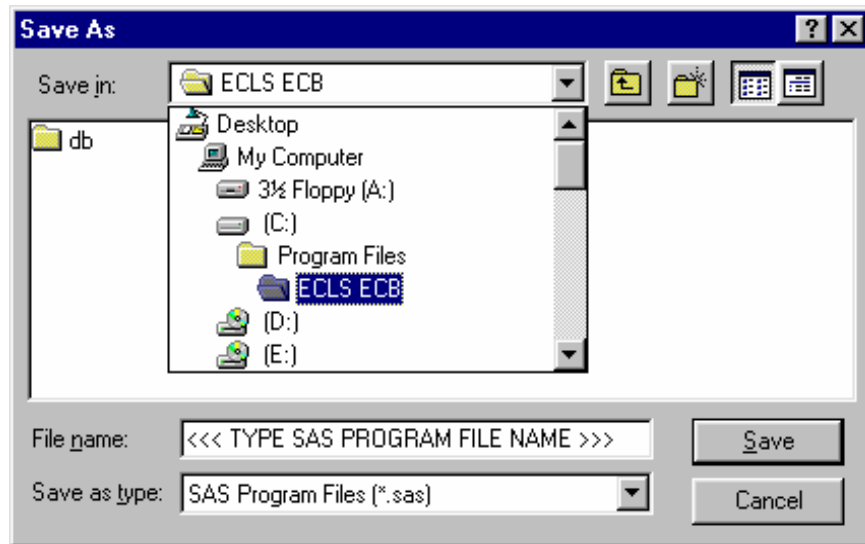
1. Complete any changes you wish to make to the displayed taglist.
2. Click on the Extract pulldown menu and select the SAS option.
3. The Limiting Fields screen for the open catalog appears. Make your selections for each limiting variable indicator.
4. Verify that the ECB CD-ROM is loaded in your PC's default CD-ROM drive and then click on the OK button.
5. Type the desired name of the extract program file in the file name field of the screen shown in exhibit 8-47.

Exhibit 8-47. Save SAS program file dialog box



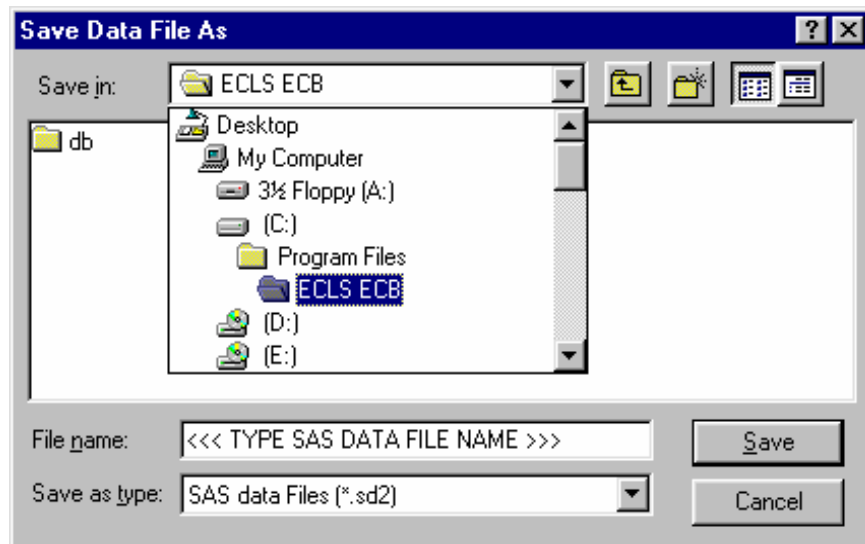
6. To save the file to another directory, click on the “Save in” dropdown menu button to browse to the new location, as shown in exhibit 8-48.

Exhibit 8-48. Save SAS program file location browse screen



7. Click on the Save button to store the file.
8. In the Save Data File As window (exhibit 8-49) type in the file name you want the data file to save to and then click on Save.

Exhibit 8-49. Save SAS data file dialog box



9. Run the saved extract program in SAS to extract the data.

How To Extract a File to SPSS Format:

1. Complete any changes you wish to make to the displayed taglist.
2. Click on the Extract pulldown menu and select the SPSS option.
3. The Limiting Fields screen for the open catalog appears. Make your selections for each limiting variable indicator.
4. Verify that the ECB CD-ROM is loaded in your PC's default CD-ROM drive and then click on the OK button.
5. Type the desired name of the extract program file in the file name field of the screen shown in exhibit 8-50.
6. To save the file to another directory, click on the "Save in" dropdown menu button to browse to the new location, as shown in exhibit 8-51.
7. Click on the Save button to store the file.

Exhibit 8-50. Save SPSS program file dialog box

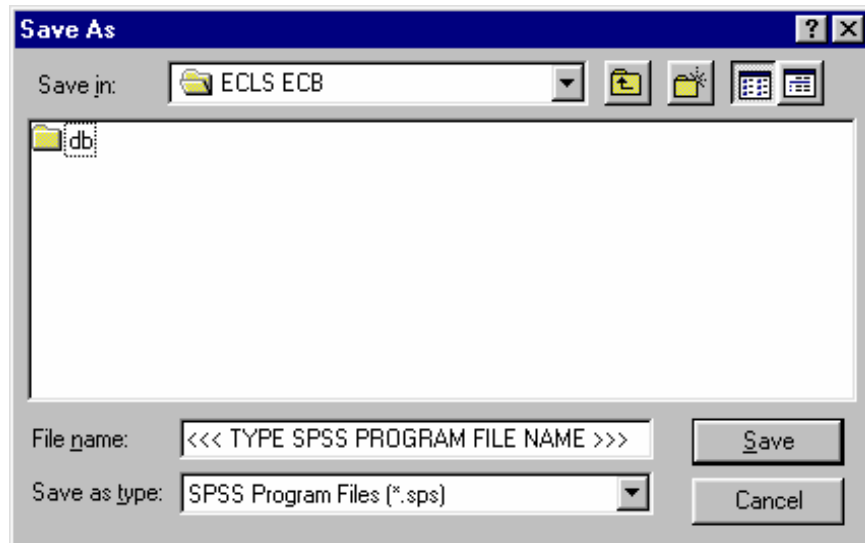
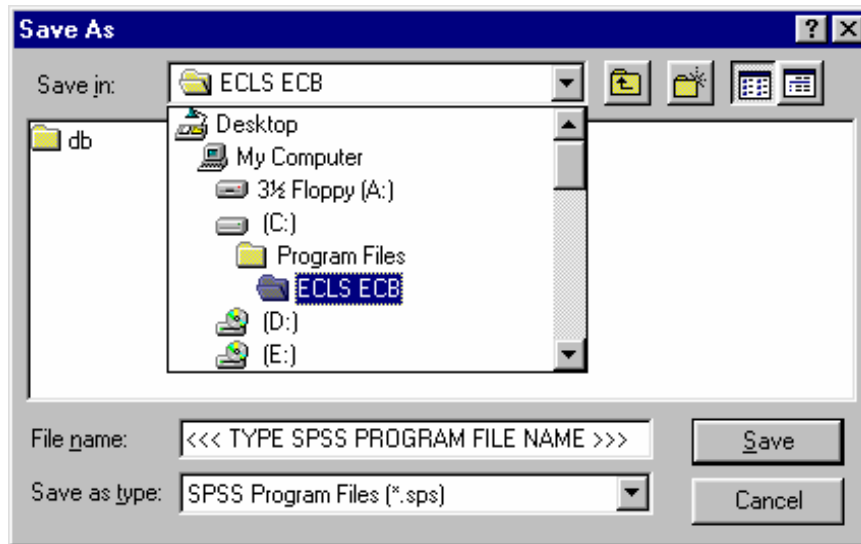
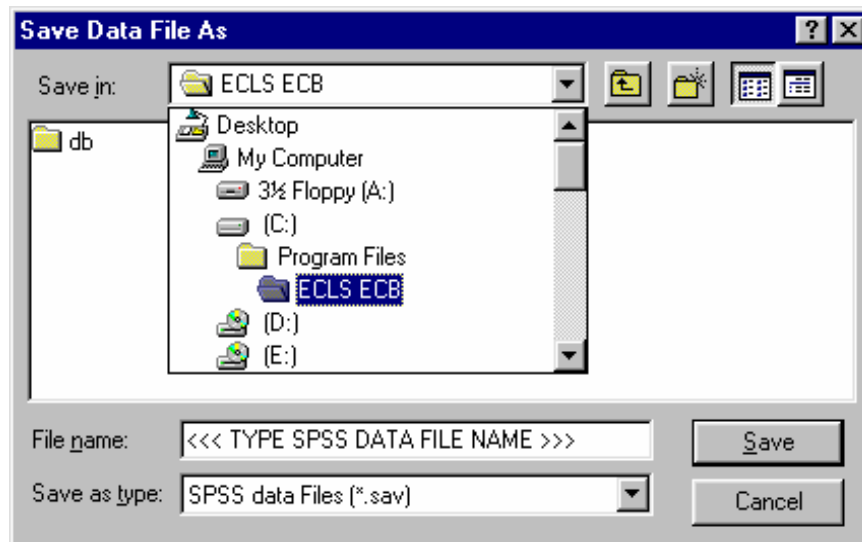


Exhibit 8-51. Save SPSS program file location browse screen



8. In the Save Data File As window (exhibit 8-52), type in the file name you want the data file to save to and then click on Save.
9. Run the saved extract program in SPSS to extract the data.

Exhibit 8-52. Save SPSS data file dialog box



How To Extract a File to Stata Format:

1. Complete any changes you wish to make to the displayed taglist.
2. Click on the Extract pulldown menu and select the Stata option.
3. The Limiting Fields screen for the open catalog appears. Make your selections for each limiting variable indicator.
4. Verify that the ECB CD-ROM is loaded in your PC's default CD-ROM drive and then click on the OK button.
5. Type the desired name of the extract program file in the file name field of the screen shown in exhibit 8-53.
6. To save the file to another directory, click on the "Save in" dropdown menu button to browse to the new location, as shown in exhibit 8-54.
7. Click on the Save button to store the file.
8. In the Save Data File As window (exhibit 8-55), type in the file name you want the data file to save to and then click on Save.
9. Run the saved extract program in Stata to extract the data.

Exhibit 8-53. Save Stata program file dialog box

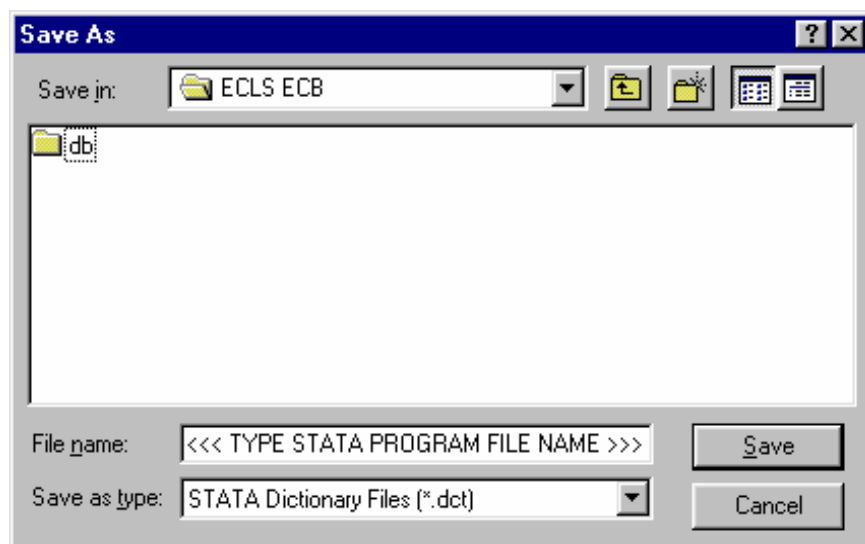


Exhibit 8-54. Save Stata program file location browse screen

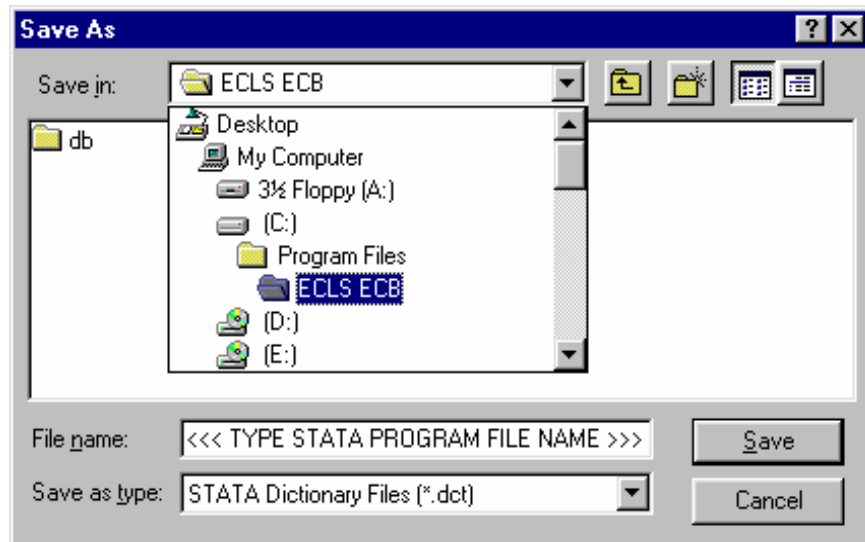
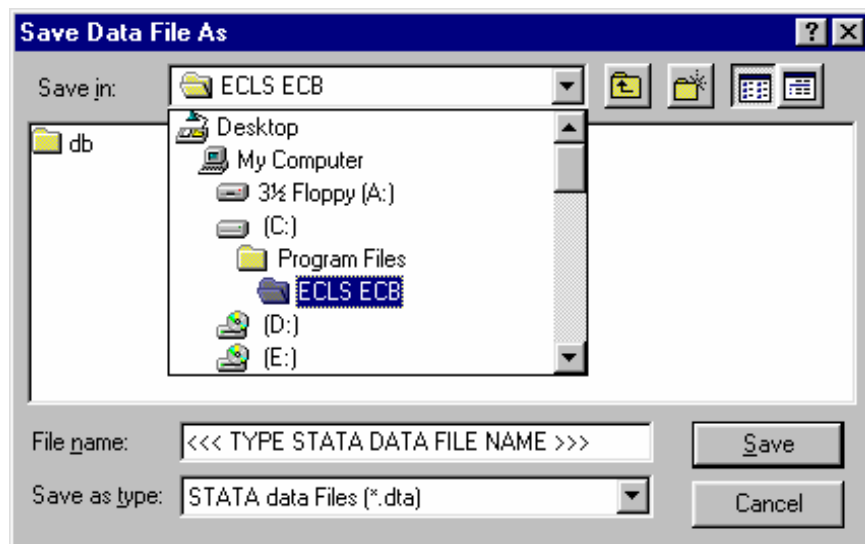


Exhibit 8-55. Save Stata data file dialog box



8.5.1 Reviewing the Extract Specifications

Users should review the SAS, SPSS, or Stata program code that is generated before running it to check that any statements subsetting the data are correct. Note that the ECB sometimes outputs superfluous code for selecting cases; this code is consistent with extract specifications, but users may wish to delete it.

If a mistake in defining the criteria is made, and it is not discovered until after writing out or running the extract program, it is very easy to correct if the taglist was saved before exiting the ECB program. Simply restart the ECB and select the appropriate catalog, open the taglist that you saved, define the extract criteria correctly by modifying the saved taglist as desired and saving it, and write out the extract program again. The program should be reviewed before running it because it may need to be customized.

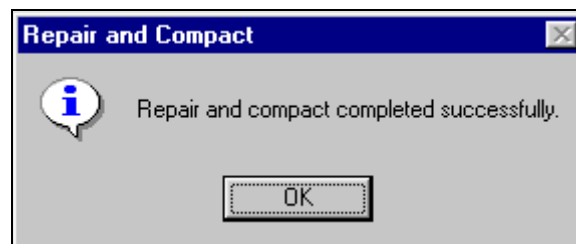
8.5.2 Repairing and Compacting the Database

Periodically users may wish to repair and compact the database that contains the data of the ECB program. If many taglists are created and deleted on a regular basis, the database will contain lingering references to old taglists that are no longer needed. When the database is repaired and compacted, the ECB program “cleans house” and makes the database more efficient. It also decreases the size of the database, so space is conserved.

How To Repair and Compact the ECB Database:

1. Select the Tools pulldown menu and select the Repair and Compact Database option.
2. After a few seconds, the screen shown in exhibit 8-56 appears indicating that the repair and compact of the database was successfully completed.

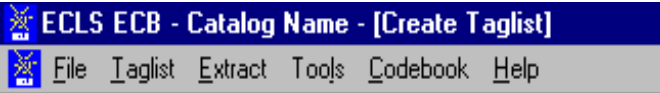
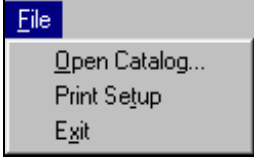

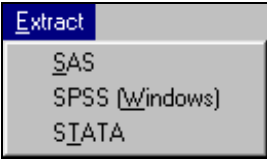



Exhibit 8-56. Repair database completed screen



3. Click on the OK button.

8.6 Menu Bar Descriptions

Exhibit 8-57. Menu Bar Descriptions

	
	<p>The File menu contains the commands needed to:</p> <ul style="list-style-type: none"> ■ Select and open a catalog; ■ Set up your software for printing; and ■ Exit the ECB.
	<p>The Taglist menu contains the commands required to manipulate the variable lists once a catalog has been selected:</p> <ul style="list-style-type: none"> ■ Create a new taglist; ■ Open a previously saved or predefined taglist; ■ Delete a previously saved taglist; ■ Add a previously saved or predefined taglist to the working taglist; ■ Save the working taglist; ■ Save a taglist with another name; ■ Import a previously exported taglist as working taglist and; ■ Export the working taglist for distribution.
	<p>The Extract menu contains options to create a syntax file for:</p> <ul style="list-style-type: none"> ■ SAS; ■ SPSS for Windows; or ■ Stata.
	<p>The Tools menu contains:</p> <ul style="list-style-type: none"> ■ The command for repairing and compacting the database.
	<p>The Code Book menu contains the command for:</p> <ul style="list-style-type: none"> ■ Viewing the entire code book based on the working taglist; and ■ Printing the entire code book based on the working taglist.
	<p>The Help menu provides access to the detailed online help system.</p>

8.7 Child Catalog

This section discusses catalog-specific information. For general information about how the ECB functions work, please refer to previous sections of chapter 8.

The third grade¹ child catalog contains information for children who participated in the third grade ECLS-K survey and includes data at the child level from the child and from the child's parents, teachers, and school. Exhibit 8-58 presents the pretagged required variables that will be automatically extracted for all user-created data files. The user cannot remove the required variables from the taglist. Three of the variables on the list, P1FIRKDG, T5GLVL, and S5PUPRI, are also limiting fields; they may be used to subset the data being extracted. See section 8.7.3 for an explanation and description of limiting fields.

Exhibit 8-58. Child catalog required variables

Field Name	Field Label
CHILDDID	CHILD IDENTIFICATION NUMBER
S5_ID	SPRING 2002 SCHOOL IDENTIFICATION NUMBER
T5_ID	SPRING 2002 TEACHR IDENTIFICATION NUMBER
D5T_ID	SPRING 2002 SPECIAL ED TEACHER ID NUMBER
R5REGION	R5 CENSUS REGION
R5URBAN	R5 LOCATION TYPE - 7 CATEGORIES
R5LOCALE	R5 LOCATION TYPE - 8 CATEGORIES
R5GENDER	R5 CHILD COMPOSITE GENDER
R5RACE	R5 CHILD COMPOSITE RACE
R5AGE	R5 COMPOSITE CHILD ASSESSMENT AGE(MNTHS)
P1FIRKDG	P1 FIRST-TIME KINDERGARTNER
W3SESQ5	W3 CATEGORICAL SES MEASURE
T5GLVL	T5 GRADE LEVEL OF CHILD
S5SCTYP	S5 SCHOOL TYPE FROM THE SFS
S5ENRLS	S5 TOTAL SCHOOL ENROLLMENT
S5PUPRI	S5 PUBLIC OR PRIVATE SCHOOL

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.

¹ Although the great majority of the sampled children were in the third grade in spring 2002, some of them were actually in other grades. For convenience, we simply refer to this as the “third grade” survey, and refer to all the sampled children as “third-graders” in spring 2002.

The child catalog is organized into blocks of data corresponding to specific content areas:

- **Identifiers** are the child, parent, school, and teacher identification number.
- **Field Management System (FMS)** data are for methodological analyses.
- **School demographic** data are obtained from the sampling frame (e.g., census region).
- **Child demographic** composite variables include age, gender, race/ethnicity, and date of birth.
- **Weights** are described in detail in chapter 4. Round 5 weights are for third grade. Cross-year panel weights are described in detail in chapter 9.
- **Flags** indicate unit response for the data source (e.g., direct child assessment, parent interview), or that the parent interview data were edited.
- **Scores** are for the direct and indirect child assessments.
- **Composite variables**, including the child demographic variables, were created to facilitate data analysis. See chapter 7, section 7.4 for a detailed description of the composite variables and table 7-11 for a list of all composite variables.
- **Direct assessment-related data** are the exclusions from the direct assessment (e.g., data from the assessment closing section). Direct assessment items are not released.
- **Parent interview data** are the data items from the parent interview.
- **Teacher questionnaire data** are the data items from the teacher questionnaires.
- **School data** are the data items from the school fact sheet, the school administrator questionnaire, and the facilities checklist.
- **Special education data** are the data items from the special education teacher questionnaires.
- **Student record abstract data** are the data items from the student record abstract form.
- **Taylor Series Sampling Strata and primary sampling units (PSUs)** are the stratum and first-stage unit identifiers that are required for estimating standard errors with the Taylor Series linearization method using software such as SUDAAN or STATA.
- **Weight replicates** are used to estimate standard errors using the paired jackknife replication method (JK2) with WesVar. Each replicate is a subsample of the full sample. The variability of the replicate estimates about the full sample estimate is used to estimate the variance of the full sample estimate.

Exhibit 8-59 presents the order of the data blocks on the child data file.

Exhibit 8-59. Child catalog blocks

Order	Block
1	Identifiers
2	FMS data
3	School demographic data
4	Round 5 school demographic data
5	Round 5 child demographic composites
	<i>Weights</i>
6	Round 5 child weight full sample
7	Panel 245 child weight full sample
8	Panel 45 child weight full sample
9	Panel 1245 child weight full sample
10	Panel 12345 child weight full sample
11	Round 5 parent weight full sample
12	Panel 245 parent weight full sample
13	Panel 45 parent weight full sample
14	Panel 1245 parent weight full sample
15	Panel 12345 parent weight full sample
16	Round 5 CPT weight full sample
	<i>Flags</i>
17	Round 5 child assessment data status flags
18	Round 5 parent interview data status flags
19	Round 5 teacher questionnaire A data status flags
20	Round 5 teacher questionnaire B data status flags
21	Round 5 teacher questionnaire C data status flags
22	Round 5 school fact sheet data status flags
23	Round 5 school administrator questionnaire data status flags
24	Round 5 facilities checklist data status flags
25	Round 5 special education questionnaire A data status flags
26	Round 5 special education questionnaire B data status flags
27	Round 5 student record abstract data status flags
28	Round 5 parent data edit flags and error flags
	<i>Scores</i>
29	Round 1 recalibrated child assessment scores
30	Round 2 recalibrated child assessment scores
31	Round 3 recalibrated child assessment scores
32	Round 4 recalibrated child assessment scores
33	Round 5 recalibrated child assessment scores
34	Round 5 child assessment science scores
35	Round 5 child assessment SDQ scores
36	Round 5 teacher questionnaire C scores

See note at end of exhibit.

Exhibit 8-59. Child catalog blocks—Continued

Order	Block
	<i>Composite Variables</i>
37	Round 5 child composites
38	Round 1 parent composites
39	Round 5 parent composites
40	Cross-round Parent composites
41	Round 5 teacher questionnaire A composites
42	Round 5 teacher questionnaire C composite
43	Round 5 school fact sheet composites
44	Round 5 school administrator questionnaire composites
45	Round 5 student record abstract composites
	<i>Direct Assessment Data</i>
46	Round 5 child assessment data
	<i>Parent Interview Data</i>
47	Round 5 parent interview data
	<i>Teacher Questionnaire Data</i>
48	Round 5 teacher questionnaire A data
49	Round 5 teacher questionnaire B data
50	Round 5 teacher questionnaire C data
	<i>School Data</i>
51	Round 5 school fact sheet data
52	Round 5 school administrator data
53	Round 5 facilities checklist data
	<i>Special Education Data</i>
54	Round 5 special education questionnaire A data
55	Round 5 special education questionnaire B data
	<i>Student Record Abstract Data</i>
56	Round 5 student record abstract data
	<i>Taylor Series Sampling Strata And Primary Sampling Units (PSU)</i>
57	Round 5 child Taylor series sampling strata and PSU
58	Panel 245 child Taylor series sampling strata and PSU
59	Panel 45 child Taylor series sampling strata and PSU
60	Panel 1245 child Taylor series sampling strata and PSU
61	Panel 12345 child Taylor series sampling strata and PSU
62	Round 5 parent Taylor series sampling strata and PSU
63	Panel 245 parent Taylor series sampling strata and PSU
64	Panel 45 parent Taylor series sampling strata and PSU
65	Panel 1245 parent Taylor series sampling strata and PSU
66	Panel 12345 parent Taylor series sampling strata and PSU
67	Round 5 CPT Taylor series sampling strata and PSU
	<i>Replicate Weights</i>
68	Round 5 child replicate weights
69	Panel 245 child replicate weights
70	Panel 45 child replicate weights
71	Panel 1245 child replicate weights
72	Panel 12345 child replicate weights
73	Round 5 parent replicate weights

See note at end of exhibit.

Exhibit 8-59. Child catalog blocks—Continued

Order	Block
74	Panel 245 parent replicate weights
75	Panel 45 parent replicate weights
76	Panel 1245 parent replicate weights
77	Panel 12345 parent replicate weights
78	Round 5 CPT replicate 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.

8.7.1 Third Grade Child Catalog Topical Variable Groupings

The variables within the third grade child catalog are organized into topical categories to help locate and identify desired variables in the ECB. These categories are identified in exhibit 8-60. The first column of the exhibit describes the topic, the second column lists the variable identifiers, which generally indicate the source of the data (e.g., the parent interview, the teacher questionnaires). The third column is a description of the topic. The last column is the Field ID, which is used to search the data file for the topics. Use the keywords in the Variable Identifier column to search for variables while using the ECB Narrow or Expand function (see sections 8.3.1.2 and 8.3.1.3). If (Variable Name) appears under the variable in the Variable Identifier column, you must search using the variable name. The field ID can be used by the Go function in the Variable List.

Exhibit 8-60 is only for your convenience in searching for variables and includes only variables that belong to certain topics. It does not reflect the contents of the entire child catalog.

Exhibit 8-60. Selected child catalog topical variable groupings

Topic	Variable identifier (Variable Name)	Description	Field ID
Identifiers	CHILDDID – U5_S_ID	Parent/Child identification numbers School identification number Teacher identification number	1 - 15
Status flags	C5COGDAT - U5SRABS	Child assessment status indicators (e.g., cognitive data available), parent data available, teacher data available, school data available, special education data available, student record abstract data available	57 - 71
Recalibrated Cognitive Skills Reading Third grade	C5R2R (Variable Name)	Reading routing test scores, IRT scale scores, T-scores, proficiency scores, and probability scores	172 - 184
Recalibrated Cognitive Skills Math Third grade	C5R2M (Variable Name)	Math routing test scores, IRT scale scores, T-scores, proficiency scores, and probability scores	185 - 195
Cognitive Skills Science	C5SROUNR - C5EARTSC (Variable Name)	Science routing test score, IRT scale score, T-score, and item cluster scores	196 - 201
Cognitive Skills SDQ	C5 SDQ	SDQ scores	202 - 207
Social Skills and Knowledge	T5 Approaches	Six factor scores from the social rating scale in teacher form C: approaches to learning, self-control, interpersonal,	212
	T5 Self	externalizing problem behaviors,	213
	T5 Interpersonal	internalizing problem behaviors, and	214
Teacher Rating	T5 Externalizing	combination of self control and interpersonal	215
	T5 Internalizing		216
	T5 Combination self/interpersonal		217
Composite/ Derived Variables	R5DOBMM – R5AGE C5HEIGHT – C5BMI (Variable Name)	Related to child (e.g., child DOB, child race, child gender, etc.)	38 - 43 218 - 220
	P1FIRKDG – IF3_ICAT (Variable Name)	Related to parent interview (e.g., parent occupation, parent highest education, etc.)	221 - 277
	A5CLSZ – T5GLVL (Variable Name)	Related to teacher information (e.g., class size, percentage minority in class, etc.)	278 - 283

See note at end of exhibit.

Exhibit 8-60. Selected child catalog topical variable groupings—Continued

Topic	Variable identifier	Description	Field ID
Composite/Derived Variables— Continued	L5SCHBMM - L5SCHEYY S5SCTYP – S5SCINC (Variable Name)	Related to school information (e.g., school year, school type, school enrollment, percent minority, etc.)	284 - 289 290 - 298
	U5RIEP (Variable Name)	Related to student record abstract: IEP record	299
Child Assessment: Closing section	C5 ACQ	Includes: child's grade level, height and weight; location, environment and accommodations of assessment, etc.	302 - 321
Parent Interview: Family Structure	P5 Roster P5 FSQ	Includes: relationship of people to child, age of people, gender of people, and race/ethnicity of people, round joined, round left	331 - 581 582 - 746
Parent Interview: Parent Involvement with Child's School	P5 PIQ	Includes: parents' contact with school, parent attendance in school related activities, barriers to attendance to school related activities	747 - 806
Parent Interview: Home Environment, Activities and Cognitive Stimulation	P5 HEQ	Includes: attending activities in the community (e.g., museums, zoos, athletic events), home computer access and use, participation in extra-curricular activities (e.g., music lessons, craft classes), information on community crime and safety, mealtimes and bedtimes	807 - 917
Parent Interview: Critical Family Processes	P5 CFQ	Includes: helps from family/friends, information on parents' relationship to each other	918 - 924
Parent Interview: Child Care	P5 CCQ	Includes: child care arrangements (i.e., relative care, non-relative care, center care)	925 - 989
Parent Interview: Discipline, Warmth and Emotional Supportiveness	P5 DWQ	Includes: parents' interaction with child and discipline strategies (e.g., spanking, time-out)	990 - 1010
Parent Interview: Non-resident Parent Question	P5 NRQ	Includes: information on non-resident parent and custody arrangements	1011 - 1083
Parent Interview: Child's Health and Well-Being	P5 CHQ	Includes: information on dentist and doctor, updates information on indicators for developmental difficulties	1084 - 1177

See note at end of exhibit.

Exhibit 8-60. Selected child catalog topical variable groupings—Continued

Topic	Variable identifier	Description	Field ID
Parent Interview: Parent's Psychological Well-Being and Health	P5 PPQ	Includes: parent depression, parent health problems	1178 - 1192
Parent Interview: Food Security	P5 FDQ	Includes: parents' food sufficiency in the household	1193 - 1210
Parent Interview: Parent Education	P5 PEQ	Includes: parents' education level, job training, currently attending courses	1211 - 1224
Parent Interview: Parent Employment	P5 EMQ	Includes: parents' employment status	1225 - 1256
Parent Interview: Welfare and Other Public Transfers	P5 WPQ	Includes: family's utilization of TANF or food stamps, free/reduced breakfast or lunch	1257 - 1282
Parent Interview: Parents Income and Assets	P5 PAQ	Includes: household income	1283 - 1288
Parent Interview: Child Mobility and Plans to Move	P5 CMQ	Includes: number of places child lived, reasons for moving, mode, language and location of interview	1289 - 1302
Teacher Questionnaire A: Classroom and Student Characteristics	A5 Q1 – A5 Q20	Includes: Grades, class age and race-ethnic composition, gender composition, percent limited English proficient, languages in classroom	1306 - 1386
Teacher Questionnaire A: Instructional Activities and Curricular Focus	A5 Q21 – A5 Q45	Includes: Instructional approach, presence of centers/learning areas, curriculum content, teacher's aide, physical resources (e.g., texts, equipment, materials for art and music).	1387 - 1490
Teacher Questionnaire A: Language Arts Instructional Information	A5 Q46 – A5 Q52	Includes: Instructional time spent on specific aspects of language arts	1491 - 1529
Teacher Questionnaire A: Mathematics Instructional Information	A5 Q53 – A5 Q54	Includes: Instructional time spent on specific aspects of mathematics	1530 - 1556

See note at end of exhibit.

Exhibit 8-60. Selected child catalog topical variable groupings—Continued

Topic	Variable identifier	Description	Field ID
Teacher Questionnaire A: Science Instructional Information	A5 Q55 – A5 Q56	Includes: Instructional time spent on specific aspects of science	1557 - 1577
Teacher Questionnaire A: Social Studies Instructional Information	A5 Q57	Includes: Instructional time spent on specific aspects of social studies	1578 - 1594
Teacher Questionnaire A: Student Evaluation	A5 Q58 – A5 Q64	Includes: Criteria and standards for evaluating students	1595 - 1617
Teacher Questionnaire A: Parent Involvement	A5 Q65 – A5 Q66	Includes: Communication with parents and percent of parents involved in school offered activities	1618 - 1621
Teacher Questionnaire B: School and Staff Activities	B5 Q1 – B5 Q2	Includes: School-related, staff development and training activities	1625 - 1636
Teacher Questionnaire B: Teacher Views	B5 Q3 – B5 Q9	Includes: Views on teaching, school climate, and environment	1637 - 1669
Teacher Questionnaire B: Teacher Background	B5 Q10 – B5 Q25	Includes: Gender, birth year, race-ethnicity, experience, education, certification, and time spent on preparation	1670 - 1710
School Fact Sheet	L5 Q1 – L5 Q10	Includes: Grade levels, school type, school year, student attendance records	1803 - 1847
School Questionnaire: School Characteristics	S5 Q1 – S5 Q16	Includes: Attendance, enrollment, breakfast/lunch programs, federal programs	1848 - 1898
School Questionnaire: School Facilities and Resources	S5 Q17 – S5 Q21	Includes: Facilities (e.g., cafeteria, music room, auditorium) and computer equipment/service	1899 - 1919
School Questionnaire: Community Characteristics and School Safety	S5 Q22 – S5 Q24	Includes: Community type, community safety, school safety	1920 - 1939

See note at end of exhibit.

Exhibit 8-60. Selected child catalog topical variable groupings—Continued

Topic	Variable identifier	Description	Field ID
School Questionnaire: School Policies and Practices	S5 Q25 – S5 Q30	Includes: Uniform code, testing and assessments, retention	1940 - 1972
School Questionnaire: School-Family- Community Connections	S5 Q31 – S5 Q32	Includes: Programs and services for families, parent involvement in school	1973 - 1982
School Questionnaire: School Programs for Special Populations	S5 Q33 – S5 Q39	Includes: Limited English proficient, special education, gifted and talented	1983 - 2002
School Questionnaire: Staffing and Teacher Characteristics	S5 Q40 – S5 Q41	Includes: School staff list, staff's race- ethnic distribution	2003 - 2034
School Questionnaire: School Governance and Climate	S5 Q43– S5 Q44	Includes: Goals/objectives for teachers, school's climate in early grades	2037 - 2054
School Questionnaire: Principal Characteristics	S5 Q45 – S5 Q51	Includes: Principal's gender, race- ethnicity, birth year, experience and education	2055 - 2067
Facilities Checklist	K5 Q1 – K5 Q6	Includes: Measures of security, learning environment, happiness, and surrounding neighborhood	2071 - 2113
Special Education Teacher	D5 Q1 – D5 Q15	Includes: Special education teacher information such as gender, years teaching, educational background and certification, and working with IEP children	2114 - 2166
Special Education Student	E5 Q1 – E5 Q24	Includes: information specific to the study child, such as disability category, Individualized Education Plan goals, special education services received, involvement in a general education program, accommodations used and evaluations performed	2167 - 2272
Student Record Abstract	U5 Q1 – U5 Q13	Includes: Student attendance and Individualized Education Plan record	2273 - 2313

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.

8.7.2 Child Catalog Predefined Taglists

There is one predefined taglist provided with the child catalog that can be used for extracting data. This taglist, third grade, is defined in exhibit 8-61. The required variables, identified with “Yes” appearing in the Required field column, are the same as those listed in exhibit 8-58. The limiting variables, identified as “Yes” in the Limiting field column, are the same as those described in section 8.7.3.

8.7.3 Child Catalog Limiting Fields

The limiting fields for the child catalog include (a) whether the child attended kindergarten for the first time in the base year (base year variable), (b) child’s grade level in spring 2002, and (c) child’s school type (public vs. private) in spring 2002. These limiting fields are specific to the child catalog and allow codes within each variable to be included or excluded from the extraction depending on the selection indicator. For example, the user can select “No” for private schools and for cases with “not ascertained” or “not applicable” data on school type if he or she would like to create a data file that only includes public school third-graders. The selection indicator will be either a “Yes” or “No” to specify whether the variable code should be included or excluded, respectively.

The limiting fields feature for the catalog allows the user to create a subset of cases based on the settings of the Select column in the Extract Specifications (exhibit 8-62) window. The default setting is all “Yes” in the Select column meaning that all records will be present in the extract file. To exclude records in a particular category of a variable, change the “Yes” associated with that code in the Select column to “No” by double-clicking on it. At least one of the codes for each limiting variable must be selected as “Yes” or no records will be extracted for analysis. For example, an extract using the default specifications above will include children from all grades. To restrict the extract to records for third-graders only, double-click the “Yes” next to all the classes other than third grade to change the other codes to “No.”

Exhibit 8-61. Third grade predefined taglist—child catalog

Field Name	Field Label	Required field	Limiting field
CHILDID	CHILD IDENTIFICATION NUMBER	Yes	No
S5_ID	SPRING 2002 SCHOOL IDENTIFICATION NUMBER	Yes	No
T5_ID	SPRING 2002 TEACHR IDENTIFICATION NUMBER	Yes	No
D5T_ID	SPRING 2002 SPECIAL ED TEACHER ID NUMBER	Yes	No
R5REGION	R5 CENSUS REGION	Yes	No
R5URBAN	R5 LOCATION TYPE - 7 CATEGORIES	Yes	No
R5LOCALE	R5 LOCATION TYPE - 8 CATEGORIES	Yes	No
R5GENDER	R5 CHILD COMPOSITE GENDER	Yes	No
R5RACE	R5 CHILD COMPOSITE RACE	Yes	No
R5AGE	R5 COMPOSITE CHILD ASSESSMENT AGE(MNTHS)	Yes	No
C5CW0	C5 CHILD WEIGHT FULL SAMPLE	No	No
C245CW0	C2C4C5 CHILD PANEL WEIGHT FULL SAMPLE	No	No
C45CW0	C4C5 CHILD PANEL WEIGHT FULL SAMPLE	No	No
C1_5FC0	C1245 CHILD PANEL WEIGHT FULL SAMPLE	No	No
C1_5SC0	C12345 CHILD PANEL WEIGHT FULL SAMPLE	No	No
C5PW0	C5 PARENT WEIGHT FULL SAMPLE	No	No
C245PW0	C2C4C5 PARENT PANEL WEIGHT FULL SAMPLE	No	No
C45PW0	C4C5 PARENT PANEL WEIGHT FULL SAMPLE	No	No
C1_5FP0	C1245 PARENT PANEL WEIGHT FULL SAMPLE	No	No
C1_5SP0	C12345 PARENT PANEL WEIGHT FULL SAMPLE	No	No
C5CPTW0	C5 CHILD-PARENT-TCHER WEIGHT FULL SAMPLE	No	No
C5R2RSCL	C5 RC2 READING IRT SCALE SCORE	No	No
C5R2RPF	C5 RC2 READING HIGHEST PROF LVL MASTERED	No	No
C5R2RPB1	C5 RC2 PROB1 - LETTER RECOGNITION	No	No
C5R2RPB2	C5 RC2 PROB2 - BEGINNING SOUNDS	No	No
C5R2RPB3	C5 RC2 PROB3 - ENDING SOUNDS	No	No
C5R2RPB4	C5 RC2 PROB4 - SIGHT WORDS	No	No
C5R2RPB5	C5 RC2 PROB5 - WORD IN CONTEXT	No	No
C5R2RPB6	C5 RC2 PROB6 - LITERAL INFERENCE	No	No
C5R2RPB7	C5 RC2 PROB7 - EXTRAPOLATION	No	No
C5R2RPB8	C5 RC2 PROB8 - EVALUATION	No	No
C5R2RDEC	C5 RC2 DECODING SCORE GR3	No	No
C5R2MSCL	C5 RC2 MATH IRT SCALE SCORE	No	No
C5R2MPB2	C5 RC2 PROB2 - RELATIVE SIZE	No	No
C5R2MPB3	C5 RC2 PROB3 - ORDINALITY, SEQUENCE	No	No
C5R2MPB4	C5 RC2 PROB4 - ADD/SUBTRACT	No	No
C5R2MPB5	C5 RC2 PROB5 - MULTIPLY/DIVIDE	No	No
C5R2MPB6	C5 RC2 PROB6 - PLACE VALUE	No	No
C5R2MPB7	C5 RC2 PROB7 - RATE & MEASUREMENT	No	No
C5SSCALE	C5 SCIENCE IRT SCALE SCORE	No	No
C5LIFESC	C5 LIFE SCIENCE GR3	No	No
C5PHYSSC	C5 PHYSICAL SCIENCE GR3	No	No

See note at end of exhibit.

Exhibit 8-61. Third grade predefined taglist—child catalog—Continued

Field Name	Field Label	Required field	Limiting field
C5EARTSC	C5 EARTH SCIENCE GR3	No	No
C5SDQRDC	C5 SDQ PRCVD INTEREST/COMPETENCE-READING	No	No
C5SDQMTC	C5 SDQ PRCVD INTEREST/COMPETENCE-MATH	No	No
C5SDQSBC	C5 SDQ PRCVD INTEREST/COMPETENCE-ALL SBJ	No	No
C5SDQPRC	C5 SDQ PRCVD INTEREST/COMPETENCE-PEER RL	No	No
C5SDQEXT	C5 SDQ EXTERNALIZING PROBLEMS	No	No
C5SDQINT	C5 SDQ INTERNALIZING PROBLEMS	No	No
P1FIRKDG	P1 FIRST-TIME KINDERGARTNER	Yes	Yes
W3SESQ5	W3 CATEGORICAL SES MEASURE	Yes	No
T5GLVL	T5 GRADE LEVEL OF CHILD	Yes	Yes
S5SCTYP	S5 SCHOOL TYPE FROM THE SFS	Yes	No
S5ENRLS	S5 TOTAL SCHOOL ENROLLMENT	Yes	No
S5PUPRI	S5 PUBLIC OR PRIVATE SCHOOL	Yes	Yes

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.

Exhibit 8-62 is the child catalog Extract Specifications window.

Exhibit 8-62. Child catalog Extract Specifications window

Child Extract Specifications

The extract will include data marked "Yes" in the "Select" column below.
To customize the extraction, change the "Select" column settings.

Variable	Label	Code	Select
P1FIRKDG	P1 FIRST-TIME KINDERGARTNER	YES	Yes
		NO	Yes
		DON'T KNOW	Yes
		NOT ASCERTAINED	Yes
			Yes
T5GLVL	T5 GRADE LEVEL OF CHILD	KINDERGARTEN	Yes
		FIRST GRADE	Yes
		SECOND GRADE	Yes
		THIRD GRADE	Yes
		FOURTH GRADE	Yes
		FIFTH GRADE	Yes
		UNGRADED CLASSROOM	Yes
		NOT ASCERTAINED	Yes
S5PUPRI	S5 PUBLIC OR PRIVATE SCHOOL	PUBLIC	Yes
		PRIVATE	Yes

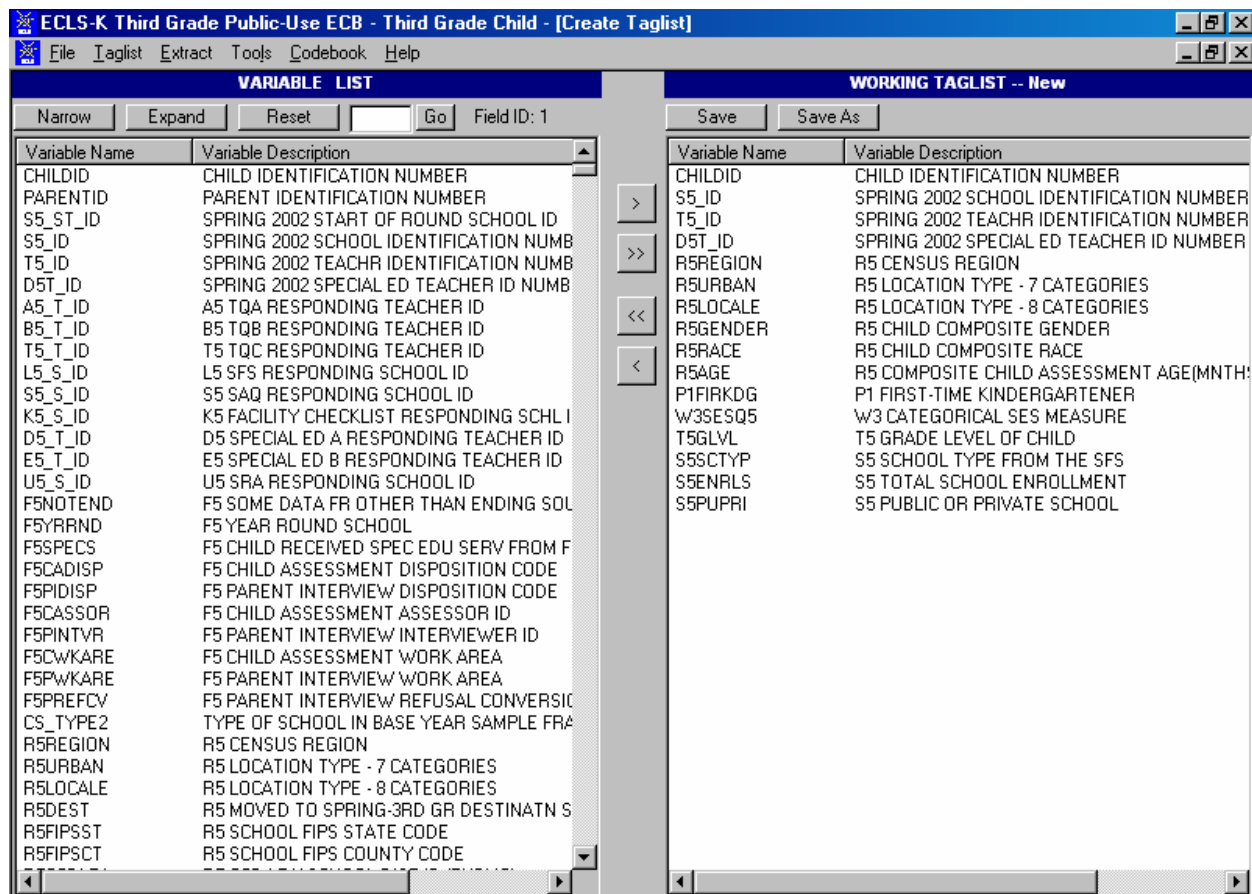
CD-ROM Drive: D: Set as permanent drive

OK Cancel

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

Exhibit 8-63 is the main screen of the Third Grade Public-Use ECB.

Exhibit 8-63. Third Grade Public-Use ECB main screen



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

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
		Replication method (WesVar, SUDAAN or AM)			Taylor Series method (SUDAAN, Stata, SAS or AM)		DEFT
		ID	Replicate weights	Jackknife method	Sample design	Nesting variables	(Average root design effect)
Spring-first grade/spring-third grade longitudinal	C45CW0 C45PW0	CHILDID PARENTID	C45CW1 – C45CW90 C45PW1 – C45PW90	JK2 JK2	WR ¹ WR	C45CSTR C45CPSU C45PSTR C45PPSU	1.815
Spring-kindergarten/spring-first grade/spring-third longitudinal	C245CW0 C245PW0	CHILDID PARENTID	C245CW1 – C245CW90 C245PW1 – C245PW90	JK2 JK2	WR WR	C245CSTR C245CPSU C245PSTR C245PPSU	1.791
Fall-kindergarten/spring-kindergarten/spring- first grade/spring- third grade longitudinal	C1_5FC0 C1_5FP0	CHILDID PARENTID	C1_5FC1 – C1_5FC90 C1_5FP1 – C1_5FP90	JK2 JK2	WR WR	C15FCSTR C15FCPSU C15FPSTR C15FPSPSU	1.749
All five round longitudinal	C1_5SC0 C1_5SP0	CHILDID PARENTID	C1_5SC1 – C1_5SC40 C1_5SP1 – C1_5SP40	JK2 JK2	WR WR	C15SCSTR C15SCPSU C15SPSTR C15SPPSU	1.635

¹WR = 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	P5HOMECEM, P5COMPWK	10,121	41.82	0.648	0.491	1.321	1.746
Had Internet access	P5HOMECEM, P5INTACC	9,883	86.06	0.431	0.348	1.237	1.529
Used computer 1-2 times/week for homework	P5HOMECEM, 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

¹Design SE is the standard error under the ECLS-K sample design. For an explanation of this statistic, see chapter 4, section 4.8.

²SRS SE is the standard error assuming simple random sample. For an explanation of this statistic, see chapter 4, section 4.8.

³DEFT is the root design effect. For an explanation of this statistic, see chapter 4, section 4.8.

⁴DEFF 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	P5HOMECEM, P5COMPWK	9,848	41.99	0.690	0.497	1.387	1.924
Had Internet access	P5HOMECEM, P5INTACC	9,615	86.08	0.458	0.353	1.297	1.683
Used computer 1-2 times/week for homework	P5HOMECEM, 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	T5RTLNG	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

¹Design SE is the standard error under the ECLS-K sample design. For an explanation of this statistic, see chapter 4, section 4.8.

²SRS SE is the standard error assuming simple random sample. For an explanation of this statistic, see chapter 4, section 4.8.

³DEFT is the root design effect. For an explanation of this statistic, see chapter 4, section 4.8.

⁴DEFF 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

¹Design SE is the standard error under the ECLS-K sample design. For an explanation of this statistic, see chapter 4, section 4.8.

²SRS SE is the standard error assuming simple random sample. For an explanation of this statistic, see chapter 4, section 4.8.

³DEFT is the root design effect. For an explanation of this statistic, see chapter 4, section 4.8.

⁴DEFF 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	T5CONTR0- T4CONTR0	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	T5CONTR0- T2CONTR0	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	P5HOMECEM, P5COMPWK	2,836	43.98	1.011	0.932	1.085	1.177
Had Internet access	P5HOMECEM, P5INTACC	2,768	86.19	0.980	0.656	1.494	2.231
Used computer 1-2 times/week for homework	P5HOMECEM, 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

¹Design SE is the standard error under the ECLS-K sample design. For an explanation of this statistic, see chapter 4, section 4.8.

²SRS SE is the standard error assuming simple random sample. For an explanation of this statistic, see chapter 4, section 4.8.

³DEFT is the root design effect. For an explanation of this statistic, see chapter 4, section 4.8.

⁴DEFF 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.

¹DEFT is the root design effect. For an explanation of this statistic, see section 4.8.

²DEFF 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.