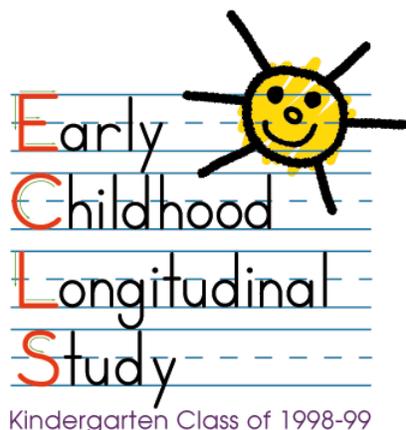

NATIONAL CENTER FOR EDUCATION STATISTICS



USER'S MANUAL FOR THE ECLS-K FIRST GRADE PUBLIC-USE DATA FILES AND ELECTRONIC CODE BOOK NCES 2002-135

Prepared by:

Westat
Rockville, Maryland
Karen Tourangeau
John Burke
Thanh Le
Siu Wan
Margaret Weant
Christine Nord
Nancy Vaden-Kiernan
Elizabeth Bissett
Richard Dulaney
Alison Fields
Lauren Byrne
Ismael Flores-Cervantes
Jean Fowler

Educational Testing Service
Princeton, New Jersey
Judith Pollack
Donald Rock

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

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

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

U.S. Department of Education
Office of Educational Research and Improvement

NCES 2002-135

U.S. Department of Education

Rod Paige

Secretary

Office of Educational Research and Improvement

Grover J. Whitehurst

Assistant Secretary

National Center for Education Statistics

Gary W. Phillips

Acting Commissioner

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

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

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

National Center for Education Statistics
Office of Educational Research and Improvement
U.S. Department of Education
555 New Jersey Avenue, N.W.
Washington, DC 20208-5574

January 2002

The NCES World Wide Web Home Page is: <http://nces.ed.gov>

The NCES World Wide Web Electronic Catalog is: <http://nces.ed.gov/pubsearch/index.asp>

TABLE OF CONTENTS

<u>Chapter</u>		<u>Page</u>
	BASE YEAR ERRATA AND COMPOSITES	1
	1 Departed Round 1	4
	2 Departed Round 2	4
1	INTRODUCTION	1-1
	1.1 Background.....	1-3
	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.....	1-8
	1.5 Contents of Manual.....	1-10
2	DESCRIPTION OF DATA COLLECTION INSTRUMENTS	2-1
	2.1 Purposes of the Fall-First Grade Enhancement	2-3
	2.2 Direct Child Assessments	2-4
	2.3 Parent Interview	2-7
	2.4 Spring-First Grade Teacher Questionnaires.....	2-9
	2.4.1 Content of the Academic Rating Scale	2-12
	2.4.2 Teacher Social Rating Scale.....	2-14
	2.5 Special Education Teacher Questionnaires.....	2-15
	2.6 Adaptive Behavior Scale	2-16
	2.7 School Administrator Questionnaire.....	2-17
	2.8 School Facilities Checklist.....	2-18
	2.9 School Records Abstract Form	2-20
	REFERENCES	2-21
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.....	3-2
	3.1.2 Item Response Theory Scale Scores	3-2

TABLE OF CONTENTS (continued)

<u>Chapter</u>		<u>Page</u>
	3.1.3 Standardized Scores (T-Scores)	3-4
	3.1.4 Proficiency Scores.....	3-6
	3.1.5 Choosing the Appropriate Score for Analysis.....	3-13
	3.1.6 Reliabilities	3-15
	3.2 Indirect Cognitive Assessment	3-17
	3.2.1 Comparison to Direct Cognitive Assessment.....	3-18
	3.3 Rasch Scores Available for the Academic Rating Scale.....	3-19
	3.4 Oral Language Development Scale	3-25
	3.5 Social Rating Scale	3-26
	3.5.1 Teacher Social Rating Scale.....	3-28
	3.5.2 Parent Social Rating Scale	3-29
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.....	4-10
	4.4 Calculation and Use of Sample Weights	4-11
	4.4.1 Types of Sample Weights	4-12
	4.4.2 Weighting Procedures	4-17
	4.4.3 Computation of Base Year Child Weights.....	4-18
	4.4.4 Computation of Fall-First Grade Child Weights.....	4-19
	4.4.5 Computation of Spring-First Grade Child Weights	4-21
	4.4.6 Types of Weights	4-27
	4.4.7 Replicate Weights	4-29
	4.4.8 Characteristics of Sample Weights	4-29
	4.5 Variance Estimation.....	4-30
	4.5.1 Paired Jackknife Replication Method	4-31
	4.5.2 Taylor Series Method.....	4-33
	4.5.3 Specifications for Computing Standard Errors	4-34

TABLE OF CONTENTS (continued)

<u>Chapter</u>		<u>Page</u>
4.6	Design Effects.....	4-36
	4.6.1 Use of Design Effects.....	4-37
	4.6.2 Average Design Effects for the ECLS-K.....	4-38
5	DATA COLLECTION METHODS AND RESPONSE RATES.....	5-1
	5.1 Data Collection Methods.....	5-1
	5.2 Field Staff Training.....	5-2
	5.2.1 Fall-First Grade Field Staff Training.....	5-2
	5.2.2 Spring-First Grade Training.....	5-3
	5.3 Fall-First Grade Data Collection.....	5-4
	5.3.1 Advance Mailings.....	5-5
	5.3.2 Preassessment Contact.....	5-5
	5.3.3 Conducting the Direct Child Assessment.....	5-5
	5.3.4 Conducting the Parent Interview.....	5-7
	5.3.5 Conducting Data Collection on Children Who Withdrew from Their Previous Round School.....	5-8
	5.4 Fall-First Grade Completion Rates.....	5-11
	5.5 Spring-First Grade Data Collection.....	5-16
	5.5.1 Advance Contact with Respondents.....	5-16
	5.5.2 Preassessment Contact.....	5-16
	5.5.3 Conducting Direct Child Assessments.....	5-19
	5.5.4 Conducting the Parent Interview.....	5-21
	5.5.5 Teacher and School Data Collection.....	5-21
	5.5.6 Conducting Data Collection on Children Who Withdrew from Their Previous Round School.....	5-22
	5.6 Data Collection Quality Control.....	5-24
	5.6.1 Child Assessment Observations.....	5-25
	5.6.2 Parent Validations.....	5-26
	5.6.3 School Validations.....	5-26
	5.6.4 Quality Control of the Oral Language Development Scale Scoring.....	5-27
	5.6.5 Assessor Effects.....	5-28

TABLE OF CONTENTS (continued)

<u>Chapter</u>		<u>Page</u>
5.7	Spring-First Grade Completion Rates.....	5-30
	5.7.1 Students Sampled in Kindergarten.....	5-30
	5.7.2 Students Sampled in First Grade.....	5-40
	5.7.3 Spring-First Grade Completion Rates – All Children.....	5-41
6	DATA PREPARATION.....	6-1
	6.1 Coding and Editing Specifications for Computer-Assisted Interviewing.....	6-1
	6.1.1 Range Specifications.....	6-2
	6.1.2 Consistency Checks (Logical Edits).....	6-2
	6.1.3 Coding.....	6-2
	6.2 Coding and Editing Specifications for Hard-Copy Questionnaires....	6-6
	6.2.1 Receipt Control.....	6-6
	6.2.2 Coding.....	6-7
	6.2.3 Data Entry.....	6-8
	6.2.4 Data Editing.....	6-8
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.....	7-9
	7.4.1 Child Composite Variables.....	7-10
	7.4.2 Family and Household Composite Variables.....	7-17
	7.4.3 Teacher Composite Variables.....	7-28
	7.4.4 School and Class Composite Variables.....	7-30
	7.4.5 Student Record Abstract and Field Management System Composite Variables.....	7-34
	7.4.6 Base Year (Kindergarten Year) Composites on the First Grade File.....	7-37
	7.4.7 Parent Identifiers and Household Composition.....	7-41
	7.4.8 Supplemental Section in the Spring-First Grade Parent Interview.....	7-46
	7.4.9 Industry and Occupation Codes Used in ECLS-K.....	7-51

TABLE OF CONTENTS (continued)

<u>Chapter</u>		<u>Page</u>
	7.5 Relationship Among Composite Variables on the Child Catalog.....	7-55
	7.6 Children Who Changed Schools.....	7-55
	7.7 Merging Base Year School-Level Data with the First Grade Child-Level Data.....	7-57
	7.8 Composite Tables.....	7-63
	7.9 Masked Variables.....	7-94
8	PUBLIC-USE ELECTRONIC CODE BOOK.....	8-1
	8.1 Introduction.....	8-1
	8.1.1 Hardware/Software Requirements.....	8-2
	8.1.2 ECLS-K First Grade Public-Use Electronic Code Book Features.....	8-2
	8.2 Installing, Starting, and Exiting the ECLS-K First Grade Public-Use Electronic Code Book.....	8-3
	8.2.1 Installing the ECLS-K First Grade Public-Use Electronic Code Book Program on Your Personal Computer.....	8-3
	8.2.2 Starting the ECLS-K First Grade Public-Use Electronic Code Book.....	8-9
	8.2.3 Exiting the ECLS-K First Grade Public-Use Electronic Code Book.....	8-12
	8.2.4 Removing the Electronic Code Book Program From Your Personal Computer.....	8-13
	8.2.5 Title Bar.....	8-14
	8.2.6 Menu Bar.....	8-14
	8.2.7 Using Shortcut Keys to Navigate.....	8-15
	8.3 Variable List.....	8-15
	8.3.1 Searching the Code Book for Variables.....	8-18
	8.4 Working Taglist.....	8-24
	8.4.1 Opening A Taglist.....	8-25
	8.4.2 Adding Variables to the Working Taglist.....	8-26
	8.4.3 Removing Variables from the Working Taglist.....	8-28
	8.4.4 Saving Taglists.....	8-29
	8.4.5 Exporting Taglists.....	8-31
	8.4.6 Importing Taglists.....	8-33
	8.4.7 Using Predefined Taglists.....	8-34

TABLE OF CONTENTS (continued)

<u>Chapter</u>		<u>Page</u>
	8.4.8 Deleting Taglists	8-34
	8.4.9 Viewing Code Book and Variable Information	8-36
8.5	Extracting Data from the Electronic Code Book	8-43
	8.5.1 Reviewing the Extract Specifications	8-50
	8.5.2 Repairing and Compacting the Database	8-50
8.6	Child Catalog	8-51
	8.6.1 First Grade Child Catalog Topical Variable Groupings.....	8-53
	8.6.2 Child Catalog Predefined Taglists.....	8-64
	8.6.3 Child Catalog Limiting Fields.....	8-64
8.7	Base Year School Catalog	8-70
	8.7.1 Base Year School Catalog Topical Variable Groupings.....	8-71
	8.7.2 Base Year School Catalog Predefined Taglists.....	8-72
	8.7.3 Base Year School Catalog Limiting Fields.....	8-72
8.8	Switching Catalogs	8-74
8.9	Menu Bar Descriptions	8-76
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 Child-Level Data.....	9-3
	9.4 K-1 Longitudinal Weights	9-4
	9.4.1 Type of K-1 Longitudinal Weights	9-4
	9.4.2 Weighting Procedures	9-6
	9.4.3 Characteristics of Longitudinal Weights.....	9-7
	9.4.4 Variance Estimation	9-8
	9.4.5 Design Effects	9-10

TABLE OF CONTENTS (continued)

List of Appendixes

<u>Appendix</u>		<u>Page</u>
A	PORTABLE DOCUMENT FORMAT (PDF) FILES CONTAINING THE FIRST GRADE INSTRUMENTS	
B	RECORD LAYOUT OF CHILD AND SCHOOL DATA FILES	
C	ECLS-K BASE YEAR PUBLIC-USE USER’S MANUAL	
D	BASE YEAR ERRATA AND COMPOSITES	

List of Tables

<u>Table</u>		
1-1	ECLS-K waves of data collection.....	1-2
2-1	Instruments used in the ECLS-K base year and fall- and spring-first grade....	2-2
2-2	Direct child assessment, by round of data collection.....	2-5
2-3	ECLS-K parent interview by major content topics and round of data collection.....	2-10
2-4	Teacher questionnaires.....	2-13
2-5	Special education teacher questionnaire topics in spring-kindergarten and spring-first grade	2-17
2-6	School administrator questionnaire, spring-kindergarten and spring-first grade	2-19
3-1	Direct cognitive assessment: number-right scores.....	3-3
3-2	Direct cognitive assessment: Item Response Theory scale scores.....	3-4
3-3	Direct cognitive assessment: standardized scores.....	3-5
3-4	Direct cognitive assessment: proficiency-level scores.....	3-9
3-5	Direct cognitive assessment: proficiency probability scores	3-11

TABLE OF CONTENTS (continued)

List of Tables (continued)

<u>Table</u>		<u>Page</u>
3-6	Direct cognitive assessment: print familiarity scores	3-12
3-7	Reliability of Item Response Theory-based scores.....	3-16
3-8	Reliability of routing test number correct (alpha coefficient).....	3-17
3-9	Split half reliability of item-cluster-based scores (proficiency-level scores)	3-17
3-10	Academic rating scale response scale	3-18
3-11	Person reliability for the Rasch-based score	3-19
3-12	First grade academic rating scale: variable names, descriptions, ranges, weighted means, and standard deviations	3-20
3-13	Kindergarten academic rating scale language and literacy item difficulties (arranged in order of difficulty)	3-21
3-14	Kindergarten academic rating scale mathematical thinking item difficulties (arranged in order of difficulty)	3-21
3-15	Kindergarten academic rating scale general knowledge item difficulties (arranged in order of difficulty)	3-21
3-16	Spring-first grade academic rating scale language and literacy item difficulties (arranged in order of difficulty).....	3-22
3-17	Spring-first grade academic rating scale mathematical thinking item difficulties (arranged in order of difficulty).....	3-22
3-18	Spring-first grade academic rating scale general knowledge item difficulties (arranged in order of difficulty).....	3-22
3-19	Kindergarten academic rating scale language and literacy standard errors	3-23
3-20	Kindergarten academic rating scale mathematical thinking standard errors ...	3-24
3-21	Kindergarten academic rating scale general knowledge standard errors	3-24

TABLE OF CONTENTS (continued)

List of Tables (continued)

<u>Table</u>		<u>Page</u>
3-22	Spring-first grade academic rating scale language and literacy standard errors	3-24
3-23	Spring-first grade academic rating scale mathematical thinking standard errors	3-25
3-24	Spring-first grade academic rating scale general knowledge standard errors	3-25
3-25	Oral language development scale: variable names, descriptions, ranges, means, and standard deviations.....	3-26
3-26	Social rating scale response scale	3-27
3-27	Split half reliability for the teacher social rating scale scores.....	3-27
3-28	Split half reliability for the parent social rating scale scores	3-27
3-29	Teacher social rating scores: variable names, descriptions, ranges, weighted means, and standard deviations	3-29
3-30	Parent social rating scores: variable names, descriptions, ranges, weighted means, and standard deviations	3-30
4-1	Distribution of the ECLS-K primary sampling unit sample by self-representing status, metropolitan statistical area status, and census region.....	4-2
4-2	Characteristics of the ECLS-K base year school sample	4-3
4-3	Characteristics of base year cooperating schools selected for fall-first grade	4-4
4-4	Characteristics of children subsampled for fall-first grade.....	4-6
4-5	Characteristics of children in spring-first grade sample, excluding freshened students.....	4-9
4-6	ECLS-K fall-first grade cross-sectional weights	4-12
4-7	ECLS-K spring-first grade cross-sectional weights.....	4-13

TABLE OF CONTENTS (continued)

List of Tables (continued)

<u>Table</u>		<u>Page</u>
4-8	ECLS-K first grade longitudinal (panel) weights	4-15
4-9	Distribution of originally sampled schools by number of children with nonzero weights by first grade sample weights	4-16
4-10	Number of children who were not assessed due to special situations.....	4-28
4-11	Characteristics of the first grade child-level weights.....	4-30
4-12	ECLS-K Taylor Series stratum and first-stage unit identifiers.....	4-34
4-13	Specifications for computing standard errors	4-35
4-14	ECLS-K, fall-first grade: standard errors and design effects for the full sample	4-41
4-15	ECLS-K, spring-first grade: standard errors and design effects for the full sample – Child and parent data	4-42
4-16	ECLS-K, panel: standard errors and design effects for the full sample.....	4-43
4-17	ECLS-K: median design effects for subgroups – child and parent data	4-44
5-1	Oral language development scale (OLDS) routing patterns in fall-first grade for previous round respondents.....	5-6
5-2	Oral language development scale (OLDS) routing results in fall-first grade ..	5-7
5-3	Number of children excluded from or accommodated in the fall-first grade assessment.....	5-8
5-4	Number of children who moved in fall-first grade by completion category....	5-10
5-5	ECLS-K fall-first grade: number of completed child-level cases and child-level completion rates, by school characteristic	5-13
5-6	ECLS-K fall-first grade: number of completed child-level cases and child-level completion rates, by child’s mover status.....	5-14
5-7	ECLS-K fall-first grade: number of completed child-level cases and child-level completion rates, by child characteristic.....	5-15

TABLE OF CONTENTS (continued)

List of Tables (continued)

<u>Table</u>		<u>Page</u>
5-8	Oral language development scale (OLDS) routing patterns in spring-first grade for previous round respondents.....	5-19
5-9	Oral language development scale (OLDS) routing results in spring-first grade	5-20
5-10	Number of children excluded from or accommodated in the spring-first grade assessment.....	5-20
5-11	Number of children who moved in spring-first grade by completion category.....	5-24
5-12	Intra-class correlations and estimated design effects for the spring-first grade direct cognitive tests (unweighted)	5-29
5-13	Components of variance for three-level model, including the interviewer effect for spring-first grade cognitive scores	5-30
5-14	ECLS-K spring-first grade: number of completed child-level cases and child-level completion rates, by school characteristic—children sampled in the base year	5-31
5-15	ECLS-K spring-first grade: number of completed child-level cases and child-level completion rates, by child’s mover status—children sampled in the base year	5-36
5-16	ECLS-K spring-first grade: number of completed child-level cases and child-level completion rates, by child characteristic—children sampled in the base year	5-38
5-17	ECLS-K spring-first grade: number of completed instruments and child-level completion rates for additional data collected—children sampled in the base year	5-40
5-18	ECLS-K spring-first grade: number of completed child-level cases and child-level completion rates for children sampled in first grade.....	5-42
5-19	ECLS-K spring-first grade: number of completed child-level cases and child-level completion rates, for children sampled in kindergarten and first grade, by survey instruments.....	5-43

TABLE OF CONTENTS (continued)

List of Tables (continued)

<u>Table</u>		<u>Page</u>
7-1	Levels of the detailed income range	7-19
7-2	Households asked to report income to the nearest \$1,000	7-19
7-3	Missing data for socioeconomic scale source variables.....	7-20
7-4	Imputed education variables	7-23
7-5	Imputed occupation variables	7-24
7-6	Imputed detailed income range	7-24
7-7	Imputed exact household income value	7-24
7-8	ECLS-K and census poverty thresholds for 1999	7-26
7-9	Pointers to parent figure questions.....	7-44
7-10	Supplemental section item matches to spring-first grade items.....	7-49
7-11	Case status and identification numbers for children not followed or located.....	7-56
7-12	Variables from the spring-kindergarten school administrator questionnaire that were included in the first grade new school administrator questionnaire but not the first grade returning school administrator questionnaire	7-59
7-13	Fall-first grade composite variables	7-64
7-14	Spring-first grade composite variables	7-69
7-15	Recoded and suppressed data on the ECLS-K First-Grade Year Public Use File	7-96
8-1	Add variables buttons	8-27
8-2	Remove variables buttons	8-28
8-3	Code book information	8-36
8-4	Navigation buttons.....	8-38

TABLE OF CONTENTS (continued)

List of Tables (continued)

<u>Table</u>		<u>Page</u>
8-5	Child catalog required variables	8-52
8-6	Child catalog blocks (As noted).....	8-54
8-7	Child catalog topical variable groupings (As noted)	8-57
8-8	Fall-first grade predefined taglist—child catalog	8-66
8-9	Spring-first grade predefined taglist—child catalog.....	8-67
8-10	First grade predefined taglist—child catalog.....	8-68
8-11	Base year school catalog required variables	8-70
8-12	Base year school catalog blocks	8-71
8-13	Base year school catalog topical variable groupings	8-73
9-1	ECLS-K: K-1 longitudinal weights	9-5
9-2	Characteristics of child-level K-1 longitudinal weights	9-7
9-3	ECLS-K Taylor Series stratum and first-stage unit identifiers	9-9
9-4	Specifications for computing standard errors	9-11
9-5	ECLS-K spring-kindergarten/spring-first grade panel: standard errors and design effects for the full sample	9-12
9-6	ECLS-K all-four round panel: standard errors and design effects for the full sample	9-13
9-7	ECLS-K panel: median design effects for subgroups.....	9-15
9-8	ECLS-K spring-kindergarten/fall-first grade and spring-kindergarten/ spring-first grade panels: standard errors and design effects of differences in scores for the full sample	9-16
9-9	ECLS-K spring-kindergarten/fall-first grade and spring-kindergarten/ spring-first grade panels: median design effects of differences in scores for subgroups	9-17

TABLE OF CONTENTS (continued)

List of Figures

<u>Figure</u>		<u>Page</u>
1-1	ECLS-K conceptual model	1-5
8-1	Windows Run screen	8-4
8-2	InstallShield Wizard	8-4
8-3	Welcome window	8-5
8-4	Choose Destination Location.....	8-6
8-5	Setup Status.....	8-7
8-6	InstallShield Wizard Complete	8-8
8-7a	Desktop icon	8-9
8-7b	Desktop screen—click start	8-10
8-8	First-time user dialog box	8-10
8-9	ECLS-K First Grade Electronic Code Book splash screen.....	8-11
8-10	Select Catalog screen	8-11
8-11	Main electronic code book screen.....	8-12
8-12	Exit screen.....	8-13
8-13	Save working taglist dialog box.....	8-13
8-14	Title Bar	8-14
8-15	Menu Bar	8-14
8-16	ECLS-K First Grade Electronic Code Book main screen.....	8-16
8-17	Variable List	8-17
8-18	“Go” button.....	8-19
8-19	Narrow Search Text dialog box	8-21

TABLE OF CONTENTS (continued)

List of Figures (continued)

<u>Figure</u>		<u>Page</u>
8-20	No Matches Found message	8-21
8-21	Example of narrowing a search.....	8-22
8-22	Expand Search Text dialog box	8-23
8-23	No Matches Found message	8-24
8-24	Electronic code book Working Taglist	8-25
8-25	Open Taglist dialog box.....	8-26
8-26	Add Taglist dialog box	8-28
8-27	Save Taglist As dialog box	8-30
8-28	Save Taglist As dialog box (#2).....	8-31
8-29	Pull-down menu to select Taglist Export.....	8-32
8-30	Export Taglist dialog box	8-32
8-31	Pull-down menu to select Taglist Import.....	8-33
8-32	Import Taglist dialog box	8-34
8-33	Delete Taglist selection.....	8-35
8-34	Delete Taglist confirmation window	8-35
8-35	Code book view	8-37
8-36	Printing status screen	8-38
8-37	Export code book selection screen.....	8-39
8-38	Variable Quick View	8-40
8-39	Add All Variables dialog box	8-41
8-40	View of the entire code book.....	8-42

TABLE OF CONTENTS (continued)

List of Figures (continued)

<u>Figure</u>		<u>Page</u>
8-41	Printer dialog box.....	8-42
8-42	Limiting fields dialog box.....	8-44
8-43	Save SAS program file dialog box.....	8-45
8-44	Save SAS program file location browse screen.....	8-45
8-45	Save SAS data file dialog box.....	8-46
8-46	Save SPSS program file dialog box.....	8-47
8-47	Save SPSS program file location browse screen.....	8-47
8-48	Save SPSS data file dialog box.....	8-48
8-49	Save STATA program file dialog box.....	8-49
8-50	Save STATA program file location browse screen.....	8-49
8-51	Save STATA data file dialog box.....	8-50
8-52	Repair database completed screen.....	8-51
8-53	Child catalog Extract Specifications window.....	8-65
8-54	Base year school catalog limiting fields.....	8-74
8-55	Save changes.....	8-75
8-56	Select Catalog screen.....	8-75

List of Exhibits

<u>Exhibit</u>		
5-1	Timeline of first grade data collection.....	5-1

BASE YEAR ERRATA AND COMPOSITES

Since the release of the base year data files and documentation, several errata have been identified. These errata are listed here for the base year public-use data files (NCES 2001-029e, June 2001). Errors in the data file are listed first, followed by errors in the documentation. Corrected versions of these base year variables are available in appendix D for all children who participated in the base year. Users who wish to use the corrected base year variables should refer to appendix D. In addition, appendix D contains several base year composites that were created after the base year data files were released. These composites are described in chapter 7, section 7.4.6.

A. The following errors were identified in the ECLS-K Base Year Public-Use *Child* data files (child.dat)

1. WKPOVRTY

There are 383 cases that were coded as at or above the poverty level (WKPOVRTY=2) that should have been coded as below the poverty level (WKPOVRTY=1). There are two additional cases (that should have coded above the poverty level, but instead were coded below poverty level).

There were 1179 cases that had a round 1 interview, but no round 2 interview. For these cases the poverty composite was created by imputing for income and using round 1 household size. All 1179 of these cases were coded as at or above the poverty level, when 383 of the 1179 cases should have been coded as below the poverty level. Thus, in addition to the 3855 cases already coded as below the poverty level, 383 more cases should have been included in this category.

Two cases (0060013C and 1114006C) were coded as below poverty level instead of above due to an error in the income cut-off point in the program.

For a list of the 385 cases with an incorrect poverty status by child ID please refer to the file `POVERTY_IDS.TXT` on the NCES website at <http://nces.ed.gov/ecls/kindergarten/errata.htm>.

NCES has re-released the corrected base year poverty composite with the first grade data; the corrected composite is called WKPOV_R on the first grade data file.

2. KURBAN

In the public-use file, a seven-category locale code was collapsed to a three-category locale code (KURBAN). The labels of the three-category locale variable correctly read as follows:

1. Central City (Large City and Mid-Size City)
2. Urban Fringe and Large Town (Urban Fringe should include Large City Urban Fringe and Mid-Size City Urban Fringe.)
3. Small Town and Rural

However, the categories were erroneously collapsed as:

1. Large Central City and its urban fringe
2. Mid-Size Central City and its urban fringe and Large Town
3. Small Town and Rural

NCES has re-released the corrected base year school locale code with the first grade data.

3. C2ASMTST

12 children were incorrectly coded as nonrespondents instead of not assessed in the variable C2ASMTST (C2 CHILD ASSESSMENT STATUS). Instead of 4=CHILD W/ DISABILITY, NOT ASSESSED, they were coded 5=NONRESPONDENT. This error only affects this variable. These 12 children were all correctly assigned spring kindergarten (round 2) weights and they were included appropriately in table 5-12 in chapter 5 of the base year user's manual.

The child IDs (CHILDDID) for these children are as follows:

0105005C
0105008C
0105009C
0105012C
0105013C
0105018C
0105019C
0105020C
0105021C
0105023C
0105024C
2121014C

4. Academic Rating Scale (ARS) Scores

An error was identified in the base year ARS scores. Specifically, the fall and spring base year ARS scores use slightly different metrics and, therefore, are not directly comparable. The specifics of the problem are described below.

The ARS scale scores are interval level scores and may be used in analyses requiring that level of measurement. However, *the use of* gain scores (subtracting the fall score from the spring score) *is* not recommended. In order to represent the student's score within a 1-5 range, an arithmetic transformation using the mean and standard deviation was applied to the scores. Fall and spring scores were analyzed separately; therefore, the metric is slightly different due to differences in the range of the scores at the two time points. In other words, an increase of 0.1 on the fall scale is not the same as an increase of 0.1 on the spring scale. Further, because some children performed above grade level, their ratings on the ARS may be at the maximum score. Consequently, it is not possible to estimate how far beyond the skills, knowledge, and behaviors assessed on the ARS these students might have achieved. In short, any estimate of fall to spring growth will be underestimated for the highest achieving students. Although gain scores are not recommended, covariance models may be used (with the caveat that there are some ceiling and floor effects).

In the first grade data files, corrected scores for the kindergartners are included. These scores were calibrated using a combined calibration of fall and spring kindergarten ratings. Therefore, the unit for the fall and spring kindergarten scores are the same. The problem with estimating growth for students at the ceiling remain. The standard error of measurement (SEM) for the scores is provided in the User's Manual. The first grade scores are based on different items and should not be used to compare growth.

5. IF_INC

Income was collected in the round 2 spring kindergarten data collection. For those households that did not provide this information, an income value was imputed. Imputed income values are indicated on a file using a variable IF_INC. There is an error in the variable IF_INC on the file. There are 3379 cases on the file coded as "0" (not imputed) that should have been coded as "1" (imputed).

The file IMPUTE_IDS.TXT contains the child IDs of the cases with the erroneous flag on the NCES website at <http://nces.ed.gov/ecls/kindergarten/errata.htm>.

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

	Case (Child ID)	Has parent data belonging too child with ID...	Should have parent data that is currently in child ID...
1.	0162001C	3056007C	No parent data (R2 nonrespondent)
2.	0192001C	0192003C	No parent data (R2 nonrespondent)
3.	0192003C	0192015C	0192001C
4.	0192015C	None	0192003C
5.	3056007C	None	0162001C

There are two recommended strategies for correcting this error. Users may go ahead and use the cases since the effect on any analysis will be trivial. The other option would be to discard these five cases during analysis. There would not be much benefit in reassigning the correct data to the appropriate child, as parent-level weights were not created for cases 0192015C and 3056007C for round 2 as they were considered as having missing parent data during the process of the creation of round 2 and base year longitudinal parent weights.

- The format (value) labels for four variables P2AGREE1 (P2 NRQ264 AGREEMENT W/ BIOLOGICAL FATHER), P2AGREE2 (P2 NRQ264 AGREEMENT W/ BIOLOGICAL MOTHER), P2AGREE3 (P2 NRQ264 AGREEMENT W/ ADOPTIVE FATHER), and P4AGREE4 (P2 NRQ264 AGREEMENT W/ ADOPTIVE MOTHER) are incorrect.

Currently, each of the variables has value labels ranging in valued from 1 to 4 with labels (1=biological father, 2=biological mother, 3=adoptive father, 4=adoptive mother). The variables should all have had the labels 1=yes and 2=no. All cases have a value of either 1 or 2 for each of these four variables; there are no cases with values of 3 or 4.

For example, for the variable P2AGREE1 (P2 NRQ264 AGREEMENT W/ BIOLOGICAL FATHER) the frequency distribution is as follows:

Code in restricted-use file	Frequency (# of cases)	Correct Code
1 (biological father)	95	1 (yes)
2 (biological mother)	36	2 (no)
3 (adoptive father)	0	---
4 (adoptive mother)	0	---

The same situation applies to all the P2 NRQ264 variables.

8. The variable labels for the question series P2 PRRDP_* (P2 ROSTER ROUND DEPARTED - PERSON *) (where * ranges from 1 to 17) are in error. The labels currently read

1 Joined Round 1

2 Joined Round 2

The labels should read:

1 Departed Round 1

2 Departed Round 2

1. INTRODUCTION

This manual provides guidance and documentation for users of the first grade 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 and use of the data file.

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

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

Two more waves of data were collected in the fall and spring of the 1999-2000 school year when most, but not all, of the base year children were in first grade.² The fall-first grade data collection was limited to a 30 percent subsample of schools³ (see table 1-1). It was a design enhancement whose goal was to enable researchers to measure the extent of summer learning loss and the factors that contribute to such loss and to better disentangle school and home effects on children's learning. The spring-first grade data collection, on the full sample, was part of the original study design and can be used to measure annual school progress and to describe the first grade learning environment of children in the study. All children assessed during the base year were eligible to be assessed in the spring-first grade data collection regardless of whether they repeated kindergarten, were promoted to first grade, or were moved

¹ The term "first grade" is used throughout this document to refer to the data collections that took place in the 1999-2000 school year, at which time most of the sampled children—but not all of them—were in first grade.

² Though the majority of base year children were in first grade during the 1999-2000 school year, about 5 percent of the sampled children were retained in kindergarten and a handful of others were in second grade during the 1999-2000 school year.

³ Approximately 27 percent of the base year students who were eligible to participate in Year 2 attended the 30 percent subsample of schools.

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

Table 1-1.—ECLS-K waves of data collection¹

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

¹ See section 1.3 for a description of the study components.

² Fall data collection consisted of a 30 percent sample of schools containing approximately 27 percent of the base year students eligible to participate in Year 2.

The final two waves of data collection that are currently planned are scheduled for spring 2002 and spring 2004 when most of the study children will be in the third grade and fifth grade, respectively.

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

⁴ Their addition is referred to as "freshening" the sample. See chapter 4 for more detail on the freshening process.

The ECLS-K is part of a longitudinal studies program comprising two cohorts—a kindergarten cohort and a birth cohort. The birth cohort (ECLS-B) will follow a national sample of children, born in the year 2001, from birth through first grade. The ECLS-B will focus on the characteristics of children and their families that influence children’s first experiences with the demands of formal school, as well as children’s early health care and in- and out-of-home experiences. Together these cohorts will provide the range and breadth of data required to more fully describe and understand children’s health and early learning, development, and education experiences.

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

1.1 Background

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

Efforts to expand and improve early education will benefit from insights gained through analyses of data from the large-scale, nationally representative, ECLS-K data, and the study’s longitudinal design. The ECLS-K database contains information about the types of school programs in which children participate, the services they receive, and repeated measures of the children’s cognitive skills and knowledge. The ECLS-K database also contains measures of children’s physical health and growth, social

development, and emotional well-being, along with information on family background and the educational quality of their home environments.

As a study of early achievement, the ECLS-K allows researchers to examine how children's progress is affected by such factors as placement in high or low ability groups, receipt of special services or remedial instruction, grade retention, and frequent changes in schools attended because of family moves. Data on these early school experiences are collected as they occur, with the exception of their experiences before kindergarten, which are collected retrospectively. This produces a more accurate measurement of these antecedent factors and enables stronger causal inferences to be made about their relationship to later academic progress.

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

Data collected during the kindergarten year serve as baseline measures to examine how schooling shapes later individual development. The longitudinal nature of the study enables researchers to study children's cognitive, social, and emotional growth and to relate trajectories of change to variations in children's experiences in kindergarten and the early grades.

The first grade data collection has two distinct purposes. As noted previously, the fall-first grade data collection can be used to study the extent of children's summer learning loss and the factors contributing to differential loss, and to better disentangle school and home influences on children's learning. To this end, the fall-first grade data collection gathered detailed information about how the children spent their summers and about various types of cognitive stimulation that they received over the summer either from their parents or from other sources such as camps, summer school, tutoring, or enrichment programs. In addition, detailed information was gathered about the children's child care arrangements during the summer. Thus, the fall-first grade data collection can also be used to study how kindergartners' child care arrangements change during the summer.

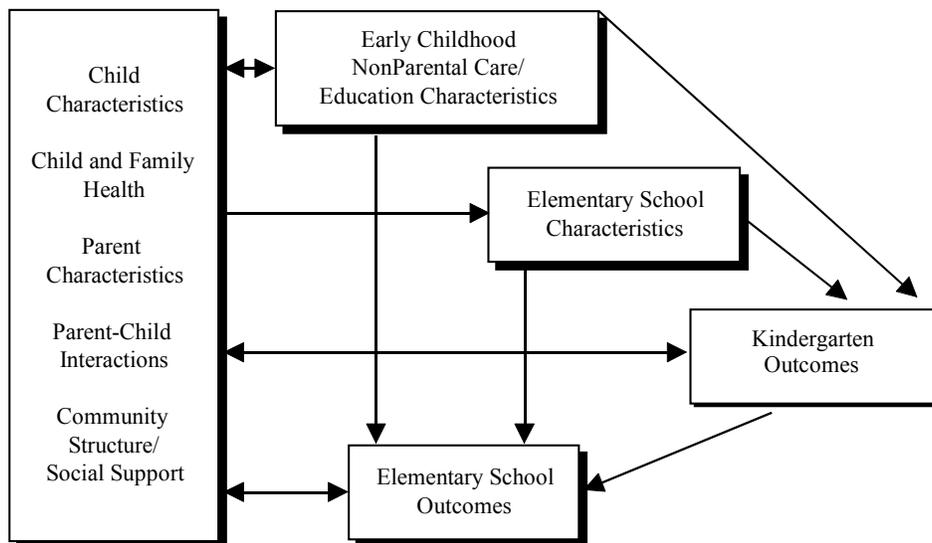
The spring-first grade data collection can be used to describe the diversity of first graders and the classrooms and schools they attend. It can also be used to study children's academic gains in the

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

1.2 Conceptual Model

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

Figure 1-1.—ECLS-K conceptual model



1.3 Study Components

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

- **Children** are asked to participate in various activities to measure the extent to which they exhibit those abilities and skills deemed important to success in school. They are asked to participate in activities designed to measure important cognitive (e.g., general knowledge, literacy, and quantitative skills) and noncognitive (e.g., fine motor and gross motor coordination, socioemotional) skills and knowledge. Most measures of a child's cognitive skills are obtained through an untimed one-on-one assessment of the child. Beginning with the third grade data collection, children will report on their own experiences in and out of school. Children are assessed in each round of data collection.
- **Parents/guardians** are an important source of information about the families of the children selected for the study and about themselves. They are asked to provide key information about their children, especially during the first years of the study. Parents are one of the important sources of information about children's development at school entry and their experiences both with family members and others. Information is collected from parents each time children are assessed using computer-assisted interviews (CAIs). Information is collected from parents/guardians in each round of data collection.
- **Teachers**, like parents, represent a valuable source of information on themselves, the children in their classrooms, and the children's learning environment (i.e., the classroom). Teachers are not only asked to provide information about their own backgrounds, teaching practices, and experience but they are also called on to provide information on the classroom setting for the sampled children they teach and to evaluate each sampled child on a number of critical cognitive and noncognitive dimensions. Teachers complete self-administered questionnaires each time children are assessed, with the exception of the fall-first grade data collection.
- **School administrators**, or their designees, are asked to provide information on the physical, organizational, and fiscal characteristics of their schools, and on the schools' learning environment and programs. Special attention is paid to the instructional philosophy of the school and its expectations for students. Information is collected from school administrators via self-administered questionnaires during spring data collection.

1.4 ECLS-K Data Files

The ECLS-K data are released in restricted-use and public-use versions. A brief overview of the differences between the restricted-use and public-use files is provided here, followed by a description of the base-year and first-grade data files that are either currently available or will be available shortly.

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

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

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

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

The modifications that are implemented to avoid identification of schools, teachers, parents, and children do not affect the overall data quality and most researchers should be able to find all that they need in the public-use files. While very few of the variables are suppressed, there are a few users who might require the restricted files. Those researchers examining certain rare subpopulations such as the

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

1.4.2 Overview of Available Data Files

Several different ECLS-K base-year and first-grade data files are available or will shortly be available for use by analysts:

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

The child file contains all the data collected from or about the children, including data from the child assessments, and from their teachers, parents, and schools. Analysts who wish to obtain descriptive information about U.S. kindergarten students or their families, or who want to examine relationships involving children and families, children and teachers, or children and schools, should make use of the child file. Analysts wishing to obtain descriptive information about the population of kindergarten teachers in the United States, or to study relationships involving teachers as the principal focus of attention, should use the teacher file. Analysts who want to obtain descriptive information about public and private schools that contain kindergarten classes, or who want to examine relationships among school characteristics, should make use of the school file. These child-, teacher-, and school-level files are available in public-use and restricted-use versions. For more information on these files, refer to the ECLS-K Base Year Public-Use User's Manual (NCES 2001-029), February 2001 or the ECLS-K Restricted-Use Base Year User's Manual (NCES 2000-097), August 2000.

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

The **special education file** is a child-based file that contains information on 784 children identified as receiving special education or related services in kindergarten. Special education teachers were asked to complete two questionnaires designed to collect information about their professional background and experience and about the nature of the special education program and special education services provided to each of the sampled children receiving services. It is only available as a restricted-use file. For more information about this file, see the User's Manual for the ECLS-K Base Year Restricted-Use Special Education Data Files and Electronic Code Book (NCES 2001-015)

The **student record abstract file** contains information from school records about children's school enrollment and attendance; Individualized Education Plan (IEP) and disability status; and home and school language. The student record abstract form was completed by school staff after the end of the school year. This file is useful in providing additional predictors and correlates of children's transitions to kindergarten and later progress in school. This file is only available as a restricted-use file. For more information about this file, see the User's Manual for the ECLS-K Base Year Restricted-Use Student Record Abstract Data Files and Electronic Code Book (NCES 2001-0016).

The purpose of the **Head Start Verification Study** was two-fold: (1) to identify which of the children reported by either their parents or their schools as having attended Head Start the year prior to kindergarten did indeed attend a Head Start program and (2) to evaluate the process of identifying Head Start participation through parent and school reports, and provide further information on the actual process of verifying these reports. The Head Start Verification file contains information collected during the verification process. This file is a restricted-use file. For more information about this file, see the User's Manual for the ECLS-K Base Year Restricted-Use Head Start Data Files and Electronic Code Book (NCES 2001-025).

- **ECLS-K first grade restricted- and public-use data files.** The first grade data (fall and spring) are available only as a child-level file. The file includes all data collected from or about the children and their schools including data from the child assessments and from their parents, teacher, and schools. First grade teacher and school files are not being released because the sample of teachers and schools is not nationally representative of first grade teachers and schools with first grades. Analysts who wish to examine children's experiences in first grade and the influence of their classroom or school characteristics on their first grade experiences should use the first grade file.

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

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

1.5 Contents of Manual

This manual provides documentation for users of the first grade public-use data of the ECLS-K. The manual contains information about the data collection instruments (chapter 2) and the psychometric properties of these instruments (chapter 3). It describes the ECLS-K sample design and weighting procedures (chapter 4); data collection procedures and response rates (chapter 5); and data processing procedures (chapter 6). In addition, this manual shows how the public-use first grade data file is structured, provides definitions of composite variables (chapter 7), explains how to use the Electronic Code Book (chapter 8) and describes how to use and merge the base year and first grade files (chapter 9). The Electronic Code Book contains unweighted frequencies for all variables. Because this manual focuses on the first grade data collection, minimal information is provided about the base year data. Users who wish to learn more about the base year data collection should refer to the ECLS-K Base Year Public-Use User's Manual (NCES 2001-029), February 2001 or the ECLS-K Restricted-Use Base Year User's Manual (NCES 2000-097), August 2000. Additional information about the ECLS program can be found on the World Wide Web at <http://nces.ed.gov/ecls>.

2. DESCRIPTION OF DATA COLLECTION INSTRUMENTS

This chapter describes the survey instruments used during the first grade data collection of the Early Childhood Longitudinal Study-Kindergarten Class of 1998-99 (ECLS-K). Table 2-1 lists all the instruments used during the first grade data collection. The instrumentation for the base year is also shown. Similarities and differences between the first grade and base year instruments are highlighted throughout this chapter.

The ECLS-K first grade data collection occurred in the fall and spring of the 1999-2000 school year. In the fall of 1999, data were collected by computer-assisted interviews (CAIs) with parents and by direct child assessments. Unlike the base year and spring-first grade data collections, no teacher or other school questionnaires were administered in the fall. As noted in chapter 1, the fall-first grade data collection was a design enhancement intended to examine changes in children's learning during the summer months. The fall instrumentation reflects this special focus.¹ In the spring of 2000, data were collected by computer-assisted parent interviews, teacher questionnaires, and direct child assessments. In addition, data were gathered from school administrators, student record abstracts, and through a facilities checklist. Because about five percent of the sampled children were retained in kindergarten, separate teacher questionnaires were developed for kindergarten and first grade teachers. Teachers of sampled students promoted to second grade received the first grade teacher questionnaires.

The Head Start verification study and the salary and benefits questionnaire administered in the spring of the base year were not repeated during the first grade data collection.² A copy of the first grade data collection instruments, except for the direct child assessment, the social rating scale (SRS)³ in the parent interview and teacher questionnaire, and the adaptive behavior scale (ABS) are available on the CD-ROM as appendix A.⁴ These latter measures contain copyright protected materials and agreements with the test publishers that restrict their distribution.

¹ See section 2.1 for more information about the purposes of the fall first-grade data collection.

² Users wanting information about the Head Start verification study should refer to the ECLS-K Base Year Restricted-Use Head Start Data Files and Electronic Code Book (NCES 2001-025), which includes a user's manual with detailed information about the verification process. Those wanting information about the salary and benefits data files should refer to the User's Manual for the ECLS-K Base Year Restricted-Use Salary and Benefits Data Files and Electronic Code Book (NCES 2001-014).

³ Adapted with permission from *Elementary Scale A ("How Often?")*, F.M. Gresham and S.N. Elliott. (1990). American Guidance Service, Inc.

⁴ Lambert, N., Nihira, K., and Leland, H. *Adaptive Behavior Scale-Second Edition*. (1993). The American Association on Mental Retardation.

Table 2-1.—Instruments used in the ECLS-K base year and fall- and spring-first grade

1998-99 school year		1999-2000 school year	
Fall-kindergarten	Spring-kindergarten	Fall-first grade ¹	Spring-first grade
Parent Interview	Parent Interview	Parent Interview	Parent Interview
Child Assessment	Child Assessment	Child Assessment	Child Assessment
Teacher Questionnaire-Part A	Teacher Questionnaire-Part A		First Grade Teacher Questionnaire - Part A
Teacher Questionnaire-Part B	Teacher Questionnaire-Part B		First Grade Teacher Questionnaire – Part B
Teacher Questionnaire-Part C	Teacher Questionnaire-Part C		First Grade Teacher Questionnaire – Part C
			Kindergarten Teacher Questionnaire - Part A
			Kindergarten Teacher Questionnaire - Part B
			Kindergarten Teacher Questionnaire - Part C
	Special Education Teacher Questionnaire-Part A		Special Education Teacher Questionnaire-Part A
	Special Education Teacher Questionnaire-Part B		Special Education Teacher Questionnaire-Part B
	Adaptive Behavior Scale		Adaptive Behavior Scale
	School Administrator Questionnaire		School Administrator Questionnaire - New Schools
			School Administrator Questionnaire - Returning Schools
	Student Record Abstract		Student Record Abstract
	School Facilities Checklist		School Facilities Checklist
	Salary and Benefits ² Questionnaire		
	Head Start Verification ³		

¹ The fall-first grade data collection consisted of a 30 percent subsample of the study schools. See section 2.1 for information about the purposes of the fall-first grade data collection.

² The salary and benefits questionnaire collected information on the base salary, merit pay, and benefit pay of teachers and principals. It was completed by the school or district business administrator or by a private school administrator or headmaster.

³ The Head Start Verification Study verified parent and school reports of children’s Head Start participation by matching information on the name and location of the Head Start facilities the children were reported to have attended against a database of Head Start centers. For each match, the center was contacted to confirm that the child had attended the center in the year before kindergarten.

2.1 Purposes of the Fall-First Grade Enhancement

The fall-first grade data collection was limited to a 30 percent subsample of schools. Approximately 27 percent of the base year children who were eligible to be interviewed in Year 2 attended these schools. Nearly 5,300 study children in these schools were assessed and their parents interviewed. The fall-first grade data collection was designed to study an important aspect of children's learning experience—summer learning. The existing research on summer learning shows conflicting findings about whether children experience decline or growth in learning over the summer as compared with learning during the school year (Entwisle and Alexander, 1992; Heyns, 1978). There are also conflicting findings regarding how children's learning of various subject areas varies differentially over the summer months.

Factors that have been related to learning rates over the summer include school poverty level, family socioeconomic status, year-round schooling, the child's attendance at summer school, the child's participation in structured summer programs, and reading materials and activities in the home (Entwisle and Alexander, 1992; Entwisle and Alexander, 1994; Cooper et al., 1996; Heyns, 1978). Although there are many common findings across studies (e.g., that learning loss over the summer is often greater for disadvantaged than for advantaged students), more research is needed to examine the effects of multiple factors on children's summer learning.

There are many other factors that have rarely been included or not included at all in the summer learning research. For example, although comparisons between black and white student achievement gains, controlling for other background factors, are frequently made in the literature, such comparisons do not appear to extend to students from other racial and ethnic groups, such as Asian Americans and Hispanic students. Limited English proficiency (LEP) students have also not been included in the extant literature. In addition, few studies have included children from rural or suburban areas.

A number of variables have been identified as possible contributors to children's rate of learning during the summer months. Some of the variables that have not been studied extensively in relation to summer learning include: family structure and composition; the use of tutoring and educational support from parents; the start and end dates of summer school programs, with the possibility that

summer school programs ending closer to the start of the following school year may lead to higher achievement gains; and the use of home computers during the summer.

The ECLS-K fall-first grade data collection includes not only factors that have been studied in previous research but also those that have not yet been studied. It provides an opportunity to examine children's summer learning on a national sample of children with a wide range of characteristics represented.

2.2 Direct Child Assessments

One-on-one direct child assessments were administered using CAI in the fall and spring of the 1999-2000 school year. The children were assessed regardless of whether they were retained in kindergarten, promoted to first grade, moved ahead to second grade, or were new to the sample (freshened students). The assessments took about one hour to administer. Table 2-2 displays the major domains measured during the direct child assessments from all four rounds of data collection. As in the base year, the first grade assessments consisted of cognitive and physical components. The fall-first grade assessments, conducted with a subsample of study children, can be compared to the spring-kindergarten assessments obtained during the base year to determine children's growth over the summer months and can also be linked to the spring-first grade assessments to study the children's gain during the first grade academic year. The spring-first grade assessments can be linked to the base year assessments conducted in the fall of 1998 and the spring of 1999.

Language Screener. As in the base year, the assessment began by verifying the child's name and administering a set of warm-up exercises similar in form to the items used to administer the cognitive component. Prior to administering the cognitive assessment battery, a language-screening assessment (the OLDS), was administered to those children identified from their school records (or by their teacher, if no school records were available) as coming from a language minority background (meaning that their primary home language was not English) and who had not passed the OLDS in any of the previous rounds of interviews.⁵ This screening test was used to determine if a child was able to understand and respond to the cognitive assessment items in English. If a child did not pass the language

⁵ Students who were part of the freshened sample followed the same procedures: If their school records indicated that they were a language minority student, they received the language screener.

Table 2-2.—Direct child assessment, by round of data collection

Direct child assessment	1998-99 school year		1999-2000 school year	
	Fall-kindergarten	Spring-kindergarten	Fall-first grade	Spring-first grade
Language screener (OLDS)	X	/	/	/
Reading (language and literacy)	X	X	X	X
Mathematical thinking	X	X	X	X
General knowledge (science and social studies)	X	X	X	X
Psychomotor	X			
Height and weight	X	X	X	X

NOTE: The columns to the right of each construct correspond to the round of administration. Rounds that included the construct are marked with an “X.” A “/” indicates that the OLDS was given to language minority students new in the spring, or who did not pass the cut score in the English version during the previous OLDS administration.

screener but spoke Spanish, he or she was administered a Spanish translated form of the mathematics assessment and an alternate form of the language screener, the Spanish version of the OLDS (Spanish OLDS). The Spanish OLDS that was administered is similar in content to the English OLDS and measures the same constructs. For further information on the language screener, please refer to the ECLS-K Base Year Public-Use User’s Manual (NCES 2001-029), February 2001, or the ECLS-K Restricted-Use Base Year User’s Manual (NCES 2000-097), August 2000.

Cognitive Components. The kindergarten-first grade cognitive assessment focused on three general areas of competence: (1) language use and literacy (reading); (2) mathematics; and (3) knowledge of the social and physical world, referred to as “general knowledge.” The assessment did not ask the children to write anything or to explain their reasoning; rather, they pointed to their answers or responded orally to complete the tasks. The assessment battery included the use of small easels with pictures, letters of the alphabet, words, short sentences, numbers, or number problems and a computer laptop for the assessor to enter children’s responses. The same two-stage cognitive assessment approach used in the base year was repeated in the first grade data collection. The purpose of the two-stage design was to maximize the accuracy of measurement and reduce administration time by using the children’s responses from the first stage to route the children to the appropriate level of difficulty in the second stage.⁶ With the exception of the reading assessment domain, the same mathematics and general knowledge two-stage

⁶ For details on the two-stage assessment design, see the ECLS-K Base Year Public-Use User’s Manual (NCES 2001-029), February 2001, or the ECLS-K Restricted-Use Base Year User’s Manual (NCES 2000-097), August 2000.

assessment batteries were re-administered in first grade. Analysis of the reading scores from spring-kindergarten showed a higher than expected number of respondents scoring near the ceiling. Therefore, to eliminate the possibility of ceiling effects, the number of reading items was increased by adding more difficult vocabulary words and text. The mathematical thinking and general knowledge assessments showed no such ceiling effects and so these assessments were not modified.

Language and Literacy. The language and literacy (reading) assessment included questions designed to measure basic skills (print familiarity, letter recognition, beginning and ending sounds, rhyming sounds, word recognition), vocabulary (receptive vocabulary), and comprehension (listening comprehension, words in context). Comprehension items were targeted to measure skills in initial understanding, developing interpretation, personal reflection, and demonstrating critical stance.

The reading assessment contained five proficiency levels. These five levels reflected a progression of skills and knowledge; if a child had mastered one of the higher levels, he or she was very likely to have passed the items that comprised the earlier levels as well. These five levels were as follows: (1) identifying upper- and lower-case letters of the alphabet by name; (2) associating letters with sounds at the beginning of words; (3) associating letters with sounds at the end of words; (4) recognizing common words by sight; and (5) reading words in context.

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

The items in the mathematics assessment could also be grouped into five proficiency levels, though the math clusters were less homogeneous in content than the reading clusters. The clusters of math items included the following: (1) identifying some one-digit numerals, recognizing geometric shapes, and one-to-one counting up to ten objects; (2) reading all one-digit numerals, counting beyond ten, recognizing a sequence of patterns, and using nonstandard units of length to compare objects; (3) reading two-digit numerals, recognizing the next number in a sequence, identifying the ordinal position of an

object, and solving a simple word problem; (4) solving simple addition and subtraction problems; and (5) solving simple multiplication and division problems and recognizing more complex number patterns.

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

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

Physical Components. In the fall of the base year there were two parts to the physical component of the child assessment, psychomotor and anthropometric. The psychomotor component was not included in subsequent rounds. The anthropometric component consisted of recording the children's height and weight in order to measure their physical growth and development. A Shorr Board (for measuring height) and a digital bathroom scale were used to obtain the height and weight measurements, which were recorded on a height and weight recording form.

2.3 Parent Interview

The first grade parent interviews were conducted using a computer-assisted interview (CAI). The parent interviews were conducted primarily in English, but provisions were made to interview

parents who spoke other languages. Bilingual interviewers were trained to conduct the parent interview in either English or Spanish. In fall-first grade if the interview was conducted in Spanish, the interviewer used a hard-copy questionnaire and then entered the respondent's answers into the CAI program. In spring-first grade, a Spanish CAI instrument was used when needed. The spring-first grade interview was also translated into Hmong and Mandarin. Such interviews were conducted with paper and pencil rather than CAI.

Fall Parent Interview. Most families interviewed during fall-first grade had been interviewed in the fall and/or spring of the base year. Typically the respondent for the fall-first grade parent interview was the mother of the child; however, the respondent could be a father, stepparent, adoptive parent, foster parent, grandparent, another relative, or a nonrelative guardian. The respondent had to be knowledgeable about the child's care and education, be 18 years of age or older, and be living in the household with the child. Respondents for the parent interview were selected according to the following order of preference:

1. Respondent from the previous round;
2. The child's mother;
3. Another parent or guardian; and
4. Another household member.

The fall interview was shorter and more narrowly focused than the parent interviews conducted during the base year and during spring-first grade due to its focus on summer learning experiences. It included sections on children's summer activities, including vacations; attendance at summer school, school enrichment programs, and summer camp; receipt of tutoring; and participation in special activities such as music, dance, or swimming lessons. It also obtained information on the types of activities parents engaged in with their children during the summer; children's summer child care arrangements; and the availability of community resources such as recreation centers, community pools and parks, boys' or girls' clubs, and libraries.

Spring Parent Interview. The majority of parents participating in the spring-first grade data collection were interviewed in the fall and/or spring of the base year. However, the sample was freshened to include children who had not attended kindergarten in the United States during the 1998-99 school year (see chapter 5). The order of preference for the respondent to the parent interview was the same as

described for the fall-first grade interview: (1) the respondent from the previous round (if there was one), (2) the child's mother, (3) another parent or guardian, or (4) some other household member.

The parent interview for the spring-first grade data collection was extensive and asked questions covering first grade school experiences, child care, parent characteristics, and family health. Table 2-3 provides an overview of the topics covered in the base year and first grade data collections. As can be seen in the table, key topics such as family structure, parental involvement in school, and the child's home environment and cognitive activities are covered in most rounds. Other topics, such as parent income, employment, and education, are measured at least once in each school year. Although the general content areas are similar across the questionnaires, the items were updated, where appropriate, to reflect the first grade context. For example, in spring-first grade, the series of questions asking, "In a typical week, how often do you or any other family member do the following things with the child," the activity "practice reading, writing or working with numbers" was added to the list of activities. Similarly, in spring-first grade questions about whether the child was tutored on a regular basis and in what subjects were added to the questionnaire.

Parents or guardians of children added to the sample through freshening or who were nonrespondents in the prior rounds in spring-first grade were asked to complete a supplementary section that contained key items asked in previous rounds. The supplementary section included questions about the children's child care arrangements in the year prior to kindergarten, whether they had attended Head Start in the year before kindergarten, their birth weight, and whether they had been born prematurely. In addition, parents were asked whether the children had ever received therapy services or taken part in programs for children with disabilities and, if so, the types of services or programs they had participated in. Finally, information about the children's backgrounds were collected including questions about the primary language spoken in the home, the country in which the child was born, and whether the children's mothers had worked for pay outside the home between when the children were born and the start of kindergarten.

2.4 Spring-First Grade Teacher Questionnaires

During spring-first grade data collection, each teacher received a self-administered questionnaire consisting of three distinct parts. The first section, part A, asked about the teacher's class

Table 2-3.—ECLS-K parent interview by major content topics and round of data collection

Parent questionnaire	1998-99 school year		1999-2000 school year	
	Fall-kindergarten	Spring-kindergarten	Fall-first grade	Spring-first grade
Family structure	X	X	X	X
Demographics	X	X	X	X
Household roster	X	X	X	X
Marital status	X	X	X	X
Immigration status		X		X
Primary language(s) spoken in home	X	/	/	/
Parent's involvement with child's school		X	X	X
Child care	X		X	X
Current arrangements with relatives	X		X	X
Current arrangements with nonrelatives	X		X	X
Current arrangements with centers	X		X	X
Head Start attendance year before kindergarten	X	/	/	/
Child care arrangements year before kindergarten	X	/	/	/
Child's health and well-being	X	X		/
Birth weight	X	/	/	/
Physical functioning	X	/	/	/
Services for children with special needs	X	/	/	/
Social skills rating	X	X		X
Home environment and cognitive activities	X	X	X	X
Frequency of literacy activities	X	X	X	X
Computer use		X	X	X
Television viewing		X	X	X
Summer activities and time use			X	
Parental educational expectations for child	X		X	X
Neighborhood		X	X	X
Safety		X		X
Resources (e.g., community center, library)			X	
Parent education	X	/	/	/
Parent employment	X			X
Parent income		X		X

Table 2-3.—ECLS-K parent interview by major content topics and round of data collection (continued)

Parent questionnaire	1998-99 school year		1999-2000 school year	
	Fall-kindergarten	Spring-kindergarten	Fall-first grade	Spring-first grade
Welfare and other public assistance use	X	X		X
Parent/child interaction		X		X
Parent discipline		X		
Parent health and emotional well-being		X		
Relationships and social support	X	X		
Marital satisfaction		X		
Background data	X	X		X
Mother's age at first birth	X			
Mother's age at child's birth				/
WIC benefits during pregnancy	X	/	/	/
Whether mother worked for pay between when child was born and time child entered kindergarten	X	/	/	/

NOTE: The columns to the right of each construct correspond to the round of questionnaire administration. Rounds that included the construct are marked with an "X." Content areas asked only of new parent respondents in each round are marked with a "/."

and classroom characteristics. It was designed to collect data about the composition and demographics of the children in the class and was completed only by teachers of sampled children, unlike the base year when it was completed by all kindergarten teachers in the school, regardless of whether they taught a sampled child. Part B addressed more specific questions on class organization, typical class activities, and evaluation methods, as well as teacher views on school readiness, school environment, and overall school climate. Background questions about the teacher were also included in this section. Teachers were asked to complete one copy of part C for each of the sampled children in their classrooms; in this part, teachers were asked to respond to 21 questions about the child's academic performance. The academic rating scale (ARS) gathered data on each sampled child's skills in areas of language and literacy, general knowledge, and mathematical thinking. Part C also included questions from the social rating scale (SRS) that collected data on five areas of children's social skills. The ARS and SRS are described in more detail in sections 2.4.1 and 2.4.2, respectively.

Two different versions of the teacher questionnaire were available. The first was for teachers of children who had made the transition to the first grade or any higher elementary school grade, and the second was for teachers of children who were repeating or attending the second year of kindergarten. Any

kindergarten-specific items were updated to first grade items in the questionnaire for first grade teachers. For example, in asking about the children's skills, first grade teachers were asked whether the children read first grade books independently with comprehension or read first grade books fluently. In kindergarten, the teachers were asked about the children reading "simple" books independently. Similarly, first grade teachers were asked whether the children compose stories with a clear beginning, middle, and end, while kindergarten teachers were asked whether the children compose simple stories. Teachers were asked about transition practices for children moving from kindergarten to first grade, while in the base year, the teachers were asked about transition practices to kindergarten. The questionnaire for kindergarten teachers combined elements from both the fall- and spring-base year teacher questionnaires.

Table 2-4 shows the overall structure of the base year and spring-first grade teacher questionnaires and the distribution of topics covered.

2.4.1 Content of the Academic Rating Scale

There are three scales of the ARS: language and literacy, general knowledge, and mathematics. Each of these is described below. The areas measured in the ARS overlap and augment what is measured in the direct cognitive assessment. The items were designed to ascertain the current skill levels, knowledge, and behaviors of the child in first grade based on the teacher's past observation and experience with the child.

- The **Language and Literacy** section of the ARS asks teachers to rate each child's proficiency in expressing ideas (one item), listening (one item), reading on grade level (four items), and writing (two items). In addition, teachers rate the child's computer literacy (one item).
- The **General Knowledge** section of the ARS asks teachers to rate each child's skills and knowledge in social studies (three items) and science (three items).
- In the **Mathematics** section, teachers rate each child on seven items that tap the following skills: understanding place values, making reasonable estimates of quantities, solving number problems, using various strategies, organizing and analyzing data (graphing), and measuring accurately.

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

Table 2-4.—Teacher questionnaires

Teacher questionnaire	1998-99 school year		1999-2000 school year	
	Fall-kindergarten	Spring-kindergarten	Spring-first grade (First grade teacher)	Spring-first grade (Kindergarten teacher)
Parts A and B				
Description of class—age, race-ethnicity, and sex distribution	X ^a		X ^a	X ^a
Class organization				
Types of activity/interest areas	X ^b	/	X ^a	X ^b
Types of materials/resources		X ^a	X ^a	X ^a
Instructional time in different subjects		X ^a	X ^a	X ^a
Child vs. teacher initiated activities	X ^b	X ^a	X ^a	X ^b
Classroom characteristics				
Children with special needs		X ^a	X ^a	X ^a
Classroom aides		X ^a	X ^a	X ^a
Class assignment and grouping		X ^a	X ^a	X ^a
Behavior of children in classroom	X ^a	X ^a	X ^a	X ^a
Parent involvement				
Share progress information with parents		X ^a	X ^a	X ^a
Professional development		X ^a	X ^b	X ^a
Teachers' evaluation and grading practices	X ^b	/	X ^a	X ^b
Teachers' views on school readiness	X ^b	/	X ^b	X ^b
Perceptions about school climate	X ^b	/	X ^b	X ^b
Perception of personal influence on policies and classroom planning	X ^b	/	X ^b	X ^b
Teacher demographic information	X ^b	/	X ^b	X ^b
Teacher experience and education	X ^b	/	X ^b	X ^b
Job satisfaction	X ^b	/	X ^b	X ^b
Transition to school activities	X ^b	/	X ^b	X ^b

Table 2-4.—Teacher questionnaires (continued)

Teacher questionnaire	1998-99 school year		1999-2000 school year	
	Fall-kindergarten	Spring-kindergarten	Spring-first grade (First grade teacher)	Spring-first grade (Kindergarten teacher)
Part C				
Indirect child cognitive evaluation by teacher	X	X	X	X
Language and literacy, mathematics, general knowledge	X	X	X	X
Social skills	X	X	X	X
Additional information on sampled child		X	X	X
Participation in special services and programs		X	X	X
Overall academic skills and physical activity levels		X	X	X
Reading group participation		X	X	X

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

^a Topic is in Teacher Questionnaire A.

^b Topic is in Teacher Questionnaire B.

2.4.2 Teacher Social Rating Scale

Teachers rated individual students’ social development on part C of the teacher questionnaire. These items are intended to measure approaches to learning, self-control, and interpersonal skills. The items were rated on a scale of one (Never) to four (Very often). Five scales are formed from these items. Three of the scales capture positive aspects of children’s development, and two represent problem behaviors. See chapter 3, section 3.5.1 for variable names, ranges, means, and standard deviations for these scales.

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

- The **Self-Control** Scale (Teacher SRS) has four items that indicate the child’s ability to control behavior by respecting the property rights of others, controlling temper, accepting peer ideas for group activities, and responding appropriately to pressure from peers.
- The five **Interpersonal Skills** Scale (Teacher SRS) has five items that rate the child’s skill in forming and maintaining friendships, getting along with people who are different, comforting or helping other children, expressing feelings, ideas and opinions in positive ways, and showing sensitivity to the feelings of others.

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

- **Externalizing Problem Behaviors** Scale (Teacher SRS) include acting out behaviors. Five items on this scale rate the frequency with which a child argues, flights, gets angry, acts impulsively, and disturbs ongoing activities.
- The **Internalizing Problem Behavior** Scale (Teacher SRS) asks about the apparent presence of anxiety, loneliness, low self-esteem, and sadness. This scale comprises four items.

These measures are adapted with permission from the instrument Elementary Scale A (“How Often?”) (Gresham and Elliott, 1990).

2.5 Special Education Teacher Questionnaires

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

- The teacher who managed the child’s Individualized Education Plan (IEP),
- The teacher who spent the most amount of time providing special education services to the child, or

- The teacher who was most knowledgeable about the child's special needs and equipment.

The spring-first grade special education teacher questionnaires were very similar to the one used in the spring of the base year. The only difference was the updating of questions to refer to the transition from kindergarten into first grade rather than into kindergarten. Table 2-5 provides a summary of the content areas addressed in the special education teacher questionnaires in both the base year and spring-first grade. The questionnaires addressed topics such as the child's disability, IEP goals, the amount and type of services used by sampled students and communication with parents and general education teachers.

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

2.6 Adaptive Behavior Scale

The ABS (Lambert, Nihira, and Leland, 1993) was completed for all sampled children who were identified in each of the data collection rounds as excluded from the direct child assessment due to a disability. A child with a disability was excluded from the assessment if he or she needed the assessment administered in Braille, enlarged print, or sign language, or if the child's IEP specifically prohibited the child from taking standardized assessments. This questionnaire was completed by the child's primary special education teacher and asked the teacher to provide ratings of the sampled child in three domains: independent functioning (domain I), language development (domain IV), and numbers and time (domain V).

Table 2-5.—Special education teacher questionnaire topics in spring-kindergarten and spring-first grade*

Teachers of sampled students with IEPs questionnaire	1998-99 school year spring-kindergarten	1999-2000 school year spring-first grade
Part A (Teacher Level)		
Teacher's sex	X	X
Teacher's age	X	X
Teacher's race-ethnicity	X	X
Teaching experience	X	X
Educational background	X	X
Special education teacher background	X	X
Location of service provision	X	X
Student load per week	X	X
Part B (Child Level)		
Disability category	X	X
IEP goals for the school year	X	X
Extent of services	X	X
Types of services provided for the year	X	X
Primary placement	X	X
Teaching practices, methods, and materials	X	X
Assistive technologies used by child	X	X
General education goals, expectations and assessments	X	X
Collaboration/communication with child's general education teacher	X	X
Frequency of communicating with child's parents	X	X
Receipt of formal evaluations in the past year	X	X

* Data collected only in the spring of each school year.

2.7 School Administrator Questionnaire

The school principal, administrator, or headmaster at the school attended by the sampled child, was asked to complete the school administrator questionnaire in the spring of 2000. This self-administered questionnaire was intended to gather information about the school, student body, teachers, school policies, and administrator characteristics. The questionnaire was divided into nine sections. The first seven sections requested mainly factual information about each school and the programs offered at the school. Either a principal or a designee who was able to provide the requested information could

complete these sections. The school's principal was asked to complete the remaining two sections concerning his or her background and evaluations of the school climate.

Two versions of the questionnaire were used. One version was aimed at updating information from returning schools, that is, schools that were part of the base year. The second version of the questionnaire was given to school administrators of new schools. Both questionnaires incorporated questions from the spring-kindergarten school administrator's questionnaire. Table 2-6 shows the topics covered in the spring-kindergarten and spring-first grade questionnaires. The spring-first grade questionnaire for returning schools gathered less information than the questionnaire for new schools, as indicated in the table by "--".

For nonresponding and late-responding schools, interviewers were trained to visit the school and encourage the school administrators to complete the questionnaire. If necessary the interviewers were to sit down with the administrators to help them fill out the questionnaire. However, if the school administrators were still reluctant to complete the full questionnaire, the interviewers were instructed to obtain key information. This key information covered such topics as school sector and focus; the school environment, particularly, the safety of the school; school policies and practices; school programs for special populations; staffing and teacher characteristics; and principal characteristics.

2.8 School Facilities Checklist

ECLS-K supervisors completed the facilities checklist during their visits to the school in the spring of first grade. The facilities checklist collects information about the (1) availability and condition of selected school facilities such as classrooms, gymnasiums, toilets, etc., (2) presence and adequacy of security measures, (3) presence of environmental factors that may affect the learning environment, and (4) overall learning climate of the school. An additional set of questions on portable classrooms, that were not included in the spring-kindergarten facilities checklist, was added to the checklist completed in the spring-first grade data collection. (See chapter 5, section 5.5 for more detail on the collection of these data.)

Table 2-6.—School administrator questionnaire, spring-kindergarten and spring-first grade

School questionnaire	Spring-kindergarten	1999-2000 school year spring-first grade	
		Returning schools	New schools
School characteristics	X	--	X
Type of school	X		X
Admission requirements	X		
School size	X	X	X
Student characteristics	X	X	X
Race-ethnicity of students	X	X	X
Children eligible for special services	X	X	X
Types of kindergarten programs	X		
School facilities and resources	X	--	X
Computer equipment	X	X	X
Community characteristics and school safety	X	X	X
Teaching and other school staff characteristics	X	X	X
Range of salary paid to teachers	X		X
Race-ethnicity of staff	X	X	X
School policies and programs	X	--	X
Assessments, testing, and retention	X	X	All grades
School-family-community connections	X	--	X
Programs and activities for families	X		X
Parent involvement and participation	X	X	X
Programs for special populations	X	X	X
ESL and bilingual education	X	X	X
Special education	X	--	X
Gifted and talented	X		X
Principal characteristics	X	X	X
Sex, race-ethnicity, age of principal	X	X	X
Experience and education	X	X	X
School governance and climate	X	X	X
Goals and objectives for teachers	X	X	X
School functioning and decisionmaking	X	X	X

NOTE: "--" indicates that fewer details on the topic were collected than for new schools.

2.9 School Records Abstract Form

School staff completed the student records abstract form for each sampled child in the spring of kindergarten and first grade. This instrument was used to obtain information about the child's attendance record, presence of and details on a child's IEP and information about the type of language or English proficiency screening that the school used. A copy of each child's report card was also obtained. The spring-first grade version of the student records abstract form differed from the spring-kindergarten version in two ways. First, no data were collected on the pre-kindergarten Head Start status of children in the first grade followup. Second, two questions on the form were modified to enable the school to provide more comprehensive answers to the question of the status of the child in the previous school year (1998-99) and whether a student had an IEP. (See chapter 5, section 5.5.5 for more detail on the collection of these forms.)

REFERENCES

- Cooper, H., Nye, B., Charlton, K., Lindsay, J., and Greathouse, S. (1996). The effects of summer vacation on achievement test scores: A narrative and meta-analytic review. *Review of Educational Research* 66(3), 227-268.
- Entwisle, D.R., and Alexander, K.L. (1992). Summer setback: Race, poverty, school composition, and mathematics achievement in the first two years of school. *American Sociological Review* 57(1), 72-84.
- Entwisle, D.R., and Alexander, K.L. (1994). Winter setback: The racial composition of schools and learning to read. *American Sociological Review* 59(3), 446-460.
- Gresham, F.M. and Elliott, S.N. (1990). *Elementary Scale A ("How Often?")*. American Guidance Service, Inc.
- Heyns, B. (1978). *Summer learning and the effects of schooling*. New York: Academic Press.
- Lamert, N., Nihira, K., and Leland, H. (1993). *Adaptive Behavior Scale, Second Edition*. The American Association on Mental Retardation.
- National Center for Education Statistics. (2001). *ECLS-K Base Year Public-Use User's Manual*. (NCES 2001-029).
- National Center for Education Statistics. (2001). *ECLS-K Base Year Restricted-Use Head Start Data Files and Electronic Code Book*. (NCES 2001-025).
- National Center for Education Statistics. (2001). *ECLS-K Base Year Restricted-Use Salary and Benefits Data Files and Electronic Code Book*. (NCES 2001-014).
- National Center for Education Statistics. (2000). *ECLS-K Restricted-Use Base Year User's Manual*. (NCES 2000-097).

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

Several types of scores are used in the Early Childhood Longitudinal Study-Kindergarten Class of 1998-99 (ECLS-K) to describe children's cognitive and social development during kindergarten and first grade. These scores are for the direct cognitive assessment, the academic rating scale (ARS), and the social rating scale (SRS). Descriptions of the scores for each assessment or scale follow, along with variable names, variable descriptions, and descriptive statistics from the ECLS-K data files.¹ Guidelines for when to use each cognitive test score are also provided in this chapter.

3.1 Direct Cognitive Assessment

The direct cognitive assessment contained items in reading, mathematics, and general knowledge in the fall and spring of kindergarten, and fall and spring of first grade.² In each subject area, children received a 12- to 20-item routing test. Performance on the routing items guided the selection and administration of one of several second-stage forms. The second-stage form contained items of appropriate difficulty for the level of ability indicated by the routing items.³

There are five different types of scores that can be used to describe children's performance on the direct cognitive assessment: (1) number-right scores and (2) Item Response Theory (IRT) scale scores, which measure children's performance on a set of test questions with a broad range of difficulty; (3) standardized scores (T-scores), which report children's performance relative to their peers; (4) criterion-referenced proficiency-level scores; and (5) proficiency probability scores, which evaluate children's performance with respect to subsets of test items that mark specific skills. Kindergarten and first grade test responses were pooled to stabilize the longitudinal estimates. As a result, the re-estimated kindergarten IRT scores, T-scores, and proficiency probability scores in this database differ very slightly from the scores in the base year kindergarten-only file previously released.⁴ See chapter 2 for a description of the ECLS-K assessment battery.

¹ This user's manual is applicable to the data gathered during the 1999-2000 school year; information contained in this manual about data gathered during the 1998-99 school year (base year of the study) is provided primarily for background and comparison purposes.

² The fall-first grade sample was a 27 percent subsample of the children in 30 percent of the base-year sampled schools. See chapter 4 for more detail on the subsampling procedures.

³ See chapter 2, section 2.2, for additional information on the two-stage process for the direct cognitive assessments.

⁴ Please see ECLS-K, Base Year Public-Use Data File, Kindergarten Class of 1998-99: Data Files and Electronic Code Book (NCES 2001-029), February 2001.

Tables 3-1 through 3-6 show the types of scores, variable names, descriptions, ranges, means, and standard deviations for the direct cognitive assessment. The name and description for each variable in the tables begin with a “C,” indicating that it is a child variable, and a data collection round number, either 1 (fall-kindergarten), 2 (spring-kindergarten), 3 (fall-first grade), or 4 (spring-first grade).

3.1.1 Number-Right Scores

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

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

3.1.2 Item Response Theory Scale Scores

Scores based on the full set of test items were calculated using IRT procedures. IRT made it possible to calculate scores that could be compared regardless of which second-stage form a child took. IRT uses the pattern of right, wrong, and omitted responses to the items actually administered in a test and the difficulty, discriminating ability, and “guess-ability” of each item to place each child on a continuous ability scale. The items in the routing test, plus a core set of items shared among the different second-stage forms, made it possible to establish a common scale. It is then possible to estimate the score the child would have achieved if all of the items in all of the test forms had been administered.

Table 3-1.—Direct cognitive assessment: number-right scores *

Variable	Description	Range of values	Weighted mean	Standard deviation
C1RROUNR	C1 Reading Routing Test, - Number Right	0 - 20	5.7	3.9
C1MROUNR	C1 Mathematics Routing Test, - Number Right	0 - 16	4.4	2.9
C1GROUNR	C1 General Knowledge Routing Test, - # Right	0 - 12	4.7	2.9
C2RROUNR	C2 Reading Routing Test, - Number Right	0 - 20	9.8	4.2
C2MROUNR	C2 Mathematics Routing Test, - Number Right	0 - 16	7.1	3.3
C2GROUNR	C2 General Knowledge Routing Test, - # Right	0 - 12	6.2	3.0
C3RROUNR	C3 Reading Routing Test, - Number Right	0 - 20	11.6	4.3
C3MROUNR	C3 Mathematics Routing Test, - Number Right	0 - 16	8.8	3.4
C3GROUNR	C3 General Knowledge Routing Test, - # Right	0 - 12	7.3	2.9
C4RROUNR	C4 Reading Routing Test, - Number Right	0 - 20	16.2	3.7
C4MROUNR	C4 Mathematics Routing Test, - Number Right	0 - 16	11.6	3.1
C4GROUNR	C4 General Knowledge Routing Test, - # Right	0 - 12	8.4	2.7

* See chapter 7, section 7.3 for variable naming conventions.

IRT has several other advantages over raw number-right scoring. By using the overall pattern of right and wrong responses to estimate ability, IRT can compensate for the possibility of a low-ability student guessing several hard items correctly. If answers on several easy items are wrong, a correct difficult item is, in effect, assumed to have been guessed. Omitted items are also less likely to cause distortion of scores, as long as enough items have been answered right and wrong to establish a consistent pattern. Unlike raw scoring, which, in effect, treats omitted items as if they had been answered incorrectly, IRT procedures use the pattern of responses to estimate the probability of correct responses for all test questions. Finally, IRT scoring makes possible longitudinal measurement of gain in achievement over time, even though the tests administered are not identical at each point. The common items present in the routing test and in overlapping second-stage forms allow the test scores to be placed on the same scale, even as the two-stage test design adapts to children's growth over time. As noted earlier, kindergarten and first-grade test responses were pooled to stabilize the longitudinal estimates. As a result, the re-estimated kindergarten IRT scores in this database differ very slightly from the IRT scores in the base year kindergarten-only file previously released. In addition, the maximum value of the reading scale score has been expanded from 72 to 92 to include the more difficult items administered in the first-grade tests.

The IRT scale scores in the database represent estimates of the number of items students would have answered correctly if they had taken all of the 92 questions in the first- and second-stage reading forms, the 64 questions in all of the mathematics forms, and the 51 general knowledge items. These scores are not integers because they are probabilities of correct answers, summed over all items in the pool. Gain scores may be obtained by subtracting the estimated number-right at fall-kindergarten from the estimated number-right at spring-kindergarten, spring-kindergarten from spring-first grade, etc. (Note that scores for different subject areas are not comparable to each other because they are based on different numbers of test questions, i.e., it would not be correct to assume that a child is doing better in reading than in mathematics because his or her IRT scale score in reading is higher.)

See table 3-2 for variable names, descriptions, ranges, weighted means, and standard deviations for the IRT scale scores.

Table 3-2.—Direct cognitive assessment: Item Response Theory scale scores *

Variable	Description	Range of values	Weighted mean	Standard deviation
C1RSCALE	C1 Reading IRT Scale Score	0 - 92	22.7	8.6
C1MSCALE	C1 Mathematics IRT Scale Score	0 - 64	19.3	7.1
C1GSCALE	C1 General Knowledge IRT Scale Score	0 - 51	22.1	7.4
C2RSCALE	C2 Reading IRT Scale Score	0 - 92	32.5	10.9
C2MSCALE	C2 Mathematics IRT Scale Score	0 - 64	27.2	8.7
C2GSCALE	C2 General Knowledge IRT Scale Score	0 - 51	26.8	7.8
C3RSCALE	C3 Reading IRT Scale Score	0 - 92	38.0	12.7
C3MSCALE	C3 Mathematics IRT Scale Score	0 - 64	32.4	9.6
C3GSCALE	C3 General Knowledge IRT Scale Score	0 - 51	30.0	7.9
C4RSCALE	C4 Reading IRT Scale Score	0 - 92	54.8	14.2
C4MSCALE	C4 Mathematics IRT Scale Score	0 - 64	42.8	9.5
C4GSCALE	C4 General Knowledge IRT Scale Score	0 - 51	34.0	7.7

* See chapter 7, section 7.3 for variable naming conventions.

3.1.3 Standardized Scores (T-Scores)

T-scores provide norm-referenced measurements of achievement, that is, estimates of achievement level *relative to the population as a whole*. A high mean T-score for a particular subgroup indicates that the group's performance is high in comparison to other groups. It does not mean that group

members have mastered a particular set of skills, only that their mastery level is greater than a comparison group. Similarly, a change in mean T-scores over time reflects a change in the group’s status with respect to other groups. In other words, they provide information on *status compared to children’s peers*, while the IRT scale scores and proficiency scores represent *status with respect to achievement on a particular criterion set of test items*. The T-scores can only provide an indicator of the extent to which an individual or a subgroup ranks higher or lower than the national average and how much this relative ranking changes over time.

The standardized scores reported in the database are transformations of the IRT theta (ability) estimates, rescaled to a mean of 50 and standard deviation of ten using cross-sectional sample weights for each wave of data. For example, a T-score of 55 (C1RTSCOR) represents a reading achievement level that is one-half of a standard deviation higher than the mean for the fall-kindergarten population represented by the tested sample of ECLS-K participants.

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

Table 3-3.—Direct cognitive assessment: standardized scores *

Variable	Description	Range of values	Weighted mean	Standard deviation
C1RTSCOR	C1 Reading T-Score	0 - 90	50.0	10.0
C1MTSCOR	C1 Mathematics T-Score	0 - 90	50.0	10.0
C1GTSCOR	C1 General Knowledge T-Score	0 - 90	50.0	10.0
C2RTSCOR	C2 Reading T-Score	0 - 90	50.0	10.0
C2MTSCOR	C2 Mathematics T-Score	0 - 90	50.0	10.0
C2GTSCOR	C2 General Knowledge T-Score	0 - 90	50.0	10.0
C3RTSCOR	C3 Reading T-Score	0 - 90	50.0	10.0
C3MTSCOR	C3 Mathematics T-Score	0 - 90	50.0	10.0
C3GTSCOR	C3 General Knowledge T-Score	0 - 90	50.0	10.0
C42RTSCOR	C4 Reading T-Score	0 - 90	50.0	10.0
C4MTSCOR	C4 Mathematics T-Score	0 - 90	50.0	10.0
C4GTSCOR	C4 General Knowledge T-Score	0 - 90	50.0	10.0

* See chapter 7, section 7.3 for variable naming conventions.

3.1.4 Proficiency Scores

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

Reading:

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

Mathematics:

- **Number and shape:** identifying some one-digit numerals, recognizing geometric shapes, and one-to-one counting of up to ten objects
- **Relative size:** reading all single-digit numerals, counting beyond ten, recognizing a sequence of patterns, and using nonstandard units of length to compare objects
- **Ordinality, sequence:** reading two-digit numerals, recognizing the next number in a sequence, identifying the ordinal position of an object, and solving a simple word problem
- **Addition/subtraction:** solving simple addition and subtraction problems
- **Multiplication/division:** solving simple multiplication and division problems and recognizing more complex number patterns

Clusters of items provide a more reliable test of proficiency than do single items because of the possibility of guessing; it is very unlikely that a student who has not mastered a particular skill would be able to guess enough answers correctly to pass a four-item cluster. The proficiency levels were assumed to follow a Guttman model, that is, a student passing a particular skill level was expected to have mastered all lower levels; a failure should have indicated nonmastery at higher levels. Only a very small percentage of students in kindergarten and first grade had response patterns that did not follow the Guttman model, that is, a failing score at a lower level followed by a pass on a more difficult item cluster. Overall, including all four rounds of data collection, only about 6 percent of reading response patterns, and about 5 percent of math test results, failed to follow the expected hierarchical pattern. This does not necessarily indicate a different order of learning for these children; since most of the proficiency-level items were multiple choice, many of these reversals are due to children guessing.

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

Proficiency-Level Scores (Dichotomous)

The proficiency-level scores reflect the children's raw ECLS-K direct cognitive assessment scores. These scores are intended for very distinct kinds of analysis. When using the dichotomous proficiency-level scores, it is necessary to take into account the sources of missing data: that scores may be missing because certain items were not administered to particular children. For this reason, most analysts prefer to use the proficiency probability scores, which already compensate for missing data.

For each proficiency level, a score of one was assigned to children who correctly answered at least three of the four items in the cluster, and a score of zero was given if at least two items were incorrect or don't know. If children did not answer enough items for pass or fail to be determined for a particular cluster, a pass/fail score was assigned only if the remaining proficiency scores indicated a level that was unambiguous. That is, a "fail" may be inferred if there were easier cluster(s) that had been failed and no higher cluster passed; and a "pass" may be assumed if harder cluster(s) were passed and no easier one failed. In the case of ambiguous (e.g., pass, blank, fail, where the blank could be either a pass or a fail) or contradictory (e.g., fail, blank, pass) patterns, no imputation is possible without reference to other information, such as the IRT-based scores.

Averaging the sample subgroup's zero and one scores for a particular proficiency cluster results in an estimate of the proportion of children in the subgroup *who answered that cluster* and had mastered the material at that level. The difference between this average at two points in time represents the proportion of these children who attained mastery during that time period. To the extent that there is non-imputable missing data for these levels, the scores are *not* designed to extrapolate to the entire population of kindergarten children. These scores simply show that *of the children who took the items represented by the level*, how many passed the level. In the reading routing test, for example, the harder items (levels four and five) were not administered to children who had difficulty with the easier tasks. Therefore, the missing data for the upper levels consists primarily of children who were not able to get to the more difficult material in the test. It would be incorrect to average the zero and one scores for reading level five and assume that the average represented a population proportion, since the missing cases would have lowered the average substantially. For example, the passing rate for reading level 4 in fall-kindergarten was .04, while the passing rate for level 5 was .26. This does *not* indicate that more children in the population would be likely to do well at level 5 than at level 4. It is a result of the test being discontinued for children who could not answer the items, so the .26 proportion is based only on the small sample of the very best readers: those who were able to complete the whole reading routing test.

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

Proficiency Probability Scores (Continuous)

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

Table 3-4.—Direct cognitive assessment: proficiency-level scores*

Variable	Description	Range of values	Weighted mean	Standard deviation
C1RPROF1	C1 Prof 1 - Letter Recognition	0 - 1	0.65	0.48
C1RPROF2	C1 Prof 2 - Beginning Sounds	0 - 1	0.30	0.46
C1RPROF3	C1 Prof 3 - Ending Sounds	0 - 1	0.18	0.38
C1RPROF4	C1 Prof 4 - Sight Words	0 - 1	0.04	0.20
C1RPROF5	C1 Prof 5 - Word in Context	0 - 1	0.26	0.44
C1MPROF1	C1 Prof 1 - Count, Number, Shape	0 - 1	0.89	0.31
C1MPROF2	C1 Prof 2- Relative size	0 - 1	0.55	0.50
C1MPROF3	C1 Prof 3 - Ordinality, sequence	0 - 1	0.20	0.40
C1MPROF4	C1 Prof 4 - Add/Subtract	0 - 1	0.04	0.20
C1MPROF5	C1 Prof 5 - Multiply/Divide	0 - 1	0.02	0.13
C2RPROF1	C2 Prof 1 - Letter Recognition	0 - 1	0.92	0.28
C2RPROF2	C2 Prof 2 - Beginning Sounds	0 - 1	0.70	0.46
C2RPROF3	C2 Prof 3 - Ending Sounds	0 - 1	0.50	0.50
C2RPROF4	C2 Prof 4 - Sight Words	0 - 1	0.14	0.35
C2RPROF5	C2 Prof 5 - Word in Context	0 - 1	0.24	0.43
C2MPROF1	C2 Prof 1- Count, Number, Shape	0 - 1	0.95	0.21
C2MPROF2	C2 Prof 2 - Relative Size	0 - 1	0.82	0.38
C2MPROF3	C2 Prof 3 - Ordinality, Sequence	0 - 1	0.54	0.50
C2MPROF4	C2 Prof 4 - Add/Subtract	0 - 1	0.17	0.37
C2MPROF5	C2 Prof 5 - Multiply/Divide	0 - 1	0.04	0.19
C3RPROF1	C3 Prof 1 - Letter Recognition	0 - 1	0.95	0.22
C3RPROF2	C3 Prof 2 - Beginning Sounds	0 - 1	0.81	0.39
C3RPROF3	C3 Prof 3 - Ending Sounds	0 - 1	0.66	0.48
C3RPROF4	C3 Prof 4 - Sight Words	0 - 1	0.25	0.43
C3RPROF5	C3 Prof 5 - Word in Context	0 - 1	0.32	0.46
C3MPROF1	C3 Prof 1- Count, Number, Shape	0 - 1	0.94	0.24
C3MPROF2	C3 Prof 2 - Relative Size	0 - 1	0.89	0.31
C3MPROF3	C3 Prof 3 - Ordinality, Sequence	0 - 1	0.71	0.46
C3MPROF4	C3 Prof 4 - Add/Subtract	0 - 1	0.32	0.47
C3MPROF5	C3 Prof 5 - Multiply/Divide	0 - 1	0.08	0.27
C4RPROF1	C4 Prof 1 - Letter Recognition	0 - 1	0.99	0.11
C4RPROF2	C4 Prof 2 - Beginning Sounds	0 - 1	0.93	0.26
C4RPROF3	C4 Prof 3 - Ending Sounds	0 - 1	0.89	0.31
C4RPROF4	C4 Prof 4 - Sight Words	0 - 1	0.78	0.42
C4RPROF5	C4 Prof 5 - Word in Context	0 - 1	0.52	0.50
C4MPROF1	C4 Prof 1- Count, Number, Shape	0 - 1	0.94	0.24
C4MPROF2	C4 Prof 2 - Relative Size	0 - 1	0.96	0.19
C4MPROF3	C4 Prof 3 - Ordinality, Sequence	0 - 1	0.91	0.29
C4MPROF4	C4 Prof 4 - Add/Subtract	0 - 1	0.70	0.46
C4MPROF5	C4 Prof 5 - Multiply/Divide	0 - 1	0.25	0.43

* See chapter 7, section 7.3 for variable naming conventions.

Due to the two-stage format of the cognitive assessment battery, not all children received all items. An IRT model was employed to produce proficiency probability scores, which reflect the probability that a child would have passed a proficiency level. The item clusters were treated as single items for the purpose of IRT calibration, in order to estimate students' probabilities of mastery of each set of skills. The hierarchical nature of the skill sets justified the use of the IRT model in this way. Because the proficiency probabilities were based on overall performance, they could be calculated for all children who had scorable test data, not just those with relatively complete sets of responses to the necessary item clusters.

The proficiency probability scores can be averaged to produce estimates of mastery rates within population subgroups. These continuous measures can provide a closer look at individuals' status and change over time. Gains in probability of mastery at each proficiency level allow researchers to study not only the amount of gain in total scale score points but also where along the score scale different children are making their largest gains in achievement during a particular time interval. Thus, students' school experiences can be related to improvements in specific skills.

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

See table 3-5 for variable names, descriptions, ranges, weighted means, and standard deviations for the proficiency probability scores.

Familiarity with Conventions of Print

Some items from the child assessment measured children's familiarity with conventions of print but were not part of the set of proficiency scores because they did not fit the hierarchical pattern.

Table 3-5.—Direct cognitive assessment: proficiency probability scores*

Variable	Description	Range of values	Weighted mean	Standard deviation
C1RPROB1	C1 Prob 1 - Letter Recognition	0 - 1	0.64	0.41
C1RPROB2	C1 Prob 2 - Beginning Sounds	0 - 1	0.29	0.33
C1RPROB3	C1 Prob 3 - Ending Sounds	0 - 1	0.16	0.26
C1RPROB4	C1 Prob 4 - Sight Words	0 - 1	0.02	0.13
C1RPROB5	C1 Prob 5 - Word in Context	0 - 1	0.01	0.08
C1MPROB1	C1 Prob 1 - Count, Number, Shape	0 - 1	0.92	0.18
C1MPROB2	C1 Prob 2 - Relative Size	0 - 1	0.53	0.35
C1MPROB3	C1 Prob 3 - Ordinality, Sequence	0 - 1	0.19	0.30
C1MPROB4	C1 Prob 4 - Add/Subtract	0 - 1	0.03	0.12
C1MPROB5	C1 Prob 5 - Multiply/Divide	0 - 1	0.00	0.03
C2RPROB1	C2 Prob 1 - Letter Recognition	0 - 1	0.92	0.22
C2RPROB2	C2 Prob 2 - Beginning Sounds	0 - 1	0.67	0.33
C2RPROB3	C2 Prob 3 - Ending Sounds	0 - 1	0.48	0.35
C2RPROB4	C2 Prob 4 - Sight Words	0 - 1	0.13	0.28
C2RPROB5	C2 Prob 5 - Word in Context	0 - 1	0.04	0.16
C2MPROB1	C2 Prob 1 - Count, Number, Shape	0 - 1	0.99	0.07
C2MPROB2	C2 Prob 2 - Relative Size	0 - 1	0.83	0.25
C2MPROB3	C2 Prob 3 - Ordinality, Sequence	0 - 1	0.52	0.39
C2MPROB4	C2 Prob 4 - Add/Subtract	0 - 1	0.16	0.26
C2MPROB5	C2 Prob 5 - Multiply/Divide	0 - 1	0.02	0.08
C3RPROB1	C3 Prob 1 - Letter Recognition	0 - 1	0.96	0.16
C3RPROB2	C3 Prob 2 - Beginning Sounds	0 - 1	0.81	0.27
C3RPROB3	C3 Prob 3 - Ending Sounds	0 - 1	0.65	0.32
C3RPROB4	C3 Prob 4 - Sight Words	0 - 1	0.24	0.37
C3RPROB5	C3 Prob 5 - Word in Context	0 - 1	0.09	0.25
C3MPROB1	C3 Prob 1 - Count, Number, Shape	0 - 1	0.99	0.04
C3MPROB2	C3 Prob 2 - Relative Size	0 - 1	0.91	0.19
C3MPROB3	C3 Prob 3 - Ordinality, Sequence	0 - 1	0.71	0.36
C3MPROB4	C3 Prob 4 - Add/Subtract	0 - 1	0.32	0.34
C3MPROB5	C3 Prob 5 - Multiply/Divide	0 - 1	0.05	0.16
C4RPROB1	C4 Prob 1 - Letter Recognition	0 - 1	0.99	0.07
C4RPROB2	C4 Prob 2 - Beginning Sounds	0 - 1	0.96	0.14
C4RPROB3	C4 Prob 3 - Ending Sounds	0 - 1	0.91	0.20
C4RPROB4	C4 Prob 4 - Sight Words	0 - 1	0.76	0.37
C4RPROB5	C4 Prob 5 - Word in Context	0 - 1	0.42	0.41
C4MPROB1	C4 Prob 1 - Count, Number, Shape	0 - 1	1.00	0.02
C4MPROB2	C4 Prob 2 - Relative Size	0 - 1	0.98	0.09
C4MPROB3	C4 Prob 3 - Ordinality, Sequence	0 - 1	0.93	0.21
C4MPROB4	C4 Prob 4 - Add/Subtract	0 - 1	0.71	0.34
C4MPROB5	C4 Prob 5 - Multiply/Divide	0 - 1	0.24	0.33

* See chapter 7, section 7.3 for variable naming conventions.

The score for these questions was obtained by counting the number of correct answers (zero to three) for the following three items, administered while the child was looking at an illustrated story.

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

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

See table 3-6 for variable names, descriptions, ranges, weighted means, and standard deviations for the conventions of print scores.

Table 3-6.—Direct cognitive assessment: print familiarity scores^a

Variable	Description	Range of values	Weighted mean	Standard deviation
C1RPRINT	C1 Print Familiarity	0 - 3 ^b	1.8	1.1
C2RPRINT	C2 Print Familiarity	0 - 3 ^b	2.3	0.9
C3RPRINT	C3 Print Familiarity	0 - 3 ^b	2.6	0.8
C4RPRINT	C4 Print Familiarity	0 - 3 ^b	2.8	0.6

^a See chapter 7, section 7.3 for variable naming conventions.

^b Ranges for discrete scores (routing number right, proficiency level right/wrong, print familiarity) are reported without a decimal point because they are always integers. Ranges for continuous scores (IRT estimates and probability scores) have decimal points because the scores themselves are not integers.

3.1.5 Choosing the Appropriate Score for Analysis

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

- A measure of overall achievement versus achievement in specific skills;
- An indicator of status at a single point in time versus growth over time; and
- A criterion-referenced versus norm-referenced interpretation.

Item Response Theory-Based Scores

The scores derived from the IRT model (IRT scale scores, T-scores, proficiency probabilities) are based on all of the child’s responses to a subject area assessment. That is, the pattern of right and wrong answers, as well as the characteristics of the assessment items themselves, are used to estimate a point on an ability continuum, and this ability estimate, theta, then provides the basis for criterion-referenced and norm-referenced scores.

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

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

- The **standardized scores (T-scores)** are also overall measures of status at a point in time, but they are **norm-referenced** rather than criterion-referenced. They do not answer the question, “What skills do children have?” but rather “**How do they compare with their peers?**” The transformation to a familiar metric with a mean of 50 and standard deviation of 10 facilitates comparisons in standard deviation units. T-score means may be used longitudinally to illustrate the **increase or decrease in gaps**

in achievement among subgroups over time. T-scores are not recommended for measuring individual gains over time. The IRT scale scores are used for that purpose.

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

Non-Item Response Theory-Based Scores

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

- **Routing test number-right scores** for the reading, math, and general knowledge assessments are based on 20, 16, and 12 items respectively. They target specific sets of skills and cover a broad range of difficulty. These scores may be of interest to researchers because they are based on a specific set of test items, which was the same for all children who took the test.
- **Proficiency-level scores** are based on the same sets of items as the proficiency probability scores but are dichotomous, rather than continuous, measures of proficiency. They simply report whether children were able to answer correctly on at least three out of four actual test items in a cluster.

Users of the proficiency-level scores should be aware of possible bias due to missing data. Stopping and starting rules employed in the administration of the tests to minimize stress on low-performing children result in substantial missing scores. For example, low-performing children did not receive the items associated with the higher proficiencies, and therefore they are missing scores on the higher proficiency levels. And, higher performing children may not have received the items associated with the lower proficiency levels, and therefore are missing scores on the lower proficiency

levels. Estimates based on variables with substantial amounts of missing data can be assumed to generalize to the whole sample *only* if “missing-ness” is unrelated to what the variable is measuring. This condition is called “MAR,” or Missing-At-Random. The missing are *not* missing-at-random, they were not administered based on performance. Estimates based on proficiency-level scores, without adjustments for missing data, would overstate or understate the population performance.

Users of the proficiency-level scores will need to compensate for the missing data in order to use these scores appropriately. A simple approach would be to impute a “pass” score (=1) for missing data where the child has passed a more difficult level (and has not failed an easier one). For example, a score pattern for the 5 levels of blank-pass-pass-pass-fail could reasonably be imputed to be pass-pass-pass-pass-fail. Similarly, one might assume that a blank level following one or more failed levels could be interpreted as a potential failure at the higher level. A child who had not been administered reading level 5 because of poor performance on easier tasks might have a pattern of pass-fail-fail-fail-blank. Interpreting the missing score as a probable failure would not be unreasonable. Some score patterns do not have a clear indication of pass or fail (e.g., pass-pass-blank-fail-fail, where either a pass or a fail for level 3 would produce a consistent pattern; or a “reversal” pattern such as fail-pass-fail-blank-blank). These situations are a relatively small proportion of the missing scores. They may best be imputed by reference to the corresponding proficiency probability score, and imputing a “pass” if the probability is high, or a “fail” if it is low. This may be done either by rounding the probability to 0 or 1, or by generating a random 0 or 1, using the proficiency probability score as the probability of generating a 1.

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

3.1.6 Reliabilities

Reliability statistics appropriate for each type of score were computed for each subject area, for fall- and spring-kindergarten and fall- and spring-first grade assessments. For the IRT-based scores, the reliability of the overall ability estimate, theta, is based on the variance of repeated estimates of theta. These reliabilities, ranging from 0.88 to 0.97, apply to all of the scores derived from the theta estimate, namely, the IRT scale scores, T-scores, and proficiency probabilities. Alpha coefficients for the routing test number correct ranged from 0.78 to 0.88. Split half reliabilities were computed for the item clusters that made up the dichotomous proficiency-level scores and the conventions of print cluster. These reliabilities were higher for the reading clusters (0.60 to 0.83) than for the math levels (0.26 to 0.67). (As noted earlier, the general knowledge test did not contain hierarchical proficiency levels.) The difference in

internal consistency statistics is due to the reading items being essentially replications of the same task, while the math items had a greater diversity of content.

Note that the split half reliabilities for the low-level item clusters decreased from fall- to spring-kindergarten and beyond, while the reliabilities for the clusters at the upper end tended to increase. This is a consequence of changes in the variance of the cluster scores as children progressed in their development of skills. After the initial round of testing, the vast majority of children had mastered the lowest proficiency levels, so the sample variance was low, resulting in lower reliability for subsequent rounds. Conversely, the sample variance for the difficult tasks was very low in the fall-kindergarten, when most children had *not* mastered these skills, and the reliability tended to rise as more and more children attained high-level proficiency in spring-kindergarten and first grade, increasing the total variance. This effect is more pronounced for the math than for the reading clusters for two reasons. First, the math item clusters were more heterogeneous than the reading, in terms of content and difficulty. Second, the reading item clusters were based entirely on items from the routing test, which was taken by all children, while the lowest math cluster employed items from the low level second-stage test as well. By spring-kindergarten, fewer than one-half of the test takers were routed to the low form (and only about seven percent in spring of first grade), and this constrained variance in ability resulted in a lower reliability for children who had this score. Tables 3-7, 3-8, and 3-9 present the reliability statistics for all of the test scores.

Table 3-7.—Reliability of Item Response Theory-based scores*

Category	IRT-based scores (reliability of theta)		IRT-based scores (reliability of theta)	
	Fall- kindergarten	Spring- kindergarten	Fall- first grade	Spring- first grade
Reading	0.93	0.95	0.96	0.97
Mathematics	0.92	0.94	0.94	0.94
General Knowledge	0.88	0.89	0.89	0.89

* Though the majority of base year children were in first grade during the 1999-2000 school year, about five percent of the sampled children were retained in kindergarten or in a second year of kindergarten and a handful of others were in second grade during the 1999-2000 school year.

Table 3-8.—Reliability of routing test number correct (alpha coefficient) *

Category	Routing test number correct (alpha coefficient)		Routing test number correct (alpha coefficient)	
	Fall- kindergarten	Spring- kindergarten	Fall- first grade	Spring- first grade
Reading	0.86	0.88	.88	.86
Mathematics	0.78	0.81	.83	.80
General Knowledge	0.79	0.79	.79	.78

* Though the majority of base year children were in first grade during the 1999-2000 school year, about five percent of the sampled children were retained in kindergarten and a handful of others were in second grade during the 1999-2000 school year.

Table 3-9.—Split half reliability of item-cluster-based scores (proficiency-level scores)*

Category	Fall- kindergarten	Spring- kindergarten	Fall- first grade	Spring- first grade
Reading level 1	0.83	0.79	0.77	0.78
Reading level 2	0.76	0.76	0.73	0.70
Reading level 3	0.72	0.76	0.76	0.68
Reading level 4	0.78	0.77	0.80	0.78
Reading level 5	0.60	0.69	0.73	0.73
Conventions of print	0.70	0.68	0.68	0.60
Mathematics level 1	0.41	0.27	0.26	0.26
Mathematics level 2	0.58	0.49	0.51	0.32
Mathematics level 3	0.63	0.66	0.67	0.59
Mathematics level 4	0.54	0.63	0.66	0.63
Mathematics level 5	0.46	0.53	0.61	0.65

* Though the majority of base year children were in first grade during the 1999-2000 school year, about five percent of the sampled children were retained in kindergarten and a handful of others were in second grade during the 1999-2000 school year.

3.2 Indirect Cognitive Assessment

The academic rating scale (ARS) was developed for the ECLS-K to measure teachers' evaluations of students' academic achievement in the three domains that are also directly assessed in the cognitive battery: language and literacy (reading), general knowledge (science and social studies), and mathematical thinking. Teachers rated the child's skills, knowledge, and behaviors on a scale from "Not

Yet” to “Proficient” (see table 3-10). If a skill, knowledge, or behavior had not been introduced into the classroom yet, the teacher coded that item as N/A (not applicable). The difference between the direct and indirect cognitive assessments, and the scores available, are described here. For a discussion of the content areas of the ARS, see chapter 2, section 2.4.1.

Table 3-10.—Academic rating scale response scale

1	Not yet:	Child <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.

3.2.1 Comparison to Direct Cognitive Assessment

The ARS was designed both to overlap and to augment the information gathered through the direct cognitive assessment battery. Although the direct and indirect instruments measure children’s skills and behaviors within the same broad curricular domains with some intended overlap, several of the constructs they were designed to measure differ in significant ways. Most importantly, the ARS includes items designed to measure both the process and products of children’s learning in school, whereas the direct cognitive battery measures only the products of children’s achievement. Because of time and space limitations, the direct cognitive battery is less able to measure the process of children’s thinking, including the strategies they use to read, solve math problems, or investigate a scientific phenomenon.

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

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

3.3 Rasch Scores Available for the Academic Rating Scale

IRT was used to calculate scores for the ARS in order to compare performance of students on a hierarchy of skills, knowledge, and behavior. The Rasch Rating Scale model uses the pattern of ratings on the items to determine an estimate of the difficulty of each item and to place each student on a continuous ability scale (in this case 1-5). Rasch provides interval-level measurement. The Rasch analysis showed that the reliability of the estimates of child ability was very high for all domains (see table 3-11).

Table 3-11.—Person reliability for the Rasch-based score

Category	Spring-kindergarten	Spring-first grade
ARS Language and Literacy	.91	.94
ARS Mathematical Thinking	.94	.94
ARS General Knowledge	.95	.95

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

The variable names, descriptions, value ranges, weighted means, and standard deviations for the first grade (T4) ARS scores are shown in table 3-12. The description for each variable in the tables begins with a “T,” indicating that it is a teacher questionnaire child-level variable. The round of data collection is indicated by a “1” for fall-kindergarten, a “2” for spring-kindergarten, and a “4” for spring-first grade. As noted earlier, an error was identified in the base year ARS scores. Specifically, the fall and spring base year ARS scores used slightly different metrics and, therefore, were not directly comparable.

These scores have been recomputed to put them on the same metric. The corrected base year ARS scores are included on the first grade data files.

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

Table 3-12.—First grade academic rating scale: variable names, descriptions, ranges, weighted means, and standard deviations¹

Variable name	Description ²	Range of values	Weighted mean	Standard deviation
T1RARSLI	T1 REC Language and Literacy ARS Score	1 - 5	2.48	0.73
T1RAR SMA	T1 REC Mathematical Thinking ARS Score	1 - 5	2.54	0.82
T1RARSGE	T1 REC General Knowledge ARS Score	1 - 5	2.62	0.98
T2RARSLI	T2 REC Literacy ARS Score	1 - 5	3.33	0.81
T2RAR SMA	T2 REC Math ARS Score	1 - 5	3.50	0.86
T2RARSGE	T2 REC General Knowledge ARS Score	1 - 5	3.55	0.99
T4ARSLIT	T4 Literacy ARS Score	1 - 5	3.40	0.93
T4ARSMAT	T4 Math ARS Score	1 - 5	3.43	0.90
T4ARSGEN	T4 General Knowledge ARS Score	1 - 5	3.26	0.99

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

² In the table, the letters “REC” in the variable description indicate that the variable was recalibrated.

Tables 3-13 to 3-18 provide the estimates of difficulty for each of the items. Higher values mean that teachers rated fewer students as proficient on those items. Students would have a greater than 50 percent probability of receiving ratings of “5” on items below their ability level. Tables are provided for both kindergarten and first grade items.

Table 3-13.—Kindergarten academic rating scale language and literacy item difficulties (arranged in order of difficulty)

Item difficulty	Item number and abbreviated content
2.44	Q3. Easily and quickly names all upper- and lower- case letters of the alphabet
2.76	Q4. Produces rhyming words
2.79	Q2. Understands and interprets a story or other text read to him/her
2.83	Q1. Uses complex sentence structures
3.20	Q8. Demonstrates an understanding of some of the conventions of print
3.21	Q5. Reads simple books independently
3.30	Q6. Uses different strategies to read unfamiliar words
3.45	Q7. Composes simple stories

Table 3-14.—Kindergarten academic rating scale mathematical thinking item difficulties (arranged in order of difficulty)

Item difficulty	Item number and abbreviated content
2.72	Sorts, classifies, and compares math materials by rules and attributes
2.79	Demonstrates an understanding of graphing activities
2.82	Orders a group of objects
2.90	Shows an understanding of the relationship between quantities
3.08	Solves problems involving numbers using concrete objects
3.22	Uses a variety of strategies to solve math problems
3.45	Uses instruments accurately for measuring

Table 3-15.—Kindergarten academic rating scale general knowledge item difficulties (arranged in order of difficulty)

Item difficulty	Item number and abbreviated content
2.78	Classifies and compares living and non-living things in different ways
2.95	Uses his/her senses to explore and observe
2.99	Recognizes distinct differences in group habits and living patterns
3.11	Recognizes some ways in which people rely on each other for goods and services
3.12	Forms explanations based on observations and explorations

Table 3-16.—Spring-first grade academic rating scale language and literacy item difficulties (arranged in order of difficulty)

Item difficulty	Item number and abbreviated content
2.73	Q3. Reads words with regular vowel sounds
2.85	Q1. Contributes relevant information to classroom discussions
2.85	Q5. Reads first grade books independently with comprehension
2.89	Q2. Understands and interprets a story or other text read to him/her
2.97	Q6. Reads first grade books fluently
3.14	Q8. Demonstrates an understanding of some of the conventions of print
3.17	Q4. Reads words with irregular vowel sounds,
3.30	Q7. Composes a story with a clear beginning, middle, and end

Table 3-17.—Spring-first grade academic rating scale mathematical thinking item difficulties (arranged in order of difficulty)

Item difficulty	Item number and abbreviated content
2.39	Models, reads, writes, and compares whole numbers
2.61	Demonstrates an understanding of place value
2.70	Surveys, collects, and organizes data into simple graphs
2.73	Makes reasonable estimates of quantities
2.81	Counts change with two different types of coins
2.90	Measures to the nearest whole number using common instruments
2.90	Uses a variety of strategies to solve math problems

Table 3-18.—Spring-first grade academic rating scale general knowledge item difficulties (arranged in order of difficulty)

Item difficulty	Item number and abbreviated content
2.82	Classifies and compares living and non-living things in different ways
2.88	Identifies similarities and differences in group habits and living patterns
2.96	Shows a beginning understanding that maps represent actual places
3.02	Forms explanations and conclusions based on observation and investigation
3.02	Recognizes some ways in which people rely on each other for goods and services
3.02	Makes logical predictions

The scale was designed to provide information on children’s abilities at a given point in time, not necessarily over time. Even though a common calibration of the fall and spring scores was used for the kindergarten ratings, it is not recommended that users compute a change score. In addition, although the item stems are similar across kindergarten and first grade, the actual items include performance criteria that increase in difficulty from one grade to the next. Moreover, the kindergarten and first grade ARS scores are placed on different metrics. Therefore, change scores should not be used between kindergarten and first grade. Covariance models may be used to compare teacher’s ratings of performance in kindergarten and first grade. Before using these variables in such analyses, the distribution of the samples should be assessed to determine if the assumption of normal distribution is met.

Tables 3-19 to 3-24 provide standard errors (SE) for each of the Rasch scores for kindergarten and first grade. The “Score” column is the sum of the raw score ratings. “Measure” is the Rasch-based score. The column labeled “SE” is the corresponding standard error of measurement for those scores. These standard errors can be used in analytic models to correct for the heteroskedasticity of scores.

Table 3-19.—Kindergarten academic rating scale language and literacy standard errors

Score	Measure	S.E.	Score	Measure	S.E.	Score	Measure	S.E.
8	1.00E	.60	19	2.63	.16	30	3.43	.16
9	1.42	.34	20	2.70	.15	31	3.52	.16
10	1.69	.26	21	2.78	.15	32	3.60	.17
11	1.86	.22	22	2.85	.15	33	3.69	.17
12	2.00	.20	23	2.92	.15	34	3.78	.18
13	2.11	.19	24	2.99	.15	35	3.89	.19
14	2.22	.18	25	3.06	.15	36	4.00	.20
15	2.31	.17	26	3.13	.15	37	4.14	.22
16	2.40	.16	27	3.21	.15	38	4.32	.26
17	2.48	.16	28	3.28	.16	39	4.58	.34
18	2.56	.16	29	3.36	.16	40	5.00E	.60

Table 3-20.—Kindergarten academic rating scale mathematical thinking standard errors

Score	Measure	S.E.	Score	Measure	S.E.	Score	Measure	S.E.
7	1.00E	.46	17	2.56	.16	27	3.66	.16
8	1.33	.27	18	2.67	.16	28	3.77	.16
9	1.55	.21	19	2.79	.17	29	3.88	.16
10	1.71	.19	20	2.90	.16	30	3.99	.17
11	1.86	.18	21	3.01	.16	31	4.11	.18
12	1.99	.18	22	3.12	.16	32	4.25	.19
13	2.11	.17	23	3.23	.16	33	4.42	.22
14	2.23	.17	24	3.33	.16	34	4.66	.28
15	2.34	.16	25	3.44	.16	35	5.00E	.46
16	2.45	.16	26	3.55	.16			

Table 3-21.—Kindergarten academic rating scale general knowledge standard errors

Score	Measure	S.E.	Score	Measure	S.E.	Score	Measure	S.E.
5	1.00E	.38	12	2.41	.18	19	3.73	.20
6	1.28	.23	13	2.57	.18	20	3.95	.22
7	1.49	.19	14	2.75	.20	21	4.17	.20
8	1.66	.18	15	2.98	.22	22	4.35	.18
9	1.83	.19	16	3.21	.20	23	4.52	.19
10	2.03	.21	17	3.39	.18	24	4.72	.23
11	2.24	.19	18	3.55	.18	25	5.00E	.37

Table 3-22.—Spring-first grade academic rating scale language and literacy standard errors

Score	Measure	S.E.	Score	Measure	S.E.	Score	Measure	S.E.
8	1.02E	.50	19	2.51	.15	30	3.53	.16
9	1.37	.29	20	2.60	.15	31	3.63	.16
10	1.60	.22	21	2.69	.16	32	3.72	.16
11	1.75	.19	22	2.78	.16	33	3.82	.16
12	1.87	.17	23	2.87	.16	34	3.91	.16
13	1.97	.16	24	2.97	.16	35	4.02	.17
14	2.07	.16	25	3.06	.16	36	4.13	.18
15	2.16	.15	26	3.16	.16	37	4.26	.19
16	2.25	.15	27	3.25	.16	38	4.41	.22
17	2.34	.15	28	3.35	.16	39	4.65	.29
18	2.42	.15	29	3.44	.16	40	5.00E	.50

Table 3-23.—Spring-first grade academic rating scale mathematical thinking standard errors

Score	Measure	S.E.	Score	Measure	S.E.	Score	Measure	S.E.
7	1.00E	.46	17	2.54	.16	27	3.68	.17
8	1.33	.27	18	2.65	.16	28	3.80	.17
9	1.55	.21	19	2.76	.17	29	3.92	.17
10	1.71	.18	20	2.87	.17	30	4.05	.17
11	1.84	.17	21	2.99	.17	31	4.17	.18
12	1.96	.17	22	3.11	.17	32	4.30	.18
13	2.08	.17	23	3.23	.17	33	4.46	.21
14	2.19	.17	24	3.34	.17	34	4.68	.27
15	2.31	.17	25	3.45	.17	35	5.00E	.46
16	2.42	.17	26	3.57	.17			

Table 3-24.—Spring-first grade academic rating scale general knowledge standard errors

Score	Measure	S.E.	Score	Measure	S.E.	Score	Measure	S.E.
6	1.000E	.394	15	2.446	.163	24	3.867	.214
7	1.282	.231	16	2.577	.169	25	4.071	.196
8	1.476	.182	17	2.725	.185	26	4.234	.176
9	1.619	.167	18	2.902	.197	27	4.376	.171
10	1.751	.167	19	3.080	.186	28	4.523	.183
11	1.888	.173	20	3.232	.172	29	4.718	.231
12	2.036	.178	21	3.367	.167	30	5.001E	.394
13	2.183	.173	22	3.504	.174			
14	2.319	.165	23	3.664	.194			

The majority of teachers rated more than one student on the ARS. The number of students rated by each teacher ranged from one to 23.

3.4 Oral Language Development Scale

The language assessment scores (OLDS scores) for language minority children are located with the other child scores on the file. There are a total of 16 OLDS scores—four English and four Spanish for each of the two rounds (i.e., fall-first grade and spring-first grade). Children in households speaking languages other than English, who had not been administered the English ECLS-K cognitive assessments in the prior rounds of the study, were first given the English OLDS. Of that group, those scoring below the cut point of the English OLDS were administered the Spanish OLDS if the child's

home language was noted as Spanish by the school. (See chapter 2, section 2.2 for more detail on the content of the OLDS items.)

The variable names, descriptions, value ranges, weighted means, and standard deviations for the OLDS are shown in table 3-25. The description for each variable in the tables begin with a “C,” indicating that it is a child variable, and a data collection round number, either 3 (fall-first grade) or 4 (spring-first grade).

Table 3-25.—Oral language development scale: variable names, descriptions, ranges, means, and standard deviations*

Variable	Description	Range of values	Weighted mean	Standard deviation
C3SCTOT	C3 AIQ400 Child’s Total OLDS Score	0 - 60	36.8	17.6
C3SCORD	C3 AIQ400 Simon Says Child Score	0 - 10	8.3	2.6
C3SCART	C3 AIQ400 Art Show Child Score	0 - 10	7.2	3.0
C3SCSTO	C3 AIQ400 Tell Stories Child Score	0 - 40	21.3	13.2
C3SSCTOT	C3 SAIQ400 Spanish Total OLDS Scores	0 - 42	24.5	6.7
C3SSCORD	C3 SAIQ400 Spanish Simon Says Child Score	0 - 10	9.8	0.7
C3SSCART	C3 SAIQ400 Spanish Art Show Child Score	0 - 10	8.5	2.2
C3SSCSTO	C3 SAIQ400 Spanish Tell Stories Score	0 - 24	6.2	6.0
C4SCTOT	C4 AIQ400 Child’s Total OLDS Score	0 - 60	38.7	15.6
C4SCORD	C4 AIQ400 Simon Says Child Score	0 - 10	8.8	2.2
C4SCART	C4 AIQ400 Art Show Child Score	0 - 10	7.6	2.7
C4SCSTO	C4 AIQ400 Tell Stories Child Score	0 - 40	22.3	11.9
C4SSCTOT	C4 SAIQ400 Spanish Total OLDS Scores	0 - 40	23.9	6.7
C4SSCORD	C4 SAIQ400 Spanish Simon Says Child Score	0 - 10	9.7	1.5
C4SSCART	C4 SAIQ400 Spanish Art Show Child Score	0 - 10	8.7	1.9
C4SSCSTO	C4 SAIQ400 Spanish Tell Stories Score	0 - 20	5.5	5.7

* See chapter 7, section 7.3 for variable naming conventions.

3.5 Social Rating Scale

The SRS asked both teachers and parents to tell how often a student exhibited certain social skills and behaviors. Teachers and parents used a frequency scale (see table 3-26) to report on how often the student demonstrated the behavior described. See chapter 2, section 2.3 and 2.4 for additional

information on the parent and teacher SRS instruments. The scale scores on all SRS scales are the mean rating on the items included in the scale. Scores were computed only if the student was rated on at least two-thirds of the items in that scale. The same items were administered in fall and spring-kindergarten, and in spring-first grade, so change scores may be computed by subtracting time 1 (fall- kindergarten) from time two (spring-kindergarten) scores, etc. The reliability for the teacher SRS scales is high (see table 3-27). The reliability is lower for the parent scales (see table 3-28).

Table 3-26.—Social rating scale response scale

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

Table 3-27.—Split half reliability for the teacher social rating scale scores

Category	Fall-kindergarten	Spring-kindergarten	Spring-first grade
Approaches to Learning	0.89	0.89	.89
Self-control	0.79	0.80	.80
Interpersonal	0.89	0.89	.89
Externalizing Problem Behaviors	0.90	0.90	.86
Internalizing Problem Behaviors	0.80	0.78	.77

Table 3-28.—Split half reliability for the parent social rating scale scores

Category	Fall-kindergarten	Spring-kindergarten	Spring-first grade
Approaches to Learning	0.68	0.69	.69
Self-Control	0.74	0.75	.75
Social Interaction	0.70	0.68	.69
Impulsive/Overactive	0.46	0.47	.48
Sad/Lonely	0.60	0.61	.63

3.5.1 Teacher Social Rating Scale

Teachers rated individual students as part of a self-administered questionnaire. The five social skill teacher scales are as follows: approaches to learning, self-control, interpersonal skills, externalizing problem behaviors, and internalizing problem behaviors.

Variable names for the teacher scores, descriptions, ranges, weighted means, and standard deviations for these scales are shown in table 3-29. Numbers in the table for round 4 are for first graders, with kindergarten repeaters' scores shown in parentheses. The description for each variable in the tables begin with a "T," indicating that it is a teacher variable, and a data collection round number, either 1 (fall-kindergarten), 2 (spring-kindergarten), or 4 (spring-first grade).

Care should be taken when entering these scales into the same analysis due to problems of multicollinearity. The intercorrelations among the five SRS factors are high. The factor intercorrelations with the internalizing problem behaviors are the lowest. The correlations between the teacher SRS factors range from 0.25 to 0.78 in fall-kindergarten and from 0.30 to 0.80 in spring-kindergarten (absolute values). In round 4, the correlations ranged from .31 to .81 for first graders, and .29 to .80 for kindergarten repeaters.

Table 3-29.—Teacher social rating scores: variable names, descriptions, ranges, weighted means, and standard deviations*

Variable	Description	Range of values	Weighted mean	Standard deviation
T1LEARN	T1 Approaches to Learning	1 - 4	3.0	0.7
T1CONTRO	T1 Self-Control	1 - 4	3.1	0.6
T1INTERP	T1 Interpersonal	1 - 4	3.0	0.6
T1EXTERN	T1 Externalizing Problem Behaviors	1 - 4	1.6	0.6
T1INTERN	T1 Internalizing Problem Behaviors	1 - 4	1.6	0.5
T2LEARN	T2 Approaches to Learning	1 - 4	3.1	0.7
T2CONTRO	T2 Self-Control	1 - 4	3.2	0.6
T2INTERP	T2 Interpersonal	1 - 4	3.1	0.6
T2EXTERN	T2 Externalizing Problem Behaviors	1 - 4	1.7	0.7
T2INTERN	T2 Internalizing Problem Behaviors	1 - 4	1.6	0.5
T4LEARN	T4 Approaches to Learning	1 - 4	3.0 (2.9)	0.7 (0.7)
T4CONTRO	T4 Self-Control	1 - 4	3.2 (3.0)	0.6 (0.6)
T4INTERP	T4 Interpersonal	1 - 4	3.1 (3.0)	0.6 (0.7)
T4EXTERN	T4 Externalizing Problem Behaviors	1 - 4	1.7 (1.8)	0.6 (0.7)
T4INTERN	T4 Internalizing Problem Behaviors	1 - 4	1.6 (1.7)	0.5 (0.6)

* See chapter 7, section 7.3 for variable naming conventions.

3.5.2 Parent Social Rating Scale

The items on the parent SRS were administered as part of a longer telephone or in-person survey. (See chapter 2, section 2.3 for a description of the parent scales.) The factors on the parent SRS are similar to the teacher SRS; however, the items in the parent SRS are geared to the home environment and thus are not the same items. It is also important to keep in mind that parents and teachers observe the children in very different environments. The five social skill parent scales are as follows: approaches to learning, self-control, social interaction, impulsive/overactive, and sad/lonely. The correlations between the parent SRS factors were not as high as the teacher SRS factors. They ranged from 0.05 to 0.45 in fall-kindergarten, from 0.08 to 0.45 in spring-kindergarten, and from .11 to .45 in spring-first grade (absolute values).

Variable names for the parent scores, descriptions, ranges, weighted means, and standard deviations for these scales are shown in table 3-30. The description for each variable in the tables begin with a “P,” indicating that it is a parent variable, and a data collection round number, either 1 (fall-kindergarten), 2 (spring-kindergarten), or 4 (spring-first grade).

Table 3-30.—Parent social rating scores: variable names, descriptions, ranges, weighted means, and standard deviations*

Variable	Description	Range of values	Weighted mean	Standard deviation
P1LEARN	P1 Approaches to Learning	1 - 4	3.1	0.5
P1CONTRO	P1 Self-Control	1 - 4	2.8	0.5
P1SOCIAL	P1 Social Interaction	1 - 4	3.3	0.6
P1SADLON	P1 Sad/Lonely	1 - 4	1.5	0.4
P1IMPULS	P1 Impulsive/Overactive	1 - 4	2.0	0.7
P2LEARN	P2 Approaches to Learning	1 - 4	3.1	0.5
P2CONTRO	P2 Self-Control	1 - 4	2.9	0.5
P2SOCIAL	P2 Social Interaction	1 - 4	3.4	0.5
P2SADLON	P2 Sad/Lonely	1 - 4	1.6	0.4
P2IMPULS	P2 Impulsive/Overactive	1 - 4	2.0	0.7
P4LEARN	P4 Approaches to Learning	1 - 4	3.1	0.5
P4CONTRO	P4 Self-Control	1 - 4	3.0	0.5
P4SOCIAL	P4 Social Interaction	1 - 4	3.4	0.5
P4SADLON	P4 Sad/Lonely	1 - 4	1.5	0.4
P4IMPULS	P4 Impulsive/Overactive	1 - 4	1.9	0.7

* See chapter 7, section 7.3 for variable naming conventions.

4. SAMPLE DESIGN AND IMPLEMENTATION

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

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

4.1 Base Year Sample

In the base year, children were selected for the ECLS-K using a multistage probability design. The PSUs were counties or groups of counties selected with probability proportional to size (PPS). The basic PSU measure of size was the number of 5-year-olds, but this was modified to facilitate the oversampling of Asian/Pacific Islanders (APIs) required to meet precision goals. In all, there were 100 PSUs selected for the ECLS-K. The 24 PSUs with the largest measure of size were designated self-representing (SR) and were included in the sample with certainty. The remaining non-SR PSUs were partitioned into 38 strata of roughly equal measure of size. An initial cross-classification of census region with metropolitan statistical area (MSA) status created eight superstrata. These were further subdivided by percent minority, size, and 1988 per capita income. From each non-SR stratum, two PSUs were selected PPS without replacement using Durbin's Method.¹

¹ Durbin, J. (1967). Design of multi-stage surveys for the estimation of sampling errors. *Journal of the Royal Statistical Society C*, 16, 152-164.

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

Table 4-1.—Distribution of the ECLS-K primary sampling unit sample by self-representing status, metropolitan statistical area status, and census region

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

In the second stage, public and private schools offering kindergarten programs were selected. For each PSU, a frame of public and private schools offering kindergarten programs was constructed using existing school universe files: the 1995-96 Common Core of Data² (CCD) and the 1995-96 Private School Universe Survey³ (PSS). The 1995-96 Office of Indian Education Programs Education Directory was consulted in order to complete the list of Bureau of Indian Affairs (BIA) schools in the CCD file. For Department of Defense (DOD) domestic schools, a 1996 list of schools was obtained directly from the DOD. A procedure was implemented to identify kindergarten programs that would be operational at the time of ECLS-K's Base Year data collection, but that were not included in the frame just described. These were newly opened schools that were not listed in CCD and PSS, and schools that were in CCD and PSS but did not appear to offer kindergarten programs according to those sources. The selection of schools was systematic, with probability proportional to a weighted measure of size based on the number of kindergartners enrolled. As with the PSU sample, the measure of size was constructed taking into account the desired oversampling of APIs. Public and private schools constituted distinct sampling strata. Within each stratum, schools were sorted to ensure good sample representation across other characteristics. In total, 1,280 schools were sampled from the original frame, and 133 from the freshened frame. Of these, 953 were public schools and 460 were private schools.

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

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

The characteristics of the ECLS-K school sample are presented in table 4-2. Schools that were discovered to be ineligible during recruitment have been omitted from the tabulation.

Table 4-2.—Characteristics of the ECLS-K base year school sample

School type	Census region				Total
	Northeast	Midwest	South	West	
Public	161	210	306	237	914
Private*	82	88	112	81	363
Total	243	298	418	318	1277

*120 Catholic, 149 other religious, 94 nonsectarian

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

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

4.2 Fall-First Grade Subsample

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

Base year schools in the 62 fall-first grade sampled PSUs were stratified by frame source (original public, original private, freshened, etc.) and arranged in their original selection order. A 30 percent equal probability subsample of schools was drawn in the 24 SR PSUs and a 60 percent subsample of schools was drawn in the 38 NSR PSUs. In total 311 schools that had cooperated in either fall- or spring-kindergarten were selected. The characteristics of the base year cooperating schools selected for fall-first grade data collection are presented in table 4-3.

Table 4-3.—Characteristics of base year cooperating schools selected for fall-first grade

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

Fall-first grade data collection consisted of the direct child assessment and the parent interview. Data collection was attempted for every eligible child found still attending the school in which

he or she had been sampled during kindergarten. By eligible we mean a base year respondent (i.e., a child who had either a fall- or spring-kindergarten child assessment or parent interview). Base year 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 4-4 shows the characteristics of the children subsampled for fall-first grade. Region, locale, religious affiliation, and type of school describe the school the child attended in kindergarten.

Table 4-4.—Characteristics of children subsampled for fall-first grade*

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

* School characteristics are of the original sampled schools.

Table 4-4.—Characteristics of children subsampled for fall-first grade* (continued)

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

* School characteristics are of the original sampled schools.

4.3 Spring-First Grade Sample

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

4.3.1 Subsampling Movers

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

In selecting the spring-first grade 50 percent subsample of schools where movers would be flagged for follow-up, the three primary strata were SR PSUs, NSR PSUs that had been selected for fall-first grade, and NSR PSUs that had not been selected for fall-first grade. Within these major strata,

schools were grouped by frame source (original public, original private, new from Catholic dioceses, new from local governments, etc.). Finally within each frame source, schools were stratified by response status, and arranged in original selection order. Schools that had been part of the 30 percent fall-first grade sample were automatically retained. Then equal probability sampling methods were employed to augment the sample to the desired 50 percent. The net result of these procedures was that every base year selected school had a 50 percent chance of having its ECLS-K transfer students followed during spring-first grade, and any transfer student who had been followed in fall-first grade would still be followed in spring-first grade.

Table 4-5 shows the characteristics of the children in the spring-first grade sample, excluding freshened students. Region, locale, religious affiliation, and type of school describe the school at which the child attended kindergarten.

Table 4-5.—Characteristics of children in spring-first grade sample, excluding freshened students *

Characteristic	Public	Private	Total
Total	14,248	3,836	18,084
Region			
Northeast	2,434	905	3,339
Midwest	3,474	1,104	4,578
South	5,029	1,021	6,050
West	3,311	806	4,117
Type of locale			
Large city	2,575	884	3,459
Midsize city	2,797	964	3,761
Urban fringe of large city	3,991	1,149	5,140
Urban fringe of midsize city	1,126	162	1,288
Large town	466	110	576
Small town	1,215	363	1,578
Rural	2,078	204	2,282
Religious affiliation			
Catholic	-	2,091	2,091
Other religious	-	1,139	1,139
Nonreligious, private	-	606	606
Type of school			
Regular	13,971	3,306	17,277
Ungraded	24	16	40
No grade beyond K	235	185	420
Unknown	18	329	347
Composite child race			
White	7,472	2,736	10,208
Black	2,289	308	2,597
Hispanic, with race	1,220	240	1,460
Hispanic, without race	1,456	192	1,648
Asian	939	210	1,149
Pacific Islander	186	16	202
Native American	294	38	332
Multirace	347	87	434
Unknown	45	9	54

* School characteristics are of the original sampled schools.

Table 4-5.—Characteristics of children in spring-first grade sample, excluding freshened students (continued)

Characteristic	Public	Private	Total
Highest parent level of education			
Less than high school	1,491	38	1,529
High school graduate	3,356	423	3,779
Vocational/technical	926	152	1,078
Some college	3,313	898	4,211
College graduate	2,194	1,154	3,348
Masters	719	472	1,191
Ph.D./professional	395	354	749
Unknown	1,854	345	2,199

* School characteristics are of the original sampled schools.

4.3.2 Student Freshening

The spring-first grade student freshening used a half-open interval sampling procedure.⁴ The procedure was implemented in the same 50 percent subsample of ECLS-K base year schools where transfer students were flagged for followup. Each of these schools was asked to prepare an alphabetic roster of students enrolled in first grade and the names of ECLS-K kindergarten-sampled students were identified on this list. Beginning with the name of the first kindergarten-sampled child, school records were checked to see whether the student directly below in the sorted list attended kindergarten in the United States in fall 1998. If not, (1) that child was considered to be part of the freshened sample and was linked to the base year sampled student (i.e., was assigned that student's probability of selection); (2) the record search procedure was repeated for the next listed child, etc. When the record search revealed that a child had been enrolled in kindergarten the previous year, (1) that child was not considered part of the freshened sample and (2) the procedure was begun all over again with the second base year sampled student name etc. Note: the student roster was "circularized" (i.e., the first name on the roster was considered to follow the last name on the roster in the implementation of the procedure). Student freshening brought 165 first graders into the ECLS-K sample, which increased the weighted survey estimate of the number first graders in the United States by about 2.6 percent.

⁴ Kish, L. (1965). *Survey Sampling*. John Wiley & Sons, New York., p 56

The student freshening procedure was not entirely free of bias. A first grader would have no chance of being in the ECLS-K first grade sample if he or she was enrolled in a school where neither the child nor any of his or her classmates had attended kindergarten in the United States in fall 1998. This would be a rare circumstance and is not thought to be an important source of bias. A more significant source of potential bias is nonresponse. One source of nonresponse inherent to the freshening plan was that the procedure only involved students who had not transferred from the school in which they had been sampled during the base year. A more detailed discussion of freshened student nonresponse can be found in section 5.7.2.

4.4 Calculation and Use of Sample Weights

As in the base year, the ECLS-K data were weighted to compensate for differential probabilities of selection at each sampling stage and to adjust for the effects of nonresponse. In the second year of the ECLS-K, which covers fall 1999 (fall-first grade) and spring 2000 (spring-first grade), only child-level weights were computed. The use of these weights is essential to produce estimates that are representative of the population of first grade children in the school year 1999-2000 and of kindergartners in 1998-99 one year later. The sample of teachers in the first grade year only represents teachers of ECLS-K children eligible for the first grade survey. The sample of teachers is not representative of the population of first grade teachers in the country. The same applies to schools. Therefore, teacher- and school-level weights are not provided.

Several sets of weights were computed for each of the two rounds of data collection (fall- and spring-first grade). First grade longitudinal weights were also computed for children with complete data from both rounds (fall and spring) of first grade. As in the base year, there are several survey instruments administered to sampled children and their parents, teachers and schools: cognitive and physical assessments for children in both fall- and spring-first grade; parent instruments in both fall- and spring-first grade; several types of teacher instruments in spring-first grade only; and school instruments, also in spring-first grade only. The stages of base year sampling in conjunction with the nonresponse at each stage and the diversity of survey instruments require that multiple sampling weights be computed for use in analyzing the ECLS-K data. Kindergarten through first grade longitudinal weights for the analysis of both kindergarten and first grade data were also created. Details on these longitudinal weights are available in chapter 9, section 9.4. This section describes the cross-sectional first grade weights and the within first grade longitudinal weights.

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

4.4.1 Types of Sample Weights

Several sets of weights were computed for each round of data collection in the first grade year and for children with complete data from both rounds. Careful consideration should be given to the choice of a weight for a specific analysis since it depends on the type of data analyzed. Each set of weights is appropriate for a different set of data or combination of sets of data.

Tables 4-6 and 4-7 summarize the different types of cross-sectional weights and how they should be used. Cross-sectional weights provide an accurate estimate for the specific round of data collection. Table 4-6 describes weights for fall-first grade estimates, and table 4-7 describes weights for spring-first grade estimates. Table 4-8 describes weights for first grade longitudinal estimates.

Table 4-6.—ECLS-K fall-first grade cross-sectional weights

<u>Weight</u>	<u>to be used for analysis of ...</u>
C3CW0	fall-first grade direct child assessment data alone or in conjunction with a) a limited set of child characteristics (e.g. age, sex, race-ethnicity), and b) school or teacher data collected in spring-first grade.
C3PW0	fall-first grade parent interview data alone or in combination with a) fall-first grade child assessment data, and b) school or teacher data collected in spring-first grade.

Table 4-7.—ECLS-K spring-first grade cross-sectional weights

<u>Weight</u>	<u>to be used for analysis of ...</u>
C4CW0	spring-first grade direct child assessment data, alone or in conjunction with any combination of a) a limited set of child characteristics (e.g. age, sex, race-ethnicity), b) any spring-first grade teacher questionnaire A, B or C data, and c) data from the school administrator questionnaire or facilities checklist
C4PW0	spring-first grade parent interview data alone or in combination with a) spring-first grade child assessment data, b) spring-first grade teacher questionnaire A, B, or C data, and c) data from the school administrator questionnaire or facilities checklist. <i>Exception:</i> If data from the parent AND child assessment AND teacher questionnaire A or B (not C) are used then C4CPTW0 should be used
C4CPTW0	spring-first grade direct child assessment data combined with spring-first grade parent interview data AND spring-first grade teacher data alone or in conjunction with data from the school administrator or facilities checklist

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

1. Is the analysis concerned with one point in time or two? If the analysis pertains to a) fall-first grade (single point in time) then table 4-6 guides the selection of weights, b) spring-first grade (single point in time) then go to table 4-7, and c) both fall- and spring-first grade (two points in time) then go to table 4-8.
2. What instruments do the data to be used in the analysis come from? There are several options when deciding on which weights to use, and the source of the data affect which weight to use. In each of the tables, details under “to be used in the analysis of . . .” provide guidance based on whether the data were collected through the child assessments, parent interviews or teacher questionnaires.

Weight C3CW0 is used to estimate child-level characteristics or assessment scores for fall-first grade, and C4CW0 is for spring-first grade. Examples of such estimates are the percent of first grade children who are male, the percent of spring-kindergarten children who are API, the percent of fall-first grade children who are seven when they begin first grade, and the mean reading score of children in fall-first grade. These weights exist not only for children who were administered a child assessment but

also for children who could not be assessed due to a disability or because they were not proficient in English due to a non-English or non-Spanish home language (LM/not Spanish). These children were not administered the ECLS-K direct cognitive battery, but their background characteristics such as age, gender, race-ethnicity, and characteristics of their parents, teachers, classrooms, and schools are available from the parent interviews, the teacher questionnaires, and the school administrator questionnaire. The social rating scores (see chapter 3, section 3.5) from parents and teachers are also available for LM/not Spanish children and children with disabilities, regardless of whether they completed the direct child assessment.

When analyzing spring child assessment data in conjunction with teacher data collected in spring-first grade, weights C4CW0 (for spring-first grade) should be used. An example for the use of C4CW0 is in the analysis of the relationship between children's approaches to learning as rated by their teachers, the teacher's type of teaching certification, and the children's cognitive skills and knowledge. Some data may be missing because some teachers did not complete the questionnaire, but these are the most appropriate weights for this type of analysis. However, different weights should be used for analysis of child data in conjunction with both parent and teacher data (C4CPTW0).

C3PW0 (for fall-first grade) and C4PW0 (for spring-first grade) are used for child-level estimates associated with data collected through the parent interview. Examples are the percent of children whose mothers are currently employed, the percent of children who are in a particular type of child care, and the percent of children who are read to at least every day. These weights should not be used for estimates solely using direct child assessment data but should be used when analyzing parent and child assessment data together, for example, when exploring the relationship between home literacy behaviors and children's reading skills.

C4CPTW0 (for spring-first grade) is used when child direct assessment *and* teacher *and* parent data are combined in an analysis; for example, in the analysis of the relationship between parent education, teacher education, and children's reading knowledge and skills. These weights should be not be used for estimates using only direct child assessment data or only parent interview data.

Careful consideration should be given to which set of weights is appropriate for the desired analysis. Using the wrong weights will result in more biased or inefficient estimates. For example, if C4CPTW0 were used in an analysis of child and teacher/classroom data only, then the resulting estimates will be inefficient compared to estimates using C4CW0. The lower parent response causes C4CPTW0 to

result in lower sample size with positive weights. There may be combinations of data from a different source for which no weights were developed, but most analyses are possible from the weights provided. No child-teacher weights were computed for analyzing child data in conjunction with teacher data because the response rates for the teachers are high; for the analysis of child assessment data in conjunction with teacher data, the child-level weights based solely on the presence of direct child assessment data should be used.

The longitudinal or panel weights (table 4-8) are used for estimates of differences at two points in time. Examples of analysis using longitudinal weights include the following:

- First grade fall-spring difference in mean child assessment scores (C34CW0);
- First grade fall-spring difference in mean social skills as rated by children’s parents (C34PW0); and
- The relationship between the gains children make in their reading knowledge and skills, how often their parents read to them, how often their parents take them to the library (C34PW0). This weight is used when the analysis includes data from all four components—fall and spring child assessment, and parent data.

Table 4-8.—ECLS-K first grade longitudinal (panel) weights

<u>Child-level weight</u>	<u>to be used for analysis of ...</u>
C34CW0	child direct assessment data from BOTH fall-first grade and spring-first grade, alone or in conjunction with any combination of a limited set of child characteristics (e.g., age, sex, race-ethnicity)
C34PW0	parent interview data from BOTH fall-first grade and spring-first grade alone or in combination with fall- and/or spring-first grade child assessment data.

While not all combinations of how the data will be used are presented here, the first-grade longitudinal weights described in table 4-8 can be used in the analysis of data coming from different sources and different times. For example, what weight should be used in the analysis of fall- and spring-first grade child direct assessment data in conjunction with spring-kindergarten parent data? The fact that fall- and spring-first grade data are to be analyzed together determines that a first-grade longitudinal weight should be used. When parent data are added, then the type of weight to be used with parent data is appropriate. In the example just cited, C34PW0 is the appropriate weight for this type of analysis. If

parent data from fall-first grade are to be analyzed with fall- and spring-first grade assessment data, then C34PW0 is also the weight to be used.

The distribution of schools by number of sampled students with nonzero first grade weights and the mean number of sampled students with nonzero weights per school are useful in analysis using hierarchical linear modeling. These are given in table 4-9. For the first grade year, there are a large number of schools with one to five ECLS-K students. For this reason, schools are classified on the basis of the number of students who did not transfer schools between the base year and the rounds of data collection. For example, for C3CW0, counts in table 4-9 include only fall-first grade students who have not transferred out of their original sample schools; for C34CW0, the counts include only students who are both in fall-first grade and spring-kindergarten who have not transferred out of their original sample schools.

Table 4-9.—Distribution of originally sampled schools by number of children with nonzero weights by first grade sample weights

Sample	Number of cases					Mean cases per school
	1 – 5	6 – 10	11 – 15	16 – 20	21 – 27	
Fall-first grade						
C3CW0	11	28	59	144	53	16
C3PW0	13	34	75	138	35	15
Spring-first grade						
C4CW0	55	70	269	436	129	16
C4PW0	77	128	314	363	92	14
C4CPTW0	80	150	301	318	65	14
Longitudinal						
C34CW0	9	35	85	132	32	15
C34PW0	17	47	114	97	19	14

4.4.2 Weighting Procedures

In fall-first grade, a sample of about one-third of the base year schools was selected, and all base year responding children in the sample schools were included in the study. A base year responding child is defined as one with at least one direct cognitive test score in fall- or spring-kindergarten or whose parent responded to the family structure section of the parent instrument in fall- or spring-kindergarten. In addition to subsampling schools, children who moved out of their original sample schools between fall-kindergarten and fall-first grade were subsampled to be followed into their new schools. The mover subsampling rate was 0.3 for children who moved between fall- and spring-kindergarten, and 0.5 otherwise.

The spring-first grade sample included all base year respondents as defined earlier, and a supplemental sample of first graders brought in through a sample freshening procedure. As in fall-first grade, only a subsample of children who moved from the schools they were attending when they were sampled originally were followed into their new schools. The freshening and mover followup activities targeted a 50 percent subsample of base year schools. To preserve the fall-first grade data series, the 30 percent of schools that constituted the fall-first grade sample were automatically included as part of the spring-first grade 50 percent sample.

Since base year responding children are the basis for all subsequent ECLS-K data collections, the weights that are common to both fall- and spring-first grade were as follows:

- Base year school weight adjusted for base year school nonresponse, and
- Base year child weight that is the product of the base year school nonresponse adjusted weight and the inverse of the within school child selection probability.

In subsequent stages, the base year child weights were then adjusted for:

- Subsampling of schools for the study (fall-first grade only),
- Subsampling of schools for freshening (spring-first grade only),
- School freshening nonresponse (spring-first grade only),
- Subsampling of movers (fall- and spring-first grade),

- Child freshening nonresponse (spring-first grade only), and
- Child nonresponse (fall- and spring-first grade).

The final stage of weighting was to rake the final adjusted weights to sample control totals, except for the first graders brought in through sample freshening (see section 4.4.5 on the computation of weights of children sampled in first grade). The computation of the base year child weights, common to both fall- and spring-first grade is described in section 4.4.3. The subsequent weight adjustments are described separately for fall- and spring-first grade in sections 4.4.4 and 4.4.5. Section 4.4.6 describes the different types of weights computed for fall- and spring-first grade.

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

4.4.3 Computation of Base Year Child Weights

Base Year Nonresponse Adjusted School Weights

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

A base year responding school was an original sample school with at least one child with a positive C1CW0, C2CW0, C1PW0, or C2PW0 weight. C1CW0 is positive for LM/not Spanish children, children with disabilities and children with at least one direct cognitive test score in fall-kindergarten. C1PW0 is positive for children whose parents completed the family structure questions of the parent interview in fall-kindergarten. C2CW0 and C2PW0 weights are positive under similar circumstances but for spring-kindergarten. Schools that did not meet this condition are nonrespondents and their weights distributed (at the school level) in this stage. The base year school weight was adjusted within nonresponse weighting classes created in the base year using the Chi-squared Automatic Interaction

Detector (CHAID) and variables with known values for both respondents and nonrespondents. School characteristics used for constructing nonresponse cells were the type of school (public, Catholic private, non-Catholic private, or nonsectarian private), the school locale (large city, midsize city, suburb of large city, suburb of midsize city, large town, small town, or rural area), the region where the school is located (Northeast, Midwest, South, or West), and the size classification of the school in terms of school enrollment. Once the nonresponse cells were determined, the nonresponse adjustment factors are the reciprocals of the response rates within the selected nonresponse cells.

Base Year Child Weights

Only base year child respondents were fielded in fall- and spring-first grade. A base year child respondent is a sampled child with a positive fall- or spring-kindergarten weight (i.e., C1CW0, C2CW0, C1PW0 or C2PW0 weights). The base year child weight is the product of the base year nonresponse adjusted school weight and the inverse of the within school selection probability of the child, adjusted for child-level nonresponse. The nonresponse weighting classes included school characteristics from the school nonresponse adjustments such as type of school, locale, region, school enrollment class, and child characteristics such as age group, gender and race-ethnicity. These weighting classes are similar to those used for the child weights in the base year. For a description of the computation of child weights in the base year, see chapter 4, section 4.3.4 of the ECLS-K Base Year Public-Use User's Manual (NCES 2001-029), February 2001 or the ECLS-K Restricted-Use Base Year User's Manual (NCES 2000-097), August 2000.

4.4.4 Computation of Fall-First Grade Child Weights

In the first step of the computation of the fall-first grade child weight, an initial weight was created for every child using the nonresponse adjusted base year child weight and incorporating the school subsampling factor appropriate for fall-first grade. The weight was then trimmed to reduce the weight of all the children in one private school that had a large school weight.

Next, the weight was adjusted to reflect the subsampling of movers. A follow flag was created for every child fielded in fall-first grade. If the child moved out of the original sample school and the value of the flag was 1 then the child was followed into his or her new school. If the value of the flag

was 0, then the child was not followed and no assessment was attempted. The adjustment factor for subsampling movers was computed as follows:

- 1, if the child was not a mover,
- 0, if the child was a mover and the value of the follow flag was 0, and,
- $\frac{\text{The sum of initial child weights of children who were movers over the sum of initial child weights of children who were movers and whose follow flags have value 1}}{\text{The sum of initial child weights of children who were movers and whose follow flags have value 1}}$, if the child was a mover whose follow flag had value 1.

For the third category, the adjustment factor was computed within mover cells. Two mover cells were created; the first one included children identified as movers in spring-kindergarten (subsamped at a rate of 0.3), and the second cell included children identified as movers in fall-first grade (subsamped at a rate of 0.5).

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

- 0, if the child was of unknown eligibility (group d), and,
- $\frac{\text{The sum of the mover adjusted weights of all children (any group) over the sum of the mover adjusted weights of children who were eligible respondents, eligible nonrespondents or ineligible (group a, b or c)}}{\text{The sum of the mover adjusted weights of all children (any group)}}$, 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,
- $\frac{\text{The sum of the weights adjusted in the first step of eligible children (group a or b) over the sum of the weights adjusted in the first step of eligible responding children (group a)}}{\text{The sum of the weights adjusted in the first step of eligible responding children (group a)}}$, if the child was an eligible respondent.

In both steps of the adjustment, separate nonresponse classes were created for movers and nonmovers using the type of school the child attended when he or she was originally sampled (public, Catholic, non-Catholic private, and nonsectarian private).

To remove the variability due to the subsampling of schools and movers, the child weights were then raked to sample-based control totals⁵ computed using the base year child weights adjusted for nonresponse. A record for every responding eligible child and every ineligible child in the base year is included in this process. In the previous steps, the weights of the nonresponding children have been distributed to the responding children while the weights of the ineligible children have not been affected by this weighting step. The weights of the ineligible children are set to zero at the end of this process because these children are not included in the analysis of the fall-first grade data. The reason for including them in the raking step is that these children were eligible in the base year and hence are in the estimates used as the sampled-based control totals. The raking factor was computed separately within raking cells as the sample-based control total for the raking cell over the sum of the nonresponse adjusted weights for children in the same cell. Raking cells (also known as raking dimensions) were created using school and child characteristics collected in fall-kindergarten and/or spring-kindergarten: type of school, region, urbanicity, gender, age, race-ethnicity and socioeconomic status (SES). To reflect the variation of the control totals, each replicate was raked to the corresponding replicated-based control totals. For a discussion of the replicates, see section 4.4.7.

4.4.5 Computation of Spring-First Grade Child Weights

The computation of the spring-first grade child weights was done separately for children sampled in the base year (referred to as children sampled in kindergarten) and children brought in through a sample freshening procedure (referred to as children sampled in first grade). For children sampled in kindergarten, the weighting steps are the same as for fall-first grade, except that the initial child weights did not include the adjustment for school subsample, applicable only to fall-first grade. Children sampled in first grade through freshening were first linked to children sampled in kindergarten in order to create a child base weight for each of them, then the different stages of adjustments applied.

⁵ These are called sample-based control totals because the numbers used in the numerator of the adjustments are sample estimates subject to sampling errors of roughly the same order as the sampling errors of the estimates from the fall-first grade data. When the numbers used in the numerators are known population totals, the sampling error for estimates such as totals are often substantially reduced. Since the numerators are sample estimates special procedures are needed to reflect this fact in variance estimation.

Weights of Children Sampled in Kindergarten

The spring-first grade initial child weights are the nonresponse adjusted base year child weights described in section 4.4.3. When these weights were examined together with the first grade and kindergarten through first grade longitudinal weights (see chapter 9 for a description of the kindergarten through first grade longitudinal weights), the longitudinal weights that included a fall-first grade component were very large for all children in one particular school. This was a private school with an unusual combination of school and child weights (especially for fall-first grade that only included a subsample of schools) that caused all the children in the school to have large weights. To reduce the effect of the sample from this school on the overall longitudinal estimates and the variances of the estimates, all the weights of the sampled children in the school were trimmed in half; this applies to both cross-sectional and longitudinal weights. No other adjustment was done to compensate for the reduction in weights because of the raking procedure that came later.

Next, the initial child weights were adjusted to reflect the subsampling of movers. As in fall-first grade, follow flags were created for all children in the sample. Children who have moved out of their original sample school were followed in the random 50 percent of schools where the follow flag was set to 1. The adjustment factor for subsampling movers was computed as follows:

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

For the third category, the adjustment factor was computed within mover cells. Mover cells were created using the type of school of the original sample school and the region where the original sample school is located. Three children with large weights had their weights trimmed. However, the weights were not redistributed because the total sum of weights was re-established in the raking procedure that came later.

After the adjustment for subsampling movers, the child weights were adjusted for nonresponse. As in fall-first grade, the nonresponse adjustment was done in two steps. In the first step, the

adjustment was for children whose eligibility was not determined (unknown eligibility). A portion of children of unknown eligibility was assumed to be ineligible. In the second step, the adjustment was for eligible nonrespondents. To carry out these adjustments, each child was classified as (a) an eligible respondent, (b) an eligible nonrespondent, (c) ineligible (out of the country or deceased) or (d) of unknown eligibility (mover who could not be located). The first adjustment factor (for children of unknown eligibility) was computed as:

- 0, if the child was of unknown eligibility (group d), and,
- $\frac{\text{The sum of the mover adjusted weights of all children (any group)}}{\text{The sum of the mover adjusted weights of children who were eligible respondents, eligible nonrespondents or ineligible (group a, b or c)}}$, 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,
- $\frac{\text{The sum of the weights adjusted in the first step of eligible children (group a or b)}}{\text{The sum of the weights adjusted in the first step of eligible responding children (group a)}}$, if the child was an eligible respondent.

In both steps of the adjustment, separate nonresponse classes were created for movers and nonmovers using various combinations of response status of child assessments and parent interviews in the base year as well as the type of household collected in the base year from the parent interviews.

To remove the variability due to the subsampling of schools and movers, the child weights were then raked to sample-based control totals computed using the base year child weights adjusted for nonresponse. A record for every responding eligible child and every ineligible child in the base year is included in this process. In the previous steps, the weights of the nonresponding children were distributed to the responding children while the weights of the ineligible children were not affected by this weighting step. The weights of the ineligible children are set to zero at the end of this process because these children are not included in the analysis of the spring-first grade data. The reason for including them in the raking step is that these children were eligible in the base year and hence are in the estimates used as the sampled-based control totals. The raking factor was computed separately within raking cells as the sample-based control total for the raking cell over the sum of the nonresponse adjusted weights for children in the same cell. Raking cells (also known as raking dimensions) were created using school and child characteristics collected in fall-kindergarten/spring-kindergarten: type of school, region, urbanicity,

gender, age, race-ethnicity, and SES. To reflect the variation of the control totals, each replicate was raked to the corresponding replicated-based control totals. For a discussion of the replicates, see section 4.4.7.

Weights of Children Sampled in First Grade

Since each student brought in through sample freshening was linked to a child sampled in kindergarten, the first step of the weighting procedure for children sampled in first grade was to create a weight using the children who were sampled in kindergarten, that reflected the school freshening subsampling and the school freshening nonresponse (some schools refused to provide information needed for freshening). This weight was then linked to the freshened child and further adjusted for nonresponse due to not obtaining the data from the sample of freshened children.

First the base year nonresponse adjusted school weight (as computed in section 4.4.3) was adjusted for the subsampling of schools for freshening. Student freshening was done in the same 50 percent subsample of schools that were flagged for following movers. The school freshening subsampling adjustment factor was computed as:

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

The freshening procedure could not be applied in all designated schools because some schools did not provide the information needed for freshening. These schools are considered freshening nonrespondents. The school weight adjusted for freshening subsampling was then adjusted for this type of nonresponse. The school freshening nonresponse adjustment factor was calculated as the sum of weights of the freshening adjusted schools weights for all schools designated for freshening over the sum of weights of the freshening adjusted school weights for schools who responded to freshening. In both the numerator and denominator of this factor, the school measure of size was incorporated; the school measure of size is relevant because the weights will be used for child level estimates, not school level estimates. The nonresponse cells for this adjustment were created using the same variables used to create the response cells for the base year school nonresponse adjustment, namely school type, region, locale, and school enrollment class.

Next, the school adjusted weight was multiplied by the inverse of the within school selection probability of the child in the base year to obtain a base year child weight. The base year child weight was then adjusted for base year child nonresponse because children who did not respond in the base year could not be linked to children in first grade in spring 2000. The adjustment factor was computed as the sum of the base year child weights of all base year children over the sum of the base year child weights of base year respondents within each nonresponse cell. The nonresponse cells are the same as those discussed in section 4.4.3, namely, school characteristics such as type of school, locale, region, school enrollment class, and child characteristics such as age group, gender, and race-ethnicity.

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

In the first step, adjustment was done for movers whose grade was unknown. A portion of the movers was assumed to be in first grade. In the second step, the weights of nonmovers in first grade or respondents were adjusted for the movers in first grade or nonrespondents. Each child was classified as a (a) mover in first grade, (b) mover in another grade, (c) mover with unknown grade, (d) nonmover in first grade, and (e) nonmover in another grade. The first step adjustment for movers whose grade was unknown was computed as:

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

In the second step, the weights were adjusted for children who were in first grade but who were not identified to participate in the freshening process because they moved into a new school. The adjustment factor was computed as:

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

Only two weighting cells were used for these adjustments: white children and nonwhite children.

The weights thus created for children sampled in kindergarten were then linked to the children that they brought into the sample in first grade through sample freshening. The children sampled in first grade were then separated for a final adjustment that did not concern the children sampled in kindergarten. This final adjustment was for children sampled in first grade who did not cooperate in the assessment. The nonresponse adjustment factor was computed as:

- 0, if the freshened child was a nonrespondent, and
- The sum of the weights adjusted for freshening of all freshened children over the sum of the weights adjusted for freshening of responding freshened children. if the child was a respondent.

This nonresponse adjustment was not done separately by nonresponse cells because of the small number of children brought in through freshening. Two records with large weights were trimmed by half, and the trimmed parts were distributed to the remaining records.

Note that no adjustment was made for unknown eligibility since every freshened child was assumed to be eligible as indicated by the school unless the parents explicitly stated otherwise. Also, the final weights were not raked because there were no control totals to which they could be raked.

4.4.6 Types of Weights

The different types of cross-sectional and longitudinal weights were described in tables 4-6 to 4-8. They were all created as described in sections 4.4.3 to 4.4.5, but the definition of which children were eligible respondents varied.

Weights to Be Used with Direct Child Assessment Data (C3CW0, C4CW0, C34CW0)

In fall-first grade, responding children for this type of weight were eligible children who had fall-first grade scorable direct child cognitive assessment data, or LM/not Spanish children who did not score at or above the OLDS cut score, or children with disabilities who according to specifications in their IEP could not participate in the assessments. A child was eligible if he or she was a base year respondent. Children who transferred to schools and were not flagged to be followed, who moved out of the country or were deceased were considered ineligible. In spring-first grade, responding children were classified using rules similar to those used in fall-first grade. For the longitudinal weights C34CW0, a respondent is defined as a child for whom both C3CW0 and C4CW0 were nonzero.

Table 4-10 shows the number of children who were not assessed due to the following special situations: children who were LM/not Spanish, children with disabilities, children who moved out of their original sample schools and were not flagged to be followed, children who moved and were flagged to be followed but could not be located or moved into a school in a nonsampled county, and children who moved outside of the country or who were deceased. Only the LM/not Spanish and children with disabilities had weights.

Table 4-10.—Number of children who were not assessed due to special situations

Sample	Number of children	
	Unweighted	Weighted
Fall-first grade		
Children with disabilities*	28	23,373
LM/not Spanish children	33	8,583
Moved from original sample schools		
Subsampled, not to be followed	779	465,761
Nonlocatable or moved to nonsampled PSU	121	75,023
To be followed but became ineligible in fall	4	1,992
Spring-first grade		
Children with disabilities	47	14,421
LM/not Spanish children	39	4,789
Moved from original sample schools		
Subsampled, not to be followed	2,850	543,651
Nonlocatable or moved to nonsampled PSU	719	136,056
To be followed but became ineligible in spring	48	7,852

* These children's IEPs specifically prohibited assessments.

Weights to Be Used with Parent Data (C3PW0, C4PW0, C34PW0)

The weights C3PW0 (fall-first grade) and C4PW0 (spring-first grade) are to be used with parent interview data. In both fall- and spring-first grade, a respondent was defined as a child for whom the family structure section (FSQ) in that child's parent interview for the corresponding round was completed. For the longitudinal weights C34PW0, a respondent is defined as a child for whom both C3PW0 and C4PW0 are nonzero. Note that these weights are at the child level even though the data were collected from the parents; they sum to all first grade children, not to the parents of first grade children.

Weights to Be Used with a Combination of Child Direct Assessment Data and Parent Interview Data and Teacher Data (C4CPTW0)

The weight C4CPTW0 (spring-first grade) is to be used for analysis involving child, parent, and teacher data. A respondent for this type of weight was defined as a child who had scorable cognitive assessment data for spring-first grade (or LM/not Spanish children or children with disabilities), whose parent completed the FSQ section of the parent interview for spring-first grade, and whose teacher completed part B of the teacher questionnaire. Note that this weight was not computed for fall-first grade because teacher data were not collected in this round.

4.4.7 Replicate Weights

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

For fall-first grade, there are 40 replicate weights. For spring-first grade, there are 90 replicate weights. Each set of replicate weights has the same prefix in the variable name as the full sample weight. For example, the replicate weights for C3CW0 are C3CW1 through C3CW40; the replicate weights for C4CW0 are C4CW1 through C4CW90. The method used to compute the replicate weights and how they are used to compute the sampling errors of the estimates are described in section 4.5.

4.4.8 Characteristics of Sample Weights

The statistical characteristics of the sample weights are presented in table 4-11. For each type of weight, the number of cases with nonzero weights is presented together with the mean weight, the standard deviation, the coefficient variation (i.e., the standard deviation as a percentage of the mean weight), the minimum weight, the maximum weight, the skewness, the kurtosis, and the sum of weights.

Table 4-11.—Characteristics of the first grade child-level weights

Sample	Number of cases	Mean	Standard deviation	CV ($\times 100$)	Minimum	Maximum	Skewness	Kurtosis	Sum
Fall-first grade									
C3CW0	5,291	729.33	554.89	76.08	71.10	6374.63	3.78	19.37	3,858,882
C3PW0	5,071	760.96	484.52	63.67	76.35	5246.83	2.84	11.81	3,858,850
Spring-first grade									
C4CW0	16,727	235.46	207.19	88.00	1.76	3517.71	4.31	32.38	3,938,490
C4PW0	15,626	251.96	203.49	80.76	1.83	3271.78	3.98	28.56	3,937,097
C4CPTW0	13,491	291.74	316.85	108.61	2.21	3849.49	4.35	26.07	3,935,870
First grade longitudinal									
C34CW0	5,047	762.96	571.61	74.92	71.81	6225.66	3.63	18.85	3,850,650
C34PW0	4,682	822.17	526.93	64.09	81.12	5657.06	2.61	10.65	3,849,405

The difference in the estimate of the population of students (sum of weights) between rounds of data collection and types of weight is due a combination of factors, among them: (1) the number of base year respondents who became ineligible (due to death, leaving the country, or being a nonsampled mover) after the base year, and (2) the adjustment of the weights for the children of unknown eligibility. The larger sums of weights in spring-first grade is due to the freshening of the sample that brought in a small number of first graders. Otherwise, the population of inference for all weights is the same.

4.5 Variance Estimation

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

4.5.1 Paired Jackknife Replication Method

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

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

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 4.5 were produced using WesVar and JK2.

Replicate Weights

Replicate weights were created to be used in the calculation of replicate estimates. Each replicate weight was calculated using the same adjustment steps as the full sample weight but using only the subsample of cases that constitute each replicate. For the original ECLS-K design in the base year, replicate weights were created taking into account the Durbin method of PSU selection.⁶ The Durbin method selects two first-stage units per stratum without replacement, with probability proportional to size and a known joint probability of inclusion.

In the ECLS-K PSU sample design, there were 24 SR strata and 38 NSR strata. Among the 38 NSR strata, 11 strata were identified as Durbin strata and were treated as SR strata for variance

⁶ Durbin, J. (1967). Design of Multi-Stage Surveys for the Estimation of Sampling Errors. *Journal of the Royal Statistical Society C*, 16, 152-164.

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

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

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

This replicate scheme was used for all of the spring-first grade cross-sectional weights. However, a new feature was added to take into account the freshening process. A child sampled in first grade through the freshening process was assigned to the same replicate as the originally sampled child to whom the child was linked. When the child sampled in first grade was assigned a full sample weight (see section 4.4.5), he or she was assigned the replicate weights in the same manner. To reflect the variability of the control totals in the sample-based raking, a set of replicate control totals was created. Each replicate was then raked to the corresponding replicate-based control totals. This resulted in each replicate retaining the variability associated with the original sample estimates of the control totals.

For fall-first grade cross-sectional weights and the longitudinal weights that include the fall-first grade component, a set of 40 replicates was created using the paired jackknife method. The smaller number of replicates was due to the fact that only 30 percent of the full sample of schools was included in the fall-first grade subsample. The fall-first grade weights do not account for the Durbin method of sampling PSUs, since it no longer applied. Rather, they reflect the fact that only one of the two sampled PSUs in the NSR strata was kept in the subsample. To account for this feature, pairs of similar NSR PSUs were collapsed into 19 variance strata. The SR PSUs account the remaining 21 variance strata. Replicates were formed following the original scheme for the full sample described earlier.

4.5.2 Taylor Series Method

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

Stratum and first-stage unit identifiers are provided as part of the ECLS-K data file and can be used with software such as SUDAAN and STATA. They are described in table 4-12.

Table 4-12.—ECLS-K Taylor Series stratum and first-stage unit identifiers

Variable name	Description
C3TCWSTR	Sampling stratum – fall-first grade C-weights
C3TCWPSU	First-stage sampling unit within stratum – fall-first grade C-weights
C3TPWSTR	Sampling stratum – fall-first grade P-weights
C3TPWPSU	First-stage sampling unit within stratum – fall-first grade P-weights
C4TCWSTR	Sampling stratum – spring-first grade C-weights
C4TCWPSU	First-stage sampling unit within stratum – spring-first grade C-weights
C4TPWSTR	Sampling stratum – spring-first grade P-weights
C4TPWPSU	First-stage sampling unit within stratum – spring-first grade P-weights
C4CPTSTR	Sampling stratum – spring-first grade CPT-weights
C4CPTPSU	First-stage sampling unit within stratum – spring-first grade CPT-weights
C34CSTR	Sampling stratum – fall-first grade/spring-first grade longitudinal C-weights
C34CPSU	First-stage sampling unit within stratum – fall-first grade/spring-first grade longitudinal C-weights
C34PSTR	Sampling stratum – fall-first grade/spring-first grade longitudinal P-weights
C34PPSU	First-stage sampling unit within stratum – fall-first grade/spring-first grade longitudinal P-weights

4.5.3 Specifications for Computing Standard Errors

Specifications for computing standard errors are given in table 4-13. For each type of analysis described in the table, users can choose the replication method or the Taylor Series method for computing standard errors.

Table 4-13.—Specifications for computing standard errors

Type of analysis	Computing standard errors					Approximating sampling errors DEFT (Average root design effect)
	Full sample weight	Replication method (WesVarPC)		Taylor Series method (SUDAAN & STATA)		
		ID	Replicate weights	Jackknife method	Sample design	
Fall-first grade cross-sectional	C3CW0	CHILDID	C3CW1 – C3CW40	JK2	WR*	1.947
	C3PW0	PARENTID	C3PW1 – C3PW40	JK2	C3TCWSTR C3TCWPSU C3TPWSTR C3TPWPSU	
Spring-first grade cross-sectional	C4CW0	CHILDID	C4CW1 – C4CW90	JK2	WR	2.007
	C4PW0	PARENTID	C4PW1 – C4PW90	JK2	WR	
	C4CPTW0	CHILDID	C4CPTW1 – C4CPTW90	JK2	WR WR WR	
First grade longitudinal	C34CW0	CHILDID	C34CW1 – C34CW40	JK2	WR	1.663
	C34PW0	PARENTID	C34PW1 – C34PW40	JK2	WR	

* WR = with replacement

For the replication method using WesVar, the full sample weight, the replicate weights, and the method of replication are required parameters. All analyses of the ECLS-K data should be done using JK2. As an example, to compute fall-first grade child level estimates (e.g., mean reading scores) and their standard errors, users need to specify CHILDID in the ID box of the WesVar data file screen, C3CW0 as the full sample weight, C3CW1 to C3CW40 as the replicate weights, and JK2 as the method of replication. Note that for the fall-first grade weights, there are 40 replicate weights instead of 90 replicate weights for all base year and spring-first grade weights, as explained in section 4.4.7.

For the Taylor Series method using either SUDAAN or STATA, the full sample weight, the sample design, the nesting stratum and PSU variables are required. For the same example above, the full sample weight (C3CW0), the with replacement sample design (WR), the stratum variable (C3TCWSTR) and the PSU variable (C3TCWPSU) must be specified.

The last column in table 4-13 gives the average root design effect that can be used to approximate the standard errors for each type of analysis. For a discussion of the use of design effects, see section 4.6.1.

4.6 Design Effects

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

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

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

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

where SE is the standard error of the estimate.

4.6.1 Use of Design Effects

One method of computing standard errors for the ECLS-K is JK2, as described in section 4.4, using programs designed specifically for analyzing complex survey data such as WesVar. Another approach, Taylor Series linearization (and software designed for it), is also discussed in the same section. If a statistical analysis software package such as SPSS (Statistical Program for the Social Sciences) or SAS (Statistical Analysis System) is used, the standard errors should be corrected using $DEFT$, since these programs calculate standard errors, assuming the data were collected with a simple random sample. The standard error of an estimate under the actual sample design can be approximated as follows:

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

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

Means

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

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.

4.6.2 Average Design Effects for the ECLS-K

In the ECLS-K, a large number of data items were collected from students, parents, teachers, and schools. Each item has its own design effect that can be estimated from the survey data. One way to produce design effects for analysts' use is to calculate them for a number of variables and average them. The averaging can be done overall and for selected subgroups. The tables that follow show estimates, standard errors, and design effects for selected means and proportions based on the ECLS-K first grade child, parent, teacher, and school data. For each survey item, the tables present the number of cases for which data are nonmissing, the estimate, the standard error taking into account the actual sample design (Design SE), the standard error assuming SRS (SRS SE), the root design effect (DEFT), and the design effect (DEFF). Standard errors (Design SE) were produced in WesVar using JK2 based on the actual ECLS-K complex design.

For each survey estimate, the variable name as it appears in the ECLS-K first grade Electronic Code Book is also provided in the table. If multiple variables were combined to arrive at the

estimate, then the names of all the variables used are provided. For example, the estimate of the mean number of days fall-first grade children spent on vacation was computed using two different survey items, P3SUMVAC (parent questionnaire item HEQ100, whether the child went on summer vacation with his or her family) and P3NMDVAC (parent questionnaire item HEQ130, number of days spent on vacation). For more information on the variables used in this section, refer to chapter 3, which describes the assessment and rating scale scores used in the ECLS-K, and chapter 7, which has a detailed discussion of the other variables.

Standard errors and design effects for the child-level items are presented in tables 4-14 to 4-18 for fall-first grade sample, spring-first grade sample, and for children in both the fall-first grade and spring-first grade samples. The survey items were selected so that there was a mix of items common to both fall- and spring-first grade and items that were specific to each round of data collection. For fall- and spring-first grade, the items include the different scores from the direct child assessment, the social rating scores as provided by parents and teachers (spring-first grade only), characteristics of the parents, and characteristics of the students as reported by the parents. For a small number of estimates, the data were subset to cases where the estimate is applicable; for example, the number of days spent on vacation is only for children who took summer vacation, the type of primary child care is only for children who had regular scheduled child care; the number of hours that the mothers work is only for women in the labor force. For the first grade student panel, design effects were calculated for some spring-first grade items from the parent data.

Table 4-14 presents standard errors and design effects for the fall-first grade sample, with a median design effect of 3.9. Table 4-15 presents design effect for the spring-first grade sample, with a median design effect of 4.1. For the panel of students common to both fall- and spring-first grade, the median design effect is 2.5, as shown in table 4-16. This lower median design effect is due to the smaller cluster size—or number of children sampled per school—in the panel.

Table 4-17 presents the median design effects for subgroups based on school type, child's gender and race-ethnicity, geographic region, level of urbanicity, and the socioeconomic scale (SES quintiles) of the parents. For fall-first grade, the median design effects vary from 1.3 (children of other race-ethnicity subgroups) to 11.4 (American Indians). For spring-first grade, the median design effect varies from 1.3 (Pacific Islanders) to 4.9 (children in small towns and rural areas). For the fall-first grade/spring-first grade panel, the range of variability of the median design effects is similar to that for fall-first grade. The variation in the design effects is largely a function of the sample size as well as the

homogeneity of the children within schools. In fall-first grade, the samples of Pacific Islanders and American Indians are very clustered as reflected in the large design effects both for fall-first grade and for the fall-first grade/spring-kindergarten panel.

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

Table 4-14.—ECLS-K, fall-first grade: standard errors and design effects for the full sample

Survey item	Variable name	Number of cases	Estimate	Design SE	SRS SE	DEFT	DEFF
Child scores (mean)							
Reading score	C3RRSCAL	5,053	38.01	0.410	0.178	2.300	5.291
Math score	C3RMSCAL	5,226	32.41	0.280	0.133	2.109	4.448
General knowledge score	C3RGSCAL	5,044	30.05	0.288	0.111	2.587	6.694
Child characteristics (percent)							
Expected to graduate from college	P3EXPECT	5,036	60.06	1.443	0.690	2.091	4.371
Received summer booklist	P3BKLIST	4,981	42.10	1.686	0.700	2.410	5.809
Attended summer school	P3SUMSCH	5,046	10.89	1.204	0.439	2.745	7.534
Took summer vacation	P3SUMVAC	5,045	74.91	1.285	0.610	2.105	4.432
Read book to child everyday	P3RDBKTC	5,023	43.79	0.868	0.700	1.240	1.537
Watched children program	P3CHLPRM	4,968	90.14	0.626	0.423	1.481	2.193
Used computer for education 1-2 times/week	P3COMEDU	5,023	25.20	0.584	0.613	0.953	0.908
Visited museum	P3ARTSCI	5,039	38.89	1.288	0.687	1.876	3.519
Have library in neighborhood	P3LIBRAR	5,040	79.11	1.569	0.573	2.740	7.506
Attended back to school night	P3BTSNGT P3ATTBTS	3,586	70.48	1.195	0.762	1.569	2.463
Child characteristics (mean)							
Age of child in months	R3AGE	5,261	80.01	0.142	0.062	2.298	5.281
BMI	C3BMI	5,044	16.65	0.065	0.036	1.782	3.176
Number of hours in summer school	P3SUMSCH P3SUMSH	511	108.34	5.617	3.480	1.614	2.606
Number of days spend on vacation	P3SUMVAC P3NMDVAC	3,846	11.72	0.319	0.175	1.821	3.315
Number of times visited library during summer	P3VISLIB	5,020	4.19	0.115	0.087	1.326	1.758
Watched TV between breakfast/dinner (hours)	P3TVBRDH	4,919	1.44	0.034	0.020	1.717	2.948
Watched TV after dinner (hours)	P3TVAFDH	4,942	1.03	0.031	0.015	2.125	4.517
Median						1.984	3.945
Mean						1.944	4.015
Standard deviation						0.497	1.921
Coefficient of variation						0.255	0.478
Minimum						0.953	0.908
Maximum						2.745	7.534

Table 4-15.—ECLS-K, spring-first grade: standard errors and design effects for the full sample – Child and parent data

Survey item	Variable name	Number of cases	Estimate	Design SE	SRS SE	DEFT	DEFF
Child scores (mean)							
Reading score	C4RRSCAL	16,336	54.77	0.288	0.111	2.595	6.733
Math score	C4RMSCAL	16,639	42.79	0.196	0.074	2.661	7.083
General knowledge score	C4RGSCAL	16,324	34.00	0.158	0.061	2.604	6.781
Approaches to learning-Parent	P4LEARN	15,574	3.08	0.007	0.004	1.695	2.872
Self-control-Parent	P4CONTRO	15,573	2.93	0.009	0.004	2.113	4.466
Social interaction-Parent	P4SOCIAL	15,581	3.38	0.008	0.005	1.719	2.956
Withdrawn-Parent	P4SADLON	15,569	1.55	0.006	0.003	1.826	3.333
Impulsive/overactive-Parent	P4IMPULS	15,491	1.90	0.011	0.005	2.009	4.035
Approaches to learning-Teacher	T4LEARN	14,986	3.00	0.010	0.006	1.806	3.263
Self-control-Teacher	T4CONTRO	14,871	3.16	0.009	0.005	1.831	3.351
Interpersonal-Teacher	T4INTERP	14,829	3.09	0.009	0.005	1.684	2.836
Externalizing problems-Teacher	T4EXTERN	14,895	1.67	0.009	0.005	1.688	2.849
Internalizing problems-Teacher	T4INTERN	14,809	1.61	0.009	0.005	1.966	3.866
Child characteristics (percent)							
Lived in single parent family	P4HFAMIL	15,624	24.23	0.585	0.343	1.707	2.913
Lived in two-parent family	P4HFAMIL	15,624	73.49	0.634	0.353	1.794	3.219
Mom worked 35 hours+/week	P4HMEMP	11,002	66.37	0.654	0.450	1.452	2.107
Primary case is center-based	P4PRIMNW	6,402	35.60	1.359	0.598	2.271	5.158
Primary case is home-based	P4PRIMNW	6,402	64.40	1.359	0.598	2.271	5.158
Parents had high school or less	WKPARED	15,626	37.69	0.911	0.388	2.349	5.516
Parents attended PTA	P4ATTENP	15,605	40.75	0.905	0.393	2.301	5.294
Did homework 3-4 times per week	P4HMWORK	15,612	39.83	0.927	0.392	2.367	5.604
Parents helped with homework 3-4 times/week	P4HLPHWK	15,100	37.99	0.733	0.395	1.857	3.448
Practiced reading, writing, numbers daily	P4RDWRNM	15,605	52.13	0.764	0.400	1.911	3.651
Visited library	P4LIBRAR	15,597	45.18	0.806	0.399	2.022	4.089
Used computer 1-2 times per week	P4COMPWK	10,389	45.34	0.712	0.488	1.458	2.127
Have family rule for TV	P4TVRULE	15,467	90.31	0.332	0.238	1.397	1.953
HH received foodstamp in last 12 months	P4FSTAMP	15,545	15.93	0.652	0.294	2.221	4.932
Child characteristics (mean)							
Age of child in months	R4_AGE	16,675	87.17	0.078	0.037	2.132	4.547
Child's household size	P4HTOTAL	15,624	4.55	0.024	0.011	2.155	4.646
Number of children <18 in child's HH	P4LESS18	15,624	2.56	0.022	0.010	2.288	5.233
Number of siblings in HH	P4NUMSIB	15,624	1.52	0.019	0.009	2.043	4.173
Number of hours watched TV after dinner	P4TVAFDH	15,445	0.77	0.013	0.006	2.030	4.122
Median						2.016	4.062
Mean						2.007	4.135
Standard deviation						0.332	1.352
Coefficient of variation						0.165	0.327
Minimum						1.397	1.953
Maximum						2.661	7.083

Table 4-16.—ECLS-K, panel: standard errors and design effects for the full sample

Survey item	Variable name	Number of cases	Estimate	Design SE	SRS SE	DEFT	DEFF
Child characteristics (percent)							
Lived in single parent family	P4HFAMIL	4,681	23.31	0.834	0.618	1.350	1.823
Lived in two-parent family	P4HFAMIL	4,681	74.21	0.852	0.639	1.333	1.776
Mom worked 35 hours+/week	P4HMEMP	3,314	65.77	1.057	0.824	1.283	1.645
Primary case is center-based	P4PRIMNW	1,932	36.83	1.587	1.097	1.447	2.093
Primary case is home-based	P4PRIMNW	1,932	63.17	1.587	1.097	1.447	2.093
Parents had high school or less	WKPARED	4,682	36.99	1.017	0.705	1.442	2.079
Parents attended PTA	P4ATTENP	4,677	42.64	1.640	0.723	2.268	5.145
Did homework 3-4 times per week	P4HMWORK	4,676	38.68	1.366	0.712	1.918	3.678
Parents helped with homework 3-4 times/week	P4HLPHWK	4,533	36.81	1.126	0.716	1.572	2.471
Practiced reading, writing, numbers daily	P4RDWRNM	4,676	53.02	1.123	0.730	1.539	2.368
Visited library	P4LIBRAR	4,678	45.76	1.518	0.728	2.084	4.343
Used computer 1-2 times per week	P4COMPWK	3,109	45.55	1.116	0.894	1.249	1.561
Have family rule for TV	P4TVRULE	4,647	90.89	0.433	0.422	1.026	1.052
HH received foodstamp in last 12 months	P4FSTAMP	4,664	17.00	1.184	0.550	2.153	4.634
Child characteristics (mean)							
Age of child in months	R4_AGE	5,023	87.00	0.131	0.063	2.074	4.301
Child's household size	P4HTOTAL	4,681	4.61	0.040	0.021	1.940	3.763
Number of children <18 in child's HH	P4LESS18	4,681	2.58	0.033	0.017	1.944	3.780
Number of siblings in HH	P4NUMSIB	4,681	1.54	0.031	0.017	1.866	3.481
Number of hours watched TV after dinner	P4TVAFDH	4,643	0.79	0.020	0.012	1.663	2.765
Median						1.572	2.471
Mean						1.663	2.887
Standard deviation						0.358	1.207
Coefficient of variation						0.215	0.418
Minimum						1.026	1.052
Maximum						2.268	5.145

Table 4-17.—ECLS-K: median design effects for subgroups – child and parent data

Subgroups	Fall-first grade ¹		Spring-first grade ²		Panel ³	
	DEFT	DEFF	DEFT	DEFF	DEFT	DEFF
All students	1.984	3.945	2.016	4.062	1.572	2.471
Type of school						
Public	1.847	3.414	1.866	3.482	1.451	2.104
Private	1.628	2.654	1.961	3.845	1.797	3.231
Catholic private	1.611	2.593	1.751	3.065	1.418	2.012
Other private	1.586	2.515	1.781	3.173	1.751	3.065
Gender						
Male	1.585	2.513	1.718	2.951	1.385	1.919
Female	1.735	3.011	1.577	2.487	1.509	2.276
Race-ethnicity						
White	1.812	3.284	1.834	3.364	1.594	2.541
Black	1.275	1.626	1.683	2.831	1.361	1.851
Hispanic	1.389	1.928	1.417	2.006	1.152	1.327
Asian	1.518	2.304	1.570	2.466	1.491	2.224
Pacific Islander	2.605	6.785	1.140	1.300	3.022	9.132
American Indian	3.366	11.402	1.283	1.647	3.543	12.550
Other	1.161	1.349	1.364	1.859	1.217	1.480
Region						
Northeast	1.750	3.062	1.688	2.850	1.971	3.883
Midwest	1.887	3.562	2.133	4.550	1.622	2.630
South	1.879	3.553	2.062	4.253	1.767	3.122
West	1.734	3.008	1.707	2.911	1.497	2.242
Urbanicity						
Central city	1.772	3.141	1.839	3.381	1.472	2.166
Urban fringe and large town	1.694	2.868	1.732	3.000	1.691	2.859
Small town and rural area	2.088	4.371	2.210	4.884	1.949	3.800
SES quintiles						
First	1.332	1.774	1.456	2.119	1.283	1.646
Second	1.320	1.742	1.477	2.182	1.155	1.335
Third	1.350	1.822	1.422	2.022	1.361	1.852
Fourth	1.286	1.654	1.435	2.058	1.176	1.383
Fifth	1.406	1.977	1.424	2.027	1.286	1.653

¹ Each median is based on 20 items.² Each median is based on 32 items.³ Each median is based on 19 items.

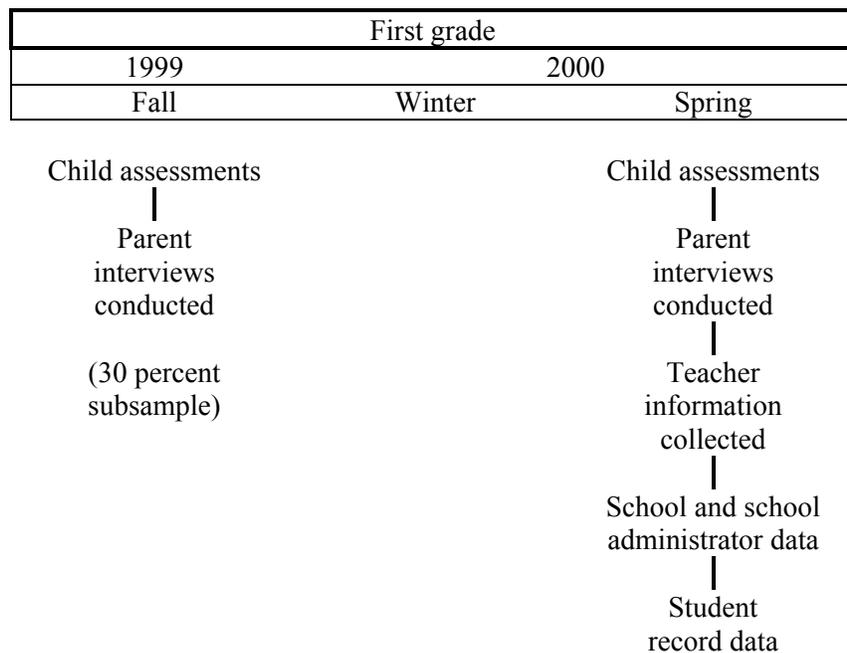
5. DATA COLLECTION METHODS AND RESPONSE RATES

5.1 Data Collection Methods

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

The ECLS-K first grade data collection was conducted in the fall and spring of the 1999-2000 school year. While a full-scale data collection was mounted in spring-first grade, the effort for fall-first grade was limited to a 30 percent subsample of schools. Spring data collection included the direct child assessments, parent interview, teacher and school questionnaires, student record abstract and the facilities checklist. Fall data collection included just the direct child assessment and parent interview. The content and timeline of first grade data collections are shown in exhibit 5-1. Computer-assisted personal interviewing (CAPI) was the mode of data collection for the child assessment, and computer-assisted interviewing (both in-person and telephone) (CAI) was the mode of data collection for the parent interview; self-administered questionnaires were used to gather information from teachers, school administrators, and student records.

Exhibit 5-1.—Timeline of first grade data collection



As in previous rounds of data collection, the field staff was organized into work areas, each with a data collection team consisting of one field supervisor and three assessors. The data collection teams were responsible for all data collection activities in their work areas; they conducted the direct child assessments and the parent interviews, distributed and collected all school and teacher questionnaires, and completed school facilities checklists. Field staff members used in first grade were either continuing from previous rounds of data collection or were new to the project. Training approaches varied depending on the project experience of the field staff—continuing staff members were trained with a self-paced home study training package and new staff members were trained in person in a classroom setting.

5.2 Field Staff Training

All field staff members assigned to fall-first grade were continuing from previous rounds of ECLS-K data collection, so the training for fall-first grade data collection was accomplished using a home study training package. Two different training modes were employed for spring-first grade: one for continuing staff and one for new staff. Continuing staff members were trained via a home study package and new staff were trained in classroom sessions.

5.2.1 Fall-First Grade Field Staff Training

Field supervisors and assessors were trained for the fall-first grade data collection in September 1999, using a home study training program. The purpose of the home study training program was to introduce changes to the instruments between the kindergarten and first grade data collection and provide sufficient review and practice with the instruments and procedures to conduct child assessments and parent interviews. In total, 39 field supervisors and 102 assessors completed this training.

Home Study Training Program. As in the base year training, field supervisor training was conducted using the automated Field Management System (FMS). The FMS was used throughout the data collection period to enter information about the sampled children, parents, teachers, and schools and to monitor production on all data collection activities. The field supervisors entered information into the FMS during training presentations, providing them with hands-on experience with the FMS and all field procedures prior to beginning data collection. The field supervisor home study program was 24 hours long. Topics included conducting the preassessment activities, identifying and locating children who

moved from the school they attended in the previous round of data collection, identifying the regular teachers of ECLS-K children and linking them to those children, and the conduct of quality control observations. The field supervisors followed role plays and completed exercises that involved entering information specifically designed for training purposes into the FMS during the home study training exercises.

Assessors and field supervisors were trained on both the parent interview and child assessment. Child assessment home study materials included written exercises and self-study role plays. Parent interview home study materials included written exercises and practice on the parent interviews in pairs on the telephone with partners assigned by either field supervisors or field managers. Assessors entered information specifically designed for training purposes into the CAPI system on laptop computers using training scripts to complete the home study exercises. The assessor home study program was 16 hours long. Field supervisors ensured that the assessors on their teams were proficient in conducting the direct child assessment by observing a practice assessment with a child of appropriate age. They ensured that the assessors on their teams were proficient in conducting the parent interview by conducting a parent interview role play over the phone at the end of training.

5.2.2 Spring-First Grade Training

Field supervisors and assessors were trained for the spring-first grade data collection over a three-week period in March 2000. The majority of the field staff members were trained primarily using a home study training program. The purpose of the home study training program was to introduce changes to the instruments since the last time these returning staff had participated in ECLS-K (either spring-kindergarten or fall-first grade), provide sufficient practice with the instruments, and provide review and practice of the procedures that were used to conduct the child assessments and parent interviews. An in-person training session was conducted for staff new to the study in the spring. This training session incorporated any changes that were made in the fall-first grade training program using the revised instruments for the spring data collection. In total, 103 field supervisors and 317 assessors completed training.

Home Study Training Program. Field supervisor training was again conducted using the FMS. The field supervisor home study program was 24 hours long. The topics covered in the field supervisor home study training program included conducting the preassessment activities such as

freshening the student sample, identifying and locating children who moved from their previous round school, identifying the regular and special education teachers of ECLS-K children and linking them to those children, distributing teacher questionnaires, distributing school administrator questionnaires, completing the facility checklist, and the conduct of quality control observations. A total of 76 field supervisors completed the home study training.

Assessors and field supervisors were trained on both the parent interview and child assessments. Child assessment home study materials included written exercises, self-study role plays, scripted role plays with partners, a training video focusing on improving assessment skills, and a final practice direct child assessment observed by the field supervisor. Parent interview home study materials included written exercises and trainees practicing the parent interviews in pairs on the telephone and completing a final role play on the parent interview with their field supervisor. Assessors again entered information designed specifically for training purposes into the CAPI system on laptop computers using training scripts to complete the home study exercises. The assessor home study program was 16 hours long. Field supervisors ensured that the assessors on their teams were proficient in conducting the direct child assessment by observing a practice assessment with a child of appropriate age and the parent interview by conducting a final role play over the telephone. A total of 175 assessors completed the home study training (see section 5.6 on data collection quality control).

Training Session for New Field Staff. The in-person training for new field staff (hired because of attrition in prior field staff) was held in March 2000 for nine days. This training incorporated any changes from the fall-first grade training, which were modified for an in-person training. The new spring-first grade parent and child home study materials were also incorporated into the training sessions. A total of 27 field supervisors and 142 assessors completed the in-person training.

5.3 Fall-First Grade Data Collection

The fall-first grade data collection was limited to 26.7 percent of the base year children in 30 percent of the ECLS-K originally sampled schools and was a design enhancement to measure the extent of and factors that contribute to summer learning/loss. See section 4.2 for details on the selection of the subsample. Only the direct child assessments and the parent interviews were included in this data collection.

5.3.1 Advance Mailings

Beginning in late summer 1999, letters were mailed to school administrators to confirm scheduled visits for the schools. A packet of material was also mailed to the school coordinators, who were identified by the school to act as a liaison with ECLS-K field staff, asking them to prepare for the preassessment visit to the school. The school coordinators were asked to distribute materials such as the study brochure, summary sheets describing the role of teachers in the study, and frequently asked questions to the first grade teachers.

5.3.2 Preassessment Contact

Most preassessment contact was made by telephone during September 1999 and at least one week before any assessment visit. During the preassessment contact, the field supervisor spoke with the school coordinator to confirm the dates of the assessment visits, answer any questions, review the list of ECLS-K children who were among the subsample selected for fall-first grade data collection, identify any of these children who were no longer enrolled at the school, collect locating information for those children, identify each enrolled child's regular teacher, review parental consent status and obtain accommodations information about the enrolled sampled children.

Reviewing Parent Consent

Although parental consent was obtained in the base year, field supervisors reviewed the parental consent with the school coordinator to determine if the base year consent was acceptable for fall-first grade. If the schools required consent to be re-obtained or they changed the type of consent that was required (e.g., from implicit to explicit), the supervisors re-obtained consent using the same procedures from the base year.

5.3.3 Conducting the Direct Child Assessment

The direct child assessment was administered during a 12-week field period that began in September and ended in late November 1999. In year-round schools, assessment teams made multiple

visits to the school to conduct direct child assessments as not all children attended school at the same time. The assessment team visited the school when each track was in session to assess the sampled children. The direct child assessment was normally conducted in a school classroom or library. Before conducting the assessment, field supervisors and assessors set up the room for the assessment. They followed procedures for meeting children that were agreed upon during the preassessment contact with the school. Each child was signed out of his or her classroom prior to the assessment and signed back into the classroom upon the conclusion of the assessment

The direct child assessment took approximately 50 to 70 minutes per child. As in the spring-kindergarten data collection, for children with a language other than English in the home, the child's score on the oral language development scale (OLDS) administered in the prior round determined what path the child would follow in fall-first grade. Refer to section 5.4.2 of the ECLS-K Restricted-Use Base Year User's Manual (NCES 2000-097), August 2000, or section 5.4.2 of the ECLS-K Base Year Public-Use User's Manual (NCES 2001-029), February 2001, for more information on how home language was identified and how the OLDS was administered. Table 5-1 summarizes the OLDS routing patterns in fall-first grade.

Table 5-1.—Oral language development scale (OLDS) routing patterns in fall-first grade for previous round respondents

Home language	OLDS score in prior round	OLDS required in fall-first grade?	Fall-first grade OLDS score	Fall-first grade assessment path
English	Not applicable	No		English
Spanish	At or above cut score	No		English
	Below cut score	Yes	At or above cut score Below cut score	English Spanish
Other language	At or above cut score	No		English
	Below cut score	Yes	At or above cut score Below cut score	English Height/weight only

Children passing the cut score for the OLDS were administered the English direct child assessment and had their height and weight measured. Children who fell below the cut score for the OLDS and whose language was Spanish were administered the Spanish language version of the OLDS

and parts of the direct child assessment that were translated into Spanish (the warm-up booklet and math). These children also had their height and weight measured. Children who fell below the cut score for the OLDS and whose language was other than Spanish had only their height and weight measured. Table 5-2 presents the percentage of children who were routed into the various assessment alternatives in fall-first grade. Overall five percent (281 children) of the sampled children were screened using the OLDS in the fall-first grade. Of the children whose home language was Spanish, 23 percent were at or above the cut score, and of the children whose home language was a language other than English or Spanish, 40 percent were at or above the cut score.

Table 5-2.—Oral language development scale (OLDS) routing results in fall-first grade

Category	Total screened (percent)	At or above cut score on OLDS (percent of those screened)	Below cut score on OLDS (percent of those screened)
Total sample	5	26	74
Spanish language children	80	23	77
Other language children	20	40	60

Slightly more than one percent of participating children were excluded from the direct child assessment due to a disability or required an accommodation offered in the assessment. Exclusion from the assessment and the identification and use of accommodations followed the procedures from the base year. Refer to section 5.4.2 of the ECLS-K Restricted-Use Base Year User’s Manual (NCES 2000-097), August 2000 or section 5.4.2 of the ECLS-K Base Year Public-Use User’s Manual (NCES 2001-029), February 2001 for more information on exclusions and identification of accommodations. Table 5-3 presents the number of children excluded from or requiring an accommodation to the direct child assessment.

5.3.4 Conducting the Parent Interview

For the fall-first grade round of data collection, the parent interview was administered between early September and mid-November 1999. The parent interview averaged 35 minutes. To administer the parent interview, assessors began by contacting parents using the contact information obtained during the previous parent interview.

Table 5-3.—Number of children excluded from or accommodated in the fall-first grade assessment

Category	Number of children
Excluded for disability	28
Setting accommodation (e.g. special lighting, adaptive chair)	10
Scheduling/timing accommodation	24
Health care aide present	5
Assistive device used/available	1

The ECLS-K fall-first grade parent interview was conducted primarily by telephone by field staff using CAI. The parent interview was conducted in person if the respondent did not have a telephone. Less than one percent of all completed parent interviews in fall-first grade were conducted in person. The parent interview was conducted primarily in English, but modifications were made to interview parents who spoke only Spanish. The questionnaire was translated into Spanish and printed on a hard-copy parent interview form. Bilingual interviewers were trained to conduct the parent interview in either English or Spanish. When the person answering the telephone was not able to speak English, and the field staff member was not bilingual and was unable to identify an English-speaking household member, the case was coded as a “language problem.” The field supervisor reviewed the case and assigned it to a bilingual field staff person if the language was Spanish. Approximately six percent of the parent interviews were conducted in Spanish. Less than one percent of parent interviews could not be conducted because of language problems, meaning that the respondent spoke a language other than English or Spanish.

5.3.5 Conducting Data Collection on Children Who Withdrew from Their Previous Round School

During the preassessment contacts, field supervisors asked school coordinators to identify children who had withdrawn from the school since the spring of kindergarten. Of the base year participants that had transferred from their base year school, a random sample was identified to be included for data collection in the fall of first grade; see section 4.2 for more details. School staff was asked whether they knew the name and address of the school the child transferred into, as well as any new information about the child’s household address. For the children who had moved from their spring-

kindergarten school and were not part of the sample to be followed, information was collected only from the school personnel and not parents. For children who had withdrawn from their spring-kindergarten school and were identified to be followed (i.e., were part of the sample of movers) supervisors also consulted parents and other contacts for information on the children's new school. This information was entered into the FMS and processed at Westat for data collection.

A total of 1,178 children (18.2 percent of total sample in fall-first grade) were identified as having transferred schools between spring-kindergarten and fall-first grade. In addition, another 273 children were identified as having changed schools between fall-kindergarten and spring-kindergarten. Combining these two types of movers, there was a total of 1,451 children identified as "movers" in fall-first grade. Of the 1,451 mover children in fall-first grade, 671 were followed (46 percent of total movers). The remaining 780 mover children were part of the sample that would not be followed and were not included in the fall-first grade data collection. No child assessment or parent interview was conducted for these children.

Various data collection strategies were used for children who moved, depending on how they were classified. The following data collection approaches were attempted for children who moved and were flagged as "follow" in fall-first grade:

- Parent interviews were attempted for all children regardless of children's mover status.
- Data collected for children moving into cooperating base year sampled schools included the child assessment in the school.
- Data collected for children moving into nonsampled schools in base year cooperating districts included the child assessment conducted in the home.
- For children moving into sample schools that refused, schools in sampled districts that refused, or ineligible sampled schools, only the child assessment was conducted in the home.
- For children moving into schools in nonsampled districts or dioceses:
 - If the school was within the PSU, data collected included the child assessment in the home.
 - If the school was outside the PSU, no child data were collected.

- For children who were not enrolled in school in the spring (including children who were home schooled), data collected included the child assessment in the home, if the child was in the sample PSU. If the child was outside the sample PSU, no data other than the parent interview were collected.

As discussed in section 4.2, a random 50 percent of children who were included in the fall-first grade subsample were flagged to be followed for fall-first grade data collection in the event that they had transferred. Slightly more than one-half (54 percent, 780 children) of the children who moved were not followed and no data were collected for them or their parents, while 46 percent of the children who moved were followed. Of those flagged as “follow,” 11 percent moved into a school outside the PSU, one percent were identified as out of the United States, and eight percent of the movers were unlocatable. One case flagged as “follow” was not fielded because, although the student was located, it was too late in the field period to field the transfer case to conduct a child assessment. This case is labeled *End of field period* in table 5-4. Of the 671 movers who were identified to be included in the fall-first grade data collection, 80 percent were fielded for data collection (i.e., were found and were eligible). Of the 533 cases that were finally located and eligible for data collection, 81 percent had a completed assessment. Table 5-4 presents the status of the 1,451 children who were movers in fall-first grade.

Table 5-4.—Number of children who moved in fall-first grade by completion category

Child in	Number of children	Percent
Total movers	1,451	100
Did not follow ¹	780	54
Followed ¹	671	46
Followed	671	100
Not fielded for assessment ²	138	20
Unlocatable ²	55	8
End of field period ²	1	0
Nonsampled PSU ²	72	11
Moved to outside the U.S. ²	9	1
Deceased ²	1	0
Fielded for assessment ²	533	80
Fielded for assessment	533	100
Completed assessment ³	430	81

¹ Percent based on total movers.

² Percent based on number of movers followed.

³ Percent based on number of movers fielded.

5.4 Fall-First Grade Completion Rates

Table 5-5 presents weighted and unweighted child-level completion rates for the fall-first grade data collection, by school characteristics. On the ECLS-K, a completion rate is a response rate conditioned on the results of an earlier stage of data collection. For the first grade year of the ECLS-K, all completion rates are conditioned on the case having been a base year respondent, since data collection was attempted only for such cases. In fall-first grade, data collection was limited to the child assessment and parent interview. A weighted 90.3 percent of base year respondents completed the child assessment in fall-first grade and 88.6 percent of the children had a completed parent interview. By and large the completion rates are quite uniform across school characteristics. Students enrolled in Catholic schools, in rural schools outside of metropolitan statistical areas (MSA), and in schools with 750 or more enrolled completed the child assessment at higher than average rates (96.1, 96.9, and 95.9 percent respectively). The completion rate for the parent interview was lowest for students in large cities and those in schools with 90 percent to 100 percent minority enrollment (85.0 percent and 83.7 percent). It was highest for students enrolled in Catholic schools, in schools in large towns and in rural schools outside of MSAs (93.4, 94.7, and 94.7 percent respectively).

However, the category labeled “Unknown” in each of the different school characteristics has a substantially lower completion rate (table 5.5). Most of the children in this category are movers. The category includes children who were unlocatable as their whereabouts were unknown, whose cases could not be processed before the end of field period and those children who had moved into a nonsampled county. Under any of these circumstances if no information concerning the child’s school was available, they were included in the “unknown” category for each of the different school characteristics. Table 5-6 gives completion rates by mover status. A full 97 percent of nonmovers completed the child assessment. Movers who were located were assessed at a rate of 76 percent and, the others, including those not located, were not assessed. Even though these children were not administered a child assessment, wherever possible, a parent interview was conducted.

Table 5-7 contains child-level fall-first grade weighted and unweighted completion rates, this time broken out by child characteristics. Again, generally speaking completion rates do not differ greatly. The lowest completion rates for the child assessment are for blacks and for American Indian/Alaska Natives (87.6 and 87.9 percent respectively), the highest for Pacific Islanders (97.6 percent). For the parent interview the lowest completion rates are for Blacks and for Asians (83.4 and 85.1 percent respectively) and the highest for Whites and Pacific Islanders (92.0 and 92.5 percent).

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

School characteristics ¹	Child assessment			Parent interview		
	Completes ²	Completion rates		Completes ³	Completion rates	
		Weighted	Unweighted		Weighted	Unweighted
All children	5,291	90.3	93.6	5,071	88.6	89.8
School type						
Public	4,191	93.9	95.5	3,926	89.0	89.4
Private	1,060	94.8	95.8	1,028	93.0	92.9
Catholic	542	96.1	97.3	522	93.4	93.7
Other private	518	93.6	94.4	506	92.5	92.2
Unknown	40	28.7	26.0	117	73.3	76.0
Type of locale						
Large city	1,033	93.6	95.5	941	85.0	87.0
Mid-size city	882	93.9	96.1	834	89.3	90.8
Urban fringe of large city	1,379	93.2	94.4	1,286	88.4	88.0
Urban fringe of midsized city	296	93.2	94.3	286	90.4	91.1
Large town	255	94.5	97.0	252	94.7	95.8
Small town	513	94.0	97.3	491	90.6	93.2
Rural – Outside MSA	536	96.9	96.8	519	94.7	93.7
Rural – Inside MSA	271	94.3	95.4	266	93.4	93.7
Unknown	126	40.6	51.0	196	75.6	79.4
School size (total enrollment)						
1 to 299	1,246	94.7	95.6	1,178	89.7	90.3
300 to 499	1,420	93.1	95.6	1,355	90.0	91.2
500 to 749	1,327	92.8	94.4	1,262	89.1	89.8
750 or more	1,183	95.9	96.7	1,088	89.3	89.0
Unknown	115	40.9	49.4	188	76.3	80.7

¹ Based on ECLS-K survey data and not data from the sampling frame.

² Reading, math, or general knowledge assessment was scorable or child was language minority (not Spanish) or disabled.

³ Family structure portion of parent interview was completed.

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

School characteristics ¹	Child assessment				Parent interview				
	Completes ²	Completion rates		Completes ³	Completion rates		Completes ³	Completion rates	
		Weighted	Unweighted		Weighted	Unweighted		Weighted	Unweighted
Percent non-White enrolled									
0 - 10%	1,604	95.1	95.9	1,568	93.3	93.8			
11 - 49%	1,635	94.0	95.9	1,572	90.7	92.2			
50 - 89%	1,005	92.2	94.5	926	87.4	87.0			
90 - 100%	888	94.6	95.8	776	83.7	83.7			
Unknown	159	46.8	56.4	229	77.4	81.2			
Region									
Northeast	931	93.4	94.8	869	87.9	88.5			
Midwest	1,299	92.0	94.7	1,232	87.8	89.8			
South	1,758	95.2	96.4	1,676	91.1	91.9			
West	1,263	94.7	95.8	1,177	89.8	89.2			
Unknown	40	28.8	26.1	117	73.8	76.5			

¹ Based on ECLS-K survey data and not data from the sampling frame.

² Reading, math, or general knowledge assessment was scorable or child was language minority (not Spanish) or disabled.

³ Family structure portion of parent interview was completed.

Table 5-6.—ECLS-K fall-first grade: number of completed child-level cases and child-level completion rates, by child's mover status

Mover status ¹	Child assessment				Parent interview	
	Completes ²	Completion rates		Completes ³	Completion rates	
		Weighted	Unweighted		Weighted	Unweighted
All children	5,291	90.3	93.6	5,071	88.6	89.8
Mover status						
Mover	404	64.6	66.4	514	82.5	84.5
Located, followed	404	76.0	78.0	445	83.8	85.9
Other ⁴	0	0.0	0.0	69	74.7	76.7
Nonmover	4,887	97.0	96.9	4,557	90.2	90.4

¹ This is the mover status used in weighting which does not consider children who moved into identified destination schools as movers.

² Reading, math, or general knowledge assessment was scorable or child was language minority (not Spanish) or disabled.

³ Family structure portion of parent interview was completed.

⁴ This category includes movers who could not be located, movers whose cases could not be processed before the end of the field period, and movers who moved into nonsampled PSUs.

Table 5-7.—ECLS-K fall-first grade: number of completed child-level cases and child-level completion rates, by child characteristic

Child characteristics ¹	Child assessment				Parent interview			
	Completes ²	Completion rates		Completes ³	Completion rates		Unweighted	
		Weighted	Unweighted		Weighted	Unweighted		
All children	5,291	90.3	93.6	5,071	88.6	89.8		
Gender								
Male	2,674	89.5	93.3	2,578	89.1	89.9		
Female	2,606	91.0	94.0	2,493	88.6	90.0		
Unknown gender	11	98.4	91.7	0	0.0	0.0		
Race/ethnicity								
White (not Hispanic)	2,949	91.7	94.2	2,912	92.0	93.0		
Black (not Hispanic)	783	87.6	92.2	718	83.4	84.6		
Hispanic	878	89.1	93.3	809	85.6	86.0		
Asian	285	89.8	93.4	260	85.1	85.2		
Pacific Islander	96	97.6	97.0	89	92.5	89.9		
American Indian or Alaskan Native	127	87.9	92.7	126	89.6	92.0		
Other	153	88.8	93.9	150	87.7	92.0		
Unknown race/ethnicity	20	81.6	80.0	7	20.2	28.0		
Year of birth								
1992	1,611	89.1	93.6	1,577	89.6	91.6		
1993	3,642	90.7	93.7	3,472	88.6	89.3		
Other/unknown	38	95.3	92.7	22	47.9	53.7		

¹ Based on ECLS-K survey data and not data from the sampling frame.

² Reading, math, or general knowledge assessment was scorable or child was language minority (not Spanish) or disabled.

³ Family structure portion of parent interview was completed.

5.5 Spring-First Grade Data Collection

All children assessed during the base year, with a completed parent interview, or with a disability were eligible to be assessed in the spring-first grade data collection regardless of whether they repeated kindergarten, were promoted to first grade, or were moved ahead to second grade. Of those children who had transferred from their base year school, only a subsample was followed. In addition, children who were not in kindergarten in the United States during the 1998-99 school year and, therefore, did not have a chance to be selected to participate in the ECLS-K were added to the spring-first grade sample through a process referred to as “freshening.” Freshening ensured that the data allowed for conclusions based on all first grade children, regardless of whether they were enrolled in kindergarten in the United States in fall 1998 or not. To accomplish the goals of following a subset of movers and freshening the sample, a random 50 percent subsample of base year schools were flagged. See section 4.3.2 for further details on how the ECLS-K sample was freshened to be representative of first graders in the United States and how movers were subsampled to be included in the survey.

5.5.1 Advance Contact with Respondents

In February 2000, letters were mailed to school coordinators confirming the scheduled visits to the school that had been arranged in the fall and winter. For those schools flagged for sample freshening, the school coordinators were also sent instructions for preparing the enrollment list of first grade students. Letters were also mailed to parents reminding them of the spring data collection activities.

5.5.2 Preassessment Contact

For schools that were not identified as part of the freshening and mover subsampling process, field supervisors conducted most preassessment activities by telephone starting in March 2000. The preassessment activities for these schools were similar to those conducted in previous rounds of data collection. For schools identified as part of the freshening and mover subsampling process, field supervisors visited each school in order to conduct freshening activities, as well as conduct other preassessment activities, such as identifying children who moved from their previous round school, identifying regular and special education teachers, linking them to children and distributing school

administrator and teacher questionnaires. Section 4.3.2 describes the procedures used for identifying the subsample of schools that were part of the freshening and mover subsampling process.

Freshening Procedures

Each school that was sampled as a school from which children for the freshened sample would be selected, was asked to provide a list of all first grade students. When the field supervisor visited the school for the preassessment visit, he/she verified that the list was current (withdrawn students deleted, newly enrolled students included) and that the list contained no duplicate entries of student names. Once the list was ready, the supervisor identified the children previously sampled for the ECLS-K by highlighting their names on the list. On the freshening roster, the supervisors recorded the name of the first ECLS-K student on the enrollment list and the name of the student listed next if it was not that of another ECLS-K student. Then the supervisor asked the school coordinator to tell him/her if this non-ECLS-K student had been enrolled in kindergarten in the United States in fall 1998 and, based on the school coordinator's response, recorded either "Yes," "No," "Don't Know," or "Refused" for that child on the freshening roster. If the response was "Yes," that child was not added to the sample and the supervisor moved on to the next ECLS-K child on the enrollment list. If the response was "No," "Don't Know," or "Refused," the supervisor added the child from the enrollment list to the freshening roster. The supervisor continued with this process until he/she received a "Yes" response to the question "Was this child enrolled in kindergarten in the United States in fall 1998?" These procedures were repeated for all the ECLS-K sampled children on the enrollment list.

The definition of "enrolled in kindergarten in the United States in fall 1998," which was provided to the school coordinators, was very specific and included the following explanation of terms:

In Kindergarten in U.S. in Fall, 1998? The "In kindergarten in U.S. in fall, 1998?" question was to be taken literally. It was not meant to determine when or where a child was in kindergarten, but whether each currently enrolled first grade child was eligible for selection for the ECLS-K sample. To have been eligible for selection in the ECLS-K sample, the first grade child must have met three criteria. He/she must have been:

1. **In kindergarten:** In the ECLS-K, kindergarten is defined as the year of school primarily for five-year-olds prior to the first grade and includes all types of programs: public, private, full-day, part-day, regular, and transition programs. In addition to

“regular kindergartners” (i.e., the traditional year of school primarily for five-year-olds prior to first grade) found in “regular” public and private schools, the following kindergarten programs should be included:

- Kindergarten programs located in early childhood programs (e.g., nursery schools, early childhood learning centers, daycare centers);
- Transitional programs:
 1. **Transitional (or readiness) Kindergarten**—extra year of school for kindergarten-age-eligible children who are judged not ready for kindergarten.
 2. **Transitional First (or Pre-First) Grade**—extra year of school for children who have attended kindergarten and have been judged not ready for first grade.
- Ungraded/multigraded programs with kindergarten-age children
- Special education programs/classes with kindergarten-age children

If a child was enrolled in any of these types of programs the answer to “In Kindergarten” is “yes.” If the child was in any other type of program or was in a different grade level (e.g., first grade), the answer to “in kindergarten” is “no.”

2. **In United States:** The 50 states of the United States, that is, the 48 contiguous states, Hawaii, Alaska, and the District of Columbia.
3. **In fall 1998:** September through December 1998.

All nonsampled ECLS-K children recorded on the freshening roster for whom the school coordinator responded “No,” “Don’t Know,” or “Refused” to the question “Was this child enrolled in kindergarten in the United States in fall 1998?” were added to the FMS and transmitted to the Westat home office. The supervisor collected contact information on each added child’s parents from the school coordinator, prepared parent information packets, including consent forms, and sent them to the parents of the added children. Meanwhile, the Westat home office created the child assessment and parent interview case and fielded the case for data collection.

Prior to assessing children who were added to the sample through freshening, the supervisor contacted the parent(s) of each freshened child for two reasons: (1) to confirm that the freshened child was not enrolled in kindergarten in the United States in fall 1998, and (2) to obtain outstanding consent forms. If the parent(s) of the freshened child indicated that the child *was* enrolled in kindergarten in fall 1998, no data were collected from the child or the parent. However, if the parent(s) of the freshened child

confirmed that the child *was not* enrolled in kindergarten in fall 1998, the child was included in the ECLS-K first grade sample, and both the child assessment and parent interview were conducted. Data for these freshened children were also collected from the teachers and school records. Of the 210 children who were added to the sample through freshening, 162 (77 percent) were confirmed by the parents as not enrolled in kindergarten in the United States in fall 1998.

5.5.3 Conducting Direct Child Assessments

The direct child assessments were conducted between March and June 2000. For children with a language other than English in the home, the child’s score on the previous OLDS determined what path was followed in the spring. Table 5-8 summarizes the OLDS routing patterns in spring-first grade.

Table 5-8.—Oral language development scale (OLDS) routing patterns in spring-first grade for previous round respondents

Home language	OLDS score in prior round	OLDS required in spring-first grade?	Spring-first grade OLDS score	Spring-first grade assessment path
English	Not applicable	No		English
Spanish	At or above cut score	No		English
	Below cut score	Yes	At or above cut score Below cut score	English Spanish
Other language	At or above cut score	No		English
	Below cut score	Yes	At or above cut score Below cut score	English Height/weight only

Children who scored at or above the cut point on the OLDS in the previous assessment were automatically routed by CAPI to take the assessment in English; the OLDS was not administered again. Children who scored below the cut point in the OLDS in the previous assessment (spring-kindergarten for most, fall-first grade for most of the remaining, and fall-kindergarten for a few) were administered the OLDS again in spring-first grade and routed according to the new spring-first grade OLDS score. Children taking the direct assessment for the first time in the spring, for example students who were included in sample through the freshening process, with a language other than English in the home, were

routed according to their home language as determined from school records or the child’s teacher and were administered the OLDS. Table 5-9 presents the percentage of children who were routed into the various assessment alternatives in spring-first grade. Five percent of the sampled children (798 children) were screened using the OLDS in the spring-first grade. Overall 56 percent of the screened children were at or above the cut score on the OLDS and were administered the English direct child assessment. Of the children whose home language was Spanish, 50 percent were at or above the cut score, and of the children whose home language was a language other than English or Spanish, 77 percent were at or above the cut score.

Table 5-9.—Oral language development scale (OLDS) routing results in spring-first grade

Category	Total screened (percent)	At or above cut score on OLDS (percent of those screened)	Below cut score on OLDS (percent of those screened)
Total sample	5	56	44
Spanish language children	77	50	50
Other language children	23	77	23

Approximately 1.5 percent of participating children were excluded from the direct child assessment or required an accommodation offered in the assessment. The patterns for accommodations were the same as in previous rounds of data collection. Table 5-10 presents the number of children excluded from or requiring an accommodation to the direct child assessment in the spring of kindergarten.

Table 5-10.—Number of children excluded from or accommodated in the spring-first grade assessment

Category	Number of children
Excluded for disability	47
Setting accommodation (e.g. special lighting, adaptive chair)	55
Scheduling/timing accommodation	119
Health care aide present	21
Assistive device	12

5.5.4 Conducting the Parent Interview

Parent interview procedures mirrored those of the fall-first grade. The parent interview was administered, primarily by telephone interview using CAI, between March and early July 2000. The parent interview averaged 45 minutes. As in previous rounds of data collection, the parent interview was conducted in person if the respondent did not have a telephone. Three percent of all completed parent interviews were conducted in person. Six percent of the completed parent interviews were conducted in a language other than English with 94 percent of completed non-English interviews conducted in Spanish. The refusal conversion workshop developed for fall-kindergarten was implemented at the end of the spring field period to improve response rates. The special effort to build parent response rates was conducted between June 5 and July 8, 2000, and yielded an additional ten percentage points in the response rate. Four percent of the parent interviews were not completed because of locating problems.

5.5.5 Teacher and School Data Collection

Data were collected from school administrators, regular classroom teachers, and special education teachers between March and June 2000. Teachers were asked to complete individual ratings for the sampled children in their classrooms, and they were reimbursed seven dollars for each child rating (teacher questionnaire C) they completed. In addition, school staff was asked to complete a student record abstract after the school year closed. The school staff was reimbursed seven dollars for every student record abstract they completed. Field supervisors also completed a facilities checklist for each sampled school.

Field supervisors distributed the school and teacher questionnaires in a variety of ways, depending on the preference of the school staff. Questionnaires were distributed during the preassessment visit (if one was held), by mail, and during the assessment visits. During the field period, field supervisors followed up with school administrators and teachers by telephone and visits to the schools to ensure that completed questionnaires were mailed to Westat. To improve response rates, in early September 2000, a package was mailed to all schools with outstanding school administrator questionnaires, teacher questionnaires, or student record abstracts with a request to complete and return questionnaires. Field staff prompted by telephone for the return of school administrator questionnaires, teacher questionnaires, and student record abstracts through October 2000.

5.5.6 Conducting Data Collection on Children Who Withdrew from Their Previous Round School

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

A total of 3,454 children (16.1 percent of total sample in spring-first grade) were identified as having transferred from the school in which they were enrolled in the previous round of data collection (their spring-kindergarten school if they were not included in the fall-first grade subsample or their fall-first grade school if they were included in the fall subsample). In addition, another 2,125 children were identified as having changed schools between fall kindergarten and spring kindergarten. Combining these two types of movers, there was a total of 5,759 children identified as 'movers' in spring-first grade. Of the 5,759 mover children in spring-first grade, 2,911 were followed (51 percent of total movers). The remaining 2,848 mover children were part of the subsample that would not be followed and were not included in the spring-first grade data collection. No child assessment or parent interview was conducted for these children.

Different data collection strategies were followed for children who moved, depending on how they were classified. Data collection was attempted for children who moved and were flagged as "follow" in spring-first grade:

- Parent interviews were attempted for all children regardless of children's mover status.
- Data collected for children moving into cooperating base year sampled schools included the child assessment in the school, school administrator questionnaire, regular or special education teacher questionnaires, facilities checklist, and student record abstract forms.

- Data collected for children moving into nonsampled schools in base year cooperating districts included the child assessment in the school, school administrator questionnaire, regular or special education teacher questionnaires, and student record abstract forms if school permission was obtained. If school permission was not obtained, the assessment was conducted in the home and no school or teacher data were collected.
- For children moving into sample schools that refused, schools in sampled districts that refused, or originally sampled schools that were ineligible when sampled because they did not have kindergarten classes, only the direct child assessment was conducted in the home. No school or teacher data were collected.
- For children moving into schools in nonsampled districts or dioceses:
 - If the school was within the PSU, data collected included the child assessment in the school, school administrator questionnaire, regular or special education teacher questionnaires, facilities checklist, and student record abstract forms if school permission was obtained. If school permission was not obtained, the assessment was conducted in the home and no school or teacher data were collected.
 - If the school was outside the PSU, no child, school, or teacher data were collected.
- For children who were not enrolled in school in the spring (including children who were home schooled), data collected included the child assessment in the home if the child was in the sample PSU. If the child was outside the sample PSU, no data were collected.

As discussed in section 4.3.1, all children in a random 50 percent subsample of base year schools were flagged to be followed for spring-first grade data collection in the event that they had transferred. Slightly less than half (49 percent, 2,848 children) of the children who moved were not followed and no data were collected for them or their parents, while 51 percent of the children who moved were followed. Of those flagged as “follow,” 16 percent moved into a school outside the PSU, two percent were identified as out of the United States, and nine percent of the movers were unlocatable. Another 2 percent were not fielded because, although the students were located, it was too late in the field period to field the transfer cases to conduct a child assessment. These cases are labeled *End of field period* in table 5-11. Of the 2,911 movers who were identified to be included in the spring-first grade data collection, 71 percent were fielded for data collection (i.e., were found and were eligible). Of the 2,070 cases that were finally located and eligible for data collection, 87 percent had a completed assessment. Table 5-11 presents the status of the 5,759 children who were movers in spring-first grade.

5.6 Data Collection Quality Control

A continuous quality assurance process was applied to all data collection activities, but with a particular focus on the assessments. The process was incorporated in all stages (i.e., during development, in the staff training program, through certification, and as part of the ongoing staff observations and evaluation activities).

Table 5-11.—Number of children who moved in spring-first grade by completion category

Child in	Number of children	Percent
Total movers	5,759	100
Did not follow ¹	2,848	49
Followed ¹	2,911	51
Followed	2,911	100
Not fielded for assessment ²	840	29
Unlocatable ²	271	9
End of field period ²	70	2
Nonsampled PSU ²	454	16
Moved to outside the U.S. ²	44	2
Deceased ²	1	0
Fielded for assessment ²	2,070	71
Fielded for assessment	2,070	100
Completed assessment ³	1,792	87

¹ Percent based on total movers.

² Percent based on number of movers followed.

³ Percent based on number of movers fielded.

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

Quality control activities continued during training and data collection. During home study training, field managers certified supervisors and supervisors certified assessors on the parent interview and the child assessment. During the in-person assessor training, field staff practiced conducting the parent interview in pairs and practiced the direct child assessment with kindergarten, first, and second grade children brought to the training site for this purpose. When the fieldwork began, field supervisors observed each assessor conducting child assessments and made telephone calls to parents to validate the interview. Field managers made telephone calls to the schools to collect information on the school activities for validation purposes. A sample of the assessor-completed OLDS score sheets was rescored in the home office for quality control purposes.

5.6.1 Child Assessment Observations

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

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

- Rapport building and working with the child—use of neutral praise and the assessor's response to various child behaviors.
- Cognitive assessment activities—reading questions verbatim, the use of acceptable probes, the use of appropriate hand motions, and the absence of coaching.
- Specific assessment activities—correctly coding answers to open-ended questions in the assessment, weighing and measuring the child correctly, and following administration procedures.

The field supervisors recorded their observations on the form and then reviewed the form with the assessor. The most frequent problems observed were not reading the items verbatim and inappropriate gesturing. Feedback was provided to the assessors on the strengths and weaknesses of their performance and, when necessary, remedial training was provided in areas of weakness.

5.6.2 Parent Validations

Approximately ten percent of the completed parent interviews were validated. The first parent interview completed by an assessor was always validated. Over the course of the field period, a running count of an assessor's completed parent interviews was maintained, and each tenth completed parent interview was selected for validation. This ensured that ten percent of each assessor's cases were selected for validation. The parent validation was approximately five minutes long, conducted by telephone.

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

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

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

5.6.3 School Validations

To ensure that assessments proceeded smoothly, a validation call was completed with the school principal in at least two of each supervisor's assigned originally sampled schools in both the fall- and spring-first grade collections.

Field managers conducted the school validations by telephone. The first school that each team completed was called to ascertain how well the preassessment and assessment activities went. If the feedback from the school was positive, the fifth school that each team completed was called. If any problems were indicated in the first validation call, immediate action was taken with the field supervisor. The validation feedback was discussed with the supervisor and remedial action was provided, including in-person observation of the supervisor's next school if necessary. In fall-first grade, a total of 72 or 23 percent of the fielded 310 originally sampled schools were validated. In spring-first grade, a total of 198 or 20 percent of the fielded 970 originally sampled schools were validated.

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

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

No problems were encountered during the school validation process.

5.6.4 Quality Control of the Oral Language Development Scale Scoring

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

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

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

5.6.5 Assessor Effects

Individual Test Administrator Effects and Design Effects

In the base year, a multilevel analysis¹ was carried out to estimate components of variance in fall- and spring-kindergarten direct cognitive scores associated with the (1) students, (2) schools, (3) team leaders, and (4) individual test administrators. A similar analysis was conducted using the spring-first grade direct cognitive scores. This secondary analysis was motivated by Westat's earlier finding of larger-than-expected design effects. In addition, the impact on the above sources of variance of the SES indicator (parent's education) was also estimated. It was expected that much of the clustering of students within neighborhood schools (hence higher design effects) could be explained by SES. To examine whether this held true for first grade, a similar analysis was conducted.

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

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

¹ Bryk, A. & Raudenbush, S.W. (1992). *Hierarchical Linear Models: Applications and data analysis methods*. New York: Sage Publications.
Snijders, T. & Bosker, R. (1999). *Multilevel Analysis – An introduction to basic and advanced multilevel modeling*. London: Sage Publications.

The following table presents information on the intra-class correlations and design effects for the Item Response Theory (IRT) scaled measures and also for the unscaled routing test in reading and mathematics. The intra-class correlations and design effects are presented for the unscaled routing tests in order to evaluate how much the IRT scaling may be contributing to the design effects. Inspection of table 5-12 indicates that the intra-class correlations are consistently albeit trivially lower for the unscaled routing tests as compared to the IRT scaled counterpart tests. The slight difference observed are probably due to the shrinking of the tails that occurs in IRT scoring since they are estimates of “true scores.”

Table 5-12.—Intra-class correlations and estimated design effects for the spring-first grade direct cognitive tests (unweighted)

Cognitive test	Level one variances (student)	Level two variances (school)	Intra-classroom correlations	Estimated design effect
Reading IRT	.318	.101	.241	2.75
Reading Routing	10.022	2.839	.221	2.61
Mathematics IRT	.503	.150	.230	2.67
Mathematics Routing	6.848	1.826	.210	2.53
General Knowledge IRT	.217	.112	.341	3.48
Highest Parent Education	.852	.443	.342	3.40

The intra-class correlation is also presented for the background variable “highest parental education” as an indicator of the potential for clustering with respect to cognitive-related variables that is likely to occur in neighborhood schools.

The estimated intra-class correlations and design effects for the cognitive tests in table 5-12 were computed on the cross-sectional spring-first grade sample using a two-level model without the sample weights. A t-test was done to examine the effects on the intra-class correlation of not using the weights. The test showed that weights and unweighted analysis gave similar results.

Table 5-13 presents the results of a three-level “null” model with student at level one, interviewer at level two, and “work area” at level three. It was suggested that the training and supervision of interviewers by a work area leader (the field supervisor) might lead to a clustering effect by work area. This model was used to examine the potential impact of the interviewer and field supervisor on the data collection process and subsequently on the intra-class correlations. Inspection of table 5-13 suggests that

the interviewer effects in spring-first grade are relatively trivial and should not have any systematic impact on analysis. Similar conclusions apply to the work area effect also.

Table 5-13.—Components of variance for three-level model, including the interviewer effect for spring-first grade cognitive scores

Cognitive tests	Level one (student)	Level two (interviewer)	Level three (work area)
Reading (IRT)	.381 (94%)	.007 (2%)	.019 (4%)
Mathematics (IRT)	.587 (92%)	.011 (2%)	.039 (6%)
General Knowledge (IRT)	.276 (85%)	.012 (4%)	.036 (11%)

5.7 Spring-First Grade Completion Rates

In the sections that follow, spring-first grade completion rates are presented for three groups of students: (1) students sampled in kindergarten, (2) students sampled in first grade through the freshening procedure, and (3) both groups combined.

5.7.1 Students Sampled in Kindergarten

Table 5-14 presents weighted and unweighted child-level completion rates for spring-first grade data collection, broken out by school characteristics. These rates pertain to children who were sampled as part of the kindergarten cohort in the base year. (Rates for students sampled in first grade through the student sample freshening procedure can be found in table 5-18.) Relative to fall-first grade the overall completion rates for the child assessment (88.0 percent) and the parent interview (84.5 percent) dropped two and four points respectively in spring-first grade. The drop in overall child assessment completion rates is tied to lower completion rates for movers, particularly those moving outside of the sampled PSUs, which accounts for approximately 30 percent (unweighted) of the movers' assessment nonresponse. Nearly half (45 percent) of the nonresponse for the child assessment is associated with those children for whom the school type was "Unknown" in the base year (see section 5.4). For the vast majority of students the child assessment completion rate increased slightly between fall- and spring-first grade, and this is true for nearly all school characteristics. The decrease in

Table 5-14.—ECLS-K spring-first grade: number of completed child-level cases and child-level completion rates, by school characteristic—children sampled in the base year

School characteristics ¹	Child assessment				Parent interview				School administrator questionnaire			
	Completes ²		Completion rates		Completes ³		Completion rates		Completes ⁴		Completion rates	
	Completed	Unweighted	Weighted	Unweighted	Completed	Unweighted	Weighted	Unweighted	Completed	Unweighted	Completed	Unweighted
All school types	16,593	91.8	88.0	91.8	15,522	84.5	85.8	85.8	14,619	76.3	81.4	81.4
School type												
Public	13,259	95.3	94.7	95.3	11,957	85.9	86.0	86.0	11,652	81.6	83.8	83.8
Private	3,273	95.1	94.7	95.1	3,125	91.3	90.8	90.8	2,967	82.2	86.2	86.2
Catholic	1,960	97.3	97.4	97.3	1,847	92.6	91.7	91.7	1,828	88.1	90.7	90.7
Other private	1,313	92.0	91.3	92.0	1,278	89.5	89.6	89.6	1,139	74.6	79.8	79.8
Unknown school type	61	8.3	7.5	8.3	440	59.4	60.1	60.1	0	0.0	0.0	0.0
Type of locale												
Large city	2,965	93.6	92.2	93.6	2,524	78.6	79.7	79.7	2,422	72.3	76.5	76.5
Mid-size city	2,925	96.0	94.9	96.0	2,687	87.2	88.2	88.2	2,585	80.4	84.8	84.8
Urban fringe of large city	4,595	94.5	94.4	94.5	4,209	86.7	86.6	86.6	4,110	81.6	84.5	84.5
Urban fringe of mid-size city	1,052	95.0	94.0	95.0	956	84.9	86.4	86.4	906	77.6	81.8	81.8
Large town	442	96.9	96.3	96.9	422	91.0	92.5	92.5	408	88.8	89.5	89.5
Small town	1,358	97.0	97.0	96.9	1,278	91.1	91.2	91.2	1,254	88.8	89.5	89.5
Rural – Outside MSA	1,853	97.0	97.0	97.0	1,757	92.5	92.0	92.0	1,792	92.9	93.8	93.8
Rural – Inside MSA	1,069	94.0	94.0	95.2	1,006	88.1	89.6	89.6	998	89.4	88.9	88.9
Unknown	334	23.1	23.1	33.1	683	64.1	67.6	67.6	144	9.2	16.2	16.2
School size (total enrollment)												
1 to 299	3,777	93.9	93.9	95.0	3,536	87.6	88.9	88.9	3,457	83.7	86.9	86.9
300 to 499	4,831	95.4	95.4	95.9	4,481	88.8	89.0	89.0	4,294	82.7	85.2	85.2
500 to 749	4,261	93.7	93.7	94.4	3,862	85.5	85.6	85.6	3,890	83.1	86.2	86.2
750 or more	3,491	95.5	95.5	95.6	3,049	84.1	83.5	83.5	2,978	80.9	81.6	81.6
Unknown	233	18.2	18.2	25.7	594	62.8	65.5	65.5	0	0.0	0.0	0.0

¹ Based on ECLS-K survey data and not data from the sampling frame.

² Reading, math, or general knowledge assessment was scorable or child was language minority (not Spanish) or disabled.

³ Family structure portion of parent interview was completed.

⁴ A completed questionnaire was defined as one that was not completely left blank.

Table 5-14.—ECLS-K spring-first grade: number of completed child-level cases and child-level completion rates, by school characteristic—children sampled in the base year (continued)

School characteristics ¹	Child assessment				Parent interview				School administrator questionnaire			
	Completes ²		Completion rates		Completes ³		Completion rates		Completes ⁴		Completion rates	
	Completed	Unweighted	Weighted	Unweighted	Completed	Unweighted	Weighted	Unweighted	Completed	Unweighted	Weighted	Unweighted
Percent non-White enrolled												
0–10%	5,320	96.0	95.9	96.0	5,102	91.9	92.1	92.1	4,879	87.0	88.1	88.1
11–49%	5,264	95.8	94.9	95.8	4,858	87.5	88.4	88.4	4,729	83.8	86.0	86.0
50–89%	2,901	93.1	92.9	93.1	2,561	83.1	82.2	82.2	2,508	78.8	80.5	80.5
90–100%	2,779	95.0	93.9	95.0	2,315	78.8	79.2	79.2	2,425	77.4	82.9	82.9
Unknown	329	32.6	25.3	32.6	686	65.5	68.0	68.0	78	7.8	8.8	8.8
Region												
Northeast	3,057	95.3	94.3	95.3	2,782	85.9	86.7	86.7	2,634	79.3	82.1	82.1
Midwest	4,239	96.3	95.3	96.3	3,921	88.3	89.1	89.1	3,682	80.2	83.7	83.7
South	5,535	95.5	95.5	95.5	5,068	86.9	87.5	87.5	5,072	86.0	87.5	87.5
West	3,700	93.5	92.5	93.5	3,310	84.0	83.6	83.6	3,231	77.3	81.6	81.6
Unknown	62	8.6	7.7	8.6	441	60.3	61.2	61.2	0	0.0	0.0	0.0

¹ Based on ECLS-K survey data and not data from the sampling frame.

² Reading, math, or general knowledge assessment was scorable or child was language minority (not Spanish) or disabled.

³ Family structure portion of parent interview was completed.

⁴ A completed questionnaire was defined as one that was not completely left blank.

Table 5-14.—ECLS-K spring-first grade: number of completed child-level cases and child-level completion rates, by school characteristic—children sampled in the base year (continued)

School characteristics ¹	Teacher questionnaire A			Teacher questionnaire B			Teacher questionnaire C		
	Completes ⁴	Completion rates		Completes ⁴	Completion rates		Completes ⁴	Completion rates	
		Weighted	Unweighted		Weighted	Unweighted		Weighted	Unweighted
All school types	15,021	78.1	83.6	14,878	77.5	82.8	14,991	78.0	83.5
School type									
Public	11,883	83.1	85.4	11,795	82.6	84.8	11,875	83.1	85.4
Private	3,138	87.2	91.2	3,083	86.0	89.6	3,116	86.6	90.5
Catholic	1,912	92.9	94.9	1,910	92.7	94.8	1,894	92.4	94.0
Other private	1,226	80.0	85.9	1,173	77.5	82.2	1,222	79.2	85.6
Unknown school type	0	0.0	0.0	0	0.0	0.0	0	0.0	0.0
Type of locale									
Large city	2,502	73.3	79.0	2,448	71.1	77.3	2,519	73.9	79.5
Mid-size city	2,705	84.7	88.7	2,658	83.9	87.2	2,700	84.8	88.6
Urban fringe of large city	4,195	82.9	86.3	4,140	81.9	85.2	4,150	82.1	85.4
Urban fringe of mid-size city	873	77.0	78.9	892	78.5	80.7	881	77.3	79.7
Large town	440	95.7	96.5	441	95.7	96.7	437	94.4	95.8
Small town	1,300	91.2	92.8	1,297	91.5	92.6	1,294	90.9	92.4
Rural – Outside MSA	1,770	92.8	92.8	1,774	93.2	93.0	1,784	93.7	93.5
Rural – Inside MSA	1,009	88.5	89.8	1,000	87.9	89.0	1,000	87.2	89.0
Unknown	227	14.3	25.5	228	14.4	25.6	226	14.4	25.4
School size (total enrollment)									
1 to 299	3,417	82.0	85.9	3,382	81.7	85.1	3,429	82.8	86.2
300 to 499	4,497	85.8	89.3	4,434	84.5	88.1	4,483	85.6	89.0
500 to 749	3,805	81.9	84.3	3,757	81.3	83.2	3,772	81.0	83.6
750 or more	3,173	85.6	86.9	3,174	85.4	87.0	3,178	85.5	87.1
Unknown	129	7.7	16.4	131	7.9	16.6	129	7.8	16.4

¹ Based on ECLS-K survey data and not data from the sampling frame.

² Reading, math, or general knowledge assessment was scorable or child was language minority (not Spanish) or disabled.

³ Family structure portion of parent interview was completed.

⁴ A completed questionnaire was defined as one that was not completely left blank.

Table 5-14.—ECLS-K spring-first grade: number of completed child-level cases and child-level completion rates, by school characteristic—children sampled in the base year (continued)

School characteristics ¹	Teacher questionnaire A			Teacher questionnaire B			Teacher questionnaire C			
	Completes ⁴	Completion rates		Completes ⁴	Completion rates		Completes ⁴	Completion rates		
		Weighted	Unweighted		Weighted	Unweighted		Weighted	Unweighted	
Percent non-White enrolled										
0 – 10%	4,898	86.5	88.4	4,899	86.7	88.4	4,905	86.7	88.5	
11 – 49%	4,931	86.8	89.8	4,854	85.9	88.4	4,869	85.4	88.6	
50 – 89%	2,586	80.7	83.0	2,523	79.1	81.0	2,563	80.3	82.3	
90 – 100%	2,389	77.2	81.7	2,383	76.6	81.5	2,437	79.1	83.3	
Unknown	217	16.1	24.4	219	16.3	24.6	217	16.3	24.4	
Region										
Northeast	2,695	80.6	84.0	2,711	81.3	84.5	2,723	81.3	84.9	
Midwest	3,792	82.4	86.2	3,776	82.1	85.9	3,791	82.6	86.2	
South	5,204	87.7	89.8	5,130	86.7	88.5	5,171	86.9	89.2	
West	3,330	79.9	84.1	3,261	78.4	82.4	3,306	79.8	83.5	
Unknown	0	0.0	0.0	0	0.0	0.0	0	0.0	0.0	

¹ Based on ECLS-K survey data and not data from the sampling frame.

² Reading, math, or general knowledge assessment was scorable or child was language minority (not Spanish) or disabled.

³ Family structure portion of parent interview was completed.

⁴ A completed questionnaire was defined as one that was not completely left blank.

parent interview rates, on the other hand, is widespread and cannot be attributed to one or two anomalous reporting categories. It is most pronounced for large cities where the 78.6 percent completion rate is 6.4 points down from the fall.

The overall weighted completion rate for the school administrator questionnaire is 76.3 percent. This is lower than the overall rate for the SAQ in the base year by close to 10 percentage points. The extremes range from 72.3 percent for children in large cities to 88.8 percent or higher for those in large towns, small towns, or rural areas. The lower rates for the SAQ in the first-grade year are due to movers. The transfer schools that received movers were given an SAQ designed for schools that were not already in the ECLS-K sample. The completion rate for these *new* SAQs is only 34.7 percent compared with 88.6 percent for *returning* SAQs given to schools that are in the original school sample. If not for the movers, the completion rate for the SAQ is in fact slightly higher than the rate for the base year (88.6 percent versus 85.9 percent). It is worth noting that the completion rates for the SAQ continue to be lower for schools with higher percentage minorities. In fact, as the percent minority enrolled in the schools increases, the completion rate for the SAQ decreases. However, this disparity decreased considerably in the first-grade year compared to the base year, reflecting increased data collection efforts targeted toward their schools.

All three of the teacher questionnaires were completed at an overall rate of approximately 78 percent. Often the rates are substantially higher, over 90 percent for Catholic schools and for the less urban areas. At the other end of the spectrum are schools in large cities which had completion rates in the low 70s, and schools with 90 percent or more minority enrollment which had rates in the mid to upper 70s.

The rate at which these survey instruments were completed varies markedly by mover status and within movers, by whether or not the child was located and followed. As presented in table 5-15 the completion rate for the child assessment was 95.9 percent for children still enrolled in their base year school. For movers it dropped over ten points to 85.5 percent for those who were located and followed, and for those not located or followed due to a move to a non-ECLS-K PSU, it was zero. The parent interview completion rates varied from 88.0 percent for nonmovers to 78.2 percent for movers who were located and followed for the purposes of the child assessment, to 60.4 percent for movers who could either not be located or were not followed for the purposes of the child assessment. Even though children who had moved to a non-ECLS-K PSU were not administered the child assessments, wherever possible, a

Table 5-15.—ECLS-K spring-first grade: number of completed child-level cases and child-level completion rates, by child’s mover status—children sampled in the base year

	Child assessment			Parent interview			School administrator questionnaire			
	Mover status ¹	Completion rates		Completes ³	Completion rates		Completes ⁴	Completion rates		
		Completes ²	Weighted		Unweighted	Weighted		Unweighted	Weighted	Unweighted
All school types		16,593	88.0	91.8	15,522	84.5	85.8	14,619	76.3	81.4
Mover status										
Mover		1,519	63.1	64.1	1,756	73.5	74.1	778	35.3	34.6
Located, followed		1,519	85.5	85.1	1,397	78.2	78.2	778	48.7	46.7
Other ⁵		0	0.0	0.0	359	60.4	61.5	0	0.0	0.0
Nonmover		15,074	95.9	95.9	13,766	88.0	87.6	13,841	88.7	88.1

¹ This is the mover status used in weighting, which does not consider children who moved into identified destination schools as movers.

² Reading, math or general knowledge assessment was scorable or child was language minority (not Spanish) or disabled.

³ Family structure portion of Parent Interview was completed.

⁴ A completed questionnaire was defined as one that was not completely left blank.

⁵ This category includes movers who could not be located, movers whose cases could not be processed before the end of the field period, and movers who moved into nonsampled PSUs.

Table 5-15.—ECLS-K, spring-first grade: number of completed child-level cases and child-level completion rates, by child’s mover status—children sampled in the base year (continued)

	Teacher questionnaire A			Teacher questionnaire B			Teacher questionnaire C			
	Mover status ¹	Completion rates		Completes ⁴	Completion rates		Completes ⁴	Completion rates		
		Completes ²	Weighted		Unweighted	Weighted		Unweighted	Weighted	Unweighted
All school types		15,021	78.1	83.6	14,878	77.5	82.8	14,991	78.0	83.5
Mover status										
Mover		856	38.5	38.0	844	38.1	37.5	858	38.5	38.1
Located, followed		856	53.2	51.4	844	52.6	50.7	858	53.2	51.5
Other ⁵		0	0.0	0.0	0	0.0	0.0	0	0.0	0.0
Nonmover		14,165	90.1	90.2	14,034	89.5	89.3	14,133	90.0	90.0

¹ This is the mover status used in weighting, which does not consider children who moved into identified destination schools as movers.

² Reading, math, or general knowledge assessment was scorable or child was language minority (not Spanish) or disabled.

³ Family structure portion of parent interview was completed.

⁴ A completed questionnaire was defined as one that was not completely left blank.

⁵ This category includes movers who could not be located, movers whose cases could not be processed before the end of the field period, and movers into nonsampled PSUs.

parent interview was conducted by telephone leading to the 60.4 percent response rate for this category. The SAQ showed a 40-point drop in completion rate for movers, even when they were located and followed. There are several reasons for this drop: located movers were not always assessed in schools; new schools in which movers enrolled had a lower level of commitment to the ECLS-K and often refused to complete the SAQ; and some of these schools were contacted too late in the school year for them to consider completing it. The completion rate for nonmovers was 88.7 percent. For located and followed movers it was 48.7 percent. For all three teacher questionnaires the completion rates were approximately 90 percent if the child had not moved, about 53 percent if the child moved, was located, and followed, and 0 if not located or followed. The reasons for lower completion rates from teachers if the child had moved are similar to the reasons that affected the SAQ completion rates for movers.

Table 5-16 presents child-level completion rates for the spring-first grade data collection, this time broken out by child characteristics for children who were sampled as part of the kindergarten cohort in the base year. The differences in completion rates by gender and by year of birth are inconsequential but for race and ethnicity they are more substantial. For the child assessment the completion rate was highest for Asians and Pacific Islanders (90.7 percent and 89.6 percent respectively) and lowest for Native Americans (84.0 percent). For the parent interview it was highest for whites (88.6 percent), lowest for Asians (73.2 percent), and second lowest for black students (77.5 percent). The low rate for Asians is partly due to language problems; 7.3 percent (unweighted) of parent interviews for Asian children were nonresponse because of language. The ECLS-K sample of Pacific Islanders is very clustered and has unusually high completion rates for the instruments filled out by school personnel, 87.1 percent for the school administrator questionnaire and over 86 percent for each of the teacher questionnaires. The lowest completion rate for the school administrator questionnaire is for Native Americans (68.6 percent). For the teacher questionnaires the lowest rates are in the 73 to 75 percent range and are associated with blacks and Hispanics. Since almost 70 percent of the black and Hispanic students are enrolled in high minority schools (50 percent or higher), this may be associated with lower levels of response for the SAQ from high minority schools. Of the 21.6 percent (unweighted) of black and Hispanic students with no SAQ data, 20 percent are enrolled in high minority schools.

In addition to the child assessment, parent interview, teacher questionnaires, and school administrator questionnaires whose completion rates have been summarized in the preceding tables, various other types of data were collected during spring-first grade as well. Table 5-17 presents counts of completes and weighted and unweighted completion rates at the overall

Table 5-16.—ECLS-K spring-first grade: number of completed child-level cases and child-level completion rates, by child characteristic—children sampled in the base year

Child characteristics ¹	Child assessment			Parent interview			School administrator questionnaire		
	Completes ²	Completion rates		Completes ³	Completion rates		Completes ⁴	Completion rates	
		Weighted	Unweighted		Weighted	Unweighted		Weighted	Unweighted
All school types	16,593	88.0	91.8	15,522	84.5	85.8	14,619	76.3	81.4
Gender									
Male	8,476	87.7	91.7	7,937	84.2	85.9	7,464	76.2	81.3
Female	8,106	88.3	91.8	7,585	84.9	85.9	7,144	76.3	81.4
Unknown gender	11	98.4	91.7	0	0.0	0.0	11	98.4	91.7
Race/ethnicity									
White (not Hispanic)	9,401	88.6	92.1	9,214	88.6	90.3	8,477	79.2	83.5
Black (not Hispanic)	2,349	87.3	90.5	2,051	77.5	79.0	2,023	72.5	78.7
Hispanic	2,838	87.2	91.3	2,559	81.0	82.3	2,391	72.0	77.7
Asian	1,071	90.7	93.5	838	73.2	73.1	919	76.3	80.6
Pacific Islander	188	89.6	93.1	165	81.9	81.7	184	87.1	91.5
American Indian or Alaska Native	296	84.0	89.2	288	82.9	86.7	247	68.6	75.5
Other	397	85.2	91.5	385	86.7	88.7	333	71.4	77.4
Unknown race/ethnicity	53	77.9	84.1	22	32.8	34.9	45	65.6	72.6
Year of birth									
1992	4,863	88.2	92.0	4,571	84.4	86.5	4,320	77.5	82.3
1993	11,616	87.9	91.7	10,856	84.7	85.7	10,197	75.8	81.0
Other/unknown	114	83.0	86.4	95	65.7	72.0	102	74.6	79.1

¹ Based on ECLS-K survey data and not data from the sampling frame.

² Reading, math or general knowledge assessment was scorable or child was language minority (not Spanish) or disabled.

³ Family structure portion of parent interview was completed.

⁴ A completed questionnaire was defined as one that was not completely left blank.

Table 5-16.—ECLS-K spring-first grade: number of completed child-level cases and child-level completion rates, by child characteristic—children sampled in the base year (continued)

Child characteristics ¹	Teacher questionnaire A			Teacher questionnaire B			Teacher questionnaire C		
	Completes ⁴	Completion rates		Completes ⁴	Completion rates		Completes ⁴	Completion rates	
		Weighted	Unweighted		Weighted	Unweighted		Weighted	Unweighted
All school types	15,021	78.1	83.6	14,878	77.5	82.8	14,991	78.0	83.5
Gender									
Male	7,648	77.6	83.3	7,569	77.0	82.5	7,631	77.4	83.2
Female	7,373	78.7	84.0	7,309	78.1	83.3	7,360	78.7	83.9
Unknown gender	0	0.0	0.0	0	0.0	0.0	0	0.0	0.0
Race/ethnicity									
White (not Hispanic)	8,699	80.7	85.7	8,638	80.4	85.1	8,657	80.2	85.3
Black (not Hispanic)	2,037	73.6	79.3	2,027	72.7	78.9	2,075	74.7	80.8
Hispanic	2,480	75.3	80.6	2,451	74.6	79.6	2,465	75.0	80.1
Asian	963	79.5	84.5	927	76.1	81.3	951	79.1	83.4
Pacific Islander	182	86.1	90.5	182	86.1	90.5	180	86.4	89.6
American Indian or Alaska Native	283	78.4	86.5	283	78.4	86.5	283	79.2	86.5
Other	345	71.9	80.2	338	71.2	78.6	348	73.1	80.9
Unknown race/ethnicity	32	36.7	51.6	32	36.7	51.6	32	36.7	51.6
Year of birth									
1992	4,462	79.4	85.0	4,415	78.8	84.1	4,462	79.7	85.0
1993	10,466	77.7	83.2	10,370	77.1	82.4	10,441	77.4	83.0
Other/unknown	93	61.6	72.1	93	61.6	72.1	88	58.6	68.2

¹ Based on ECLS-K survey data and not on data from the sampling frame.

² Reading, math, or general knowledge assessment was scorable or child was language minority (not Spanish) or disabled.

³ Family structure portion of parent interview was completed.

⁴ A completed questionnaire was defined as one that was not completely left blank.

student level for these other data collection efforts. The facilities checklist has a 79.1 percent completion rate, which is about three points higher than that for the school administrator questionnaire, the only other school-level survey instrument. The student record abstract, which was to have been completed for all students except for those who moved and could not be found in a school, had a 72.2 percent completion rate. There are fewer special education teacher questionnaires overall but their completion rates are higher, 88.7 percent for part A, which captures teacher information, and 82.9 percent for part B, which relates to individual students who receive special education services. Finally there is the adaptive behavior scale (68.5 percent), which was collected only for students who were not assessed in spring-first grade because of physical or mental disability.

Table 5-17.—ECLS-K spring-first grade: number of completed instruments and child-level completion rates for additional data collected—children sampled in the base year

Category	Completes	Completion rates	
		Weighted	Unweighted
Facility checklist*	15,319	79.1	85.3
Student record abstract*	13,928	72.2	77.5
Special education—Part A*	693	88.7	88.5
Special education—Part B*	650	82.9	83.0
Adaptive behavior scale*	22	68.5	68.8

* A completed instrument was defined as one that was not completely left blank.

5.7.2 Students Sampled in First Grade

In spring-first grade the student sample was freshened to include first graders who had no chance of selection in the base year because they did not attend kindergarten in the United States or were in first grade in the fall of 1998. (For a detailed description of the freshening procedure see section 4.3.2.) Nonresponse attrition in the freshened student sample could occur at two stages: during the procedure for sampling schools for freshening and identifying children to be used as freshening links (first component) and then during data collection after the freshened students had been identified (second component). The first component alone can further be decomposed into two sources: attrition due to entire schools refusing to implement the freshening procedure (*the school term*), and attrition due to ECLS-K sample children moving to other schools (*the child term*). To contain costs, students who transferred from schools targeted for freshening were not used as links to identify freshened students, even when they were otherwise followed for data collection. These movers were considered freshening nonrespondents in the *child term*.

Table 5-18 presents weighted completion rates for freshened students. The two components of the completion rates are presented separately in table 5-18. The actual completion rates are the products of the two components. The first component is separated into *a school term* and *a child term* as described earlier. For this component, the completion rate is defined as the freshening completion rates, as opposed to the survey instrument completion rates found in the second component. The weighted freshening completion rate for children in schools targeted for freshening (*the school term*) is 65.4 percent. The reasons that schools did not participate in the freshening process included refusing or being unable to provide the requested information in order to complete the procedures. Within the schools that agreed to freshen, the freshening completion rate is 98.3 percent, the slight loss due to students who transferred to other schools (*the child term*). Multiplying these two terms together gives a first component completion rate of 64.3 percent. The second component varies by survey instrument. The rates for the paper-and-pencil instruments range from 78.2 percent for the student record abstract to 100 percent for the special education teacher—part A questionnaire and are uniformly higher than for the kindergarten sample. The child assessment at 84.0 percent is four points lower than for the kindergarten sample and the parent interview, at 64.9 percent, is nearly 20 points down. These figures demonstrate that, except for the parent interview, data collection went very well once the freshening procedure had been implemented. The final completion rate for each instrument is the product of the two components. Because of the poor showing at the sampling stage these range from a high of 64.3 percent for the special education teacher—part A questionnaire to 41.7 percent for the parent interview.

5.7.3 Spring-First Grade Completion Rates – All Children

Table 5-19 presents final spring-first grade completion rates for children sampled in kindergarten, children sampled in first grade and all children combined. Because children sampled in first grade represent such a small fraction of the total population of children their inclusion brings down the combined rate only one- or two-tenths of a percent relative to the children sampled in kindergarten rates, even though the children sampled in first grade rates are quite low.

Table 5-18.—ECLS-K spring-first grade: number of completed child-level cases and child-level completion rates for children sampled in first grade

Category	Completes	Completion rates ¹	
		Weighted	Unweighted
First component	7,135	64.3	77.9
School term ²	7,192	65.4	78.5
Child term ³	7,135	98.3	99.2
Second component			
Child assessment ⁴	134	84.0	81.2
Parent interview ⁵	104	64.9	63.0
Teacher questionnaire—Part A ⁶	145	89.7	87.9
Teacher questionnaire—Part C ⁶	144	88.9	87.3
Teacher questionnaire—Part B ⁶	132	81.9	80.0
Special education—Part A ⁶	15	100.0	100.0
Special education—Part B ⁶	14	96.5	93.3
School administrator questionnaire ⁶	145	91.8	87.9
Facility check list ⁶	159	96.9	96.4
Student records abstract ⁶	120	78.2	72.7
Completion rates			
Child assessment ⁴	134	54.0	63.3
Parent interview ⁵	104	41.7	49.1
Teacher questionnaire—Part A ⁶	145	57.7	68.4
Teacher questionnaire—Part B ⁶	144	57.2	68.0
Teacher questionnaire—Part C ⁶	132	52.7	62.3
Special education—Part A ⁶	15	64.3	77.9
Special education—Part B ⁶	14	62.1	72.7
School administrator questionnaire ⁶	145	59.0	68.4
Facility check list ⁶	159	62.3	75.1
Student records abstract ⁶	120	50.3	56.6

¹ In the first component, this is the completion rate for freshening. In the second component, this is the completion rate for the survey instruments. The product of the two components is the overall completion rate for the survey instruments.

² The freshening completes and completion rates for children in schools targeted for freshening.

³ The freshening completes and completion rates for children in schools that agreed to the freshening procedure.

⁴ Reading, math, or general knowledge assessment was scorable or child was language minority (not Spanish) or disabled.

⁵ Family structure portion of parent interview was completed.

⁶ A completed questionnaire was defined as one that was not completely left blank.

Table 5-19.—ECLS-K spring-first grade: number of completed child-level cases and child-level completion rates, for children sampled in kindergarten and first grade, by survey instruments

Survey instrument	Children sampled in kindergarten				Children sampled in first grade				All children			
	Completes		Completion rates		Completes		Completion rates		Completes		Completion rates	
	Weighted	Unweighted	Weighted	Unweighted	Weighted	Unweighted	Weighted	Unweighted	Weighted	Unweighted	Weighted	Unweighted
Child assessment ¹	16,593	88.0	91.8	134	54.0	63.3	16,727	87.2	91.6			
Parent interview ²	15,522	84.5	85.8	104	41.7	49.1	15,626	83.5	85.6			
Teacher questionnaire—Part A ³	15,021	78.1	83.6	145	57.7	68.4	15,166	77.6	83.5			
Teacher questionnaire—Part B ³	14,878	77.5	82.8	144	57.2	68.0	15,022	77.0	82.7			
Teacher questionnaire—Part C ³	14,991	78.0	83.5	132	52.7	62.3	15,123	77.4	83.3			
Special education—Part A ³	693	88.7	88.5	15	64.3	77.9	708	88.1	88.4			
Special education—Part B ³	650	82.9	83.0	14	62.1	72.7	664	82.4	82.9			
School administrator questionnaire ³	14,619	76.3	81.4	145	59.0	68.4	14,764	75.9	81.3			
Facility check list ³	15,319	79.1	85.3	159	62.3	75.1	15,478	78.7	85.2			
Student records abstract ³	13,928	72.2	77.5	120	50.3	56.6	14,048	71.7	77.3			
Adaptive behavior scale	22	68.5	68.8	-	-	-	22	68.5	68.8			

¹ Reading, math, or general knowledge assessment was scorable or child was language minority (not Spanish) or disabled.

² Family structure portion of parent interview was completed.

³ A completed questionnaire was defined as one that was not completely left blank.