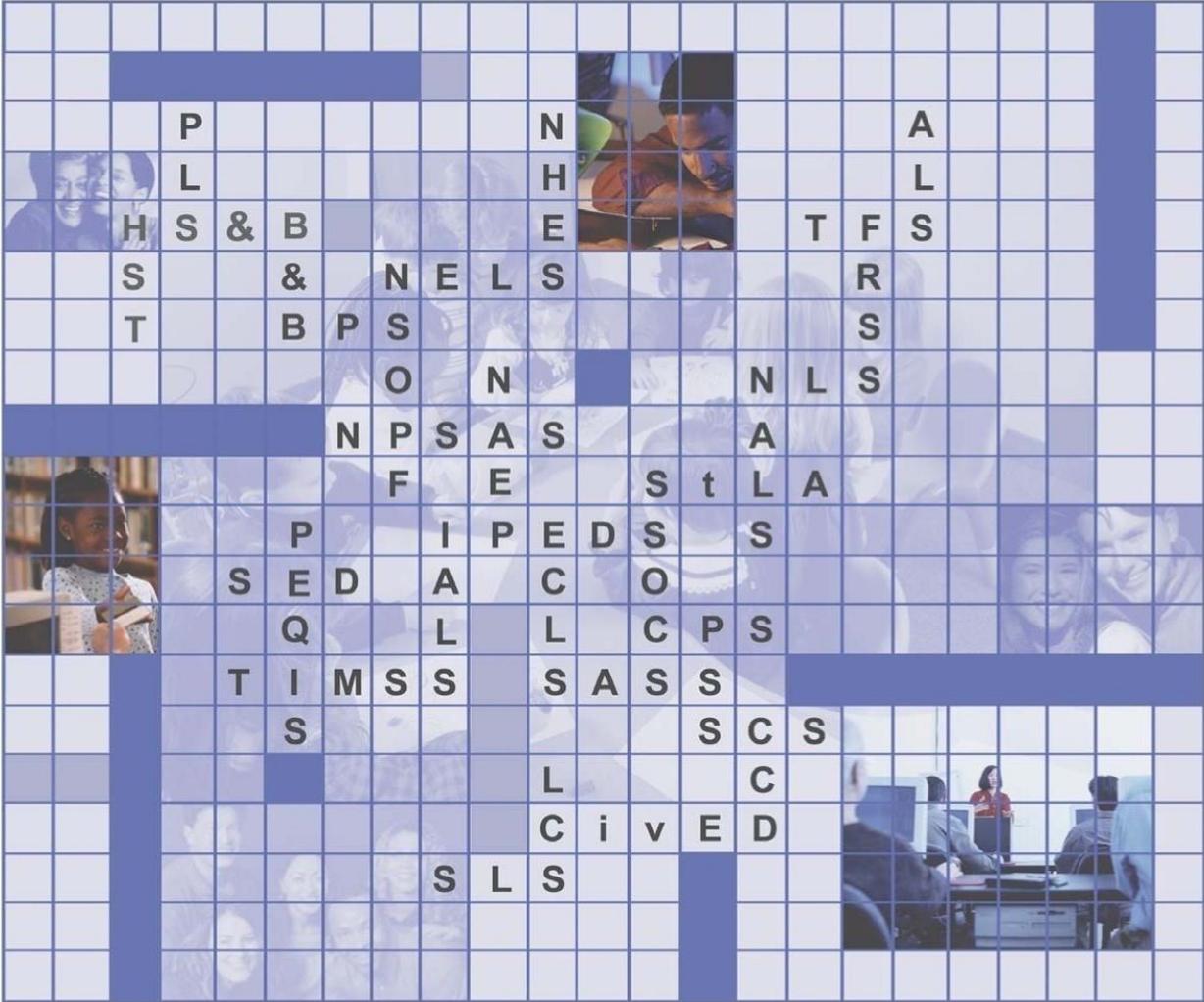


NCES Handbook of Survey Methods



NCES Handbook of Survey Methods

June 2011

Shelley Burns

Editor

National Center for Education Statistics

Xiaolei Wang

Editor

American Institutes for Research

Alexandra Henning

Editor

Quality Information Partners, Inc

NCES 2011-609

U.S. DEPARTMENT OF EDUCATION

 **ies** NATIONAL CENTER FOR
EDUCATION STATISTICS
Institute of Education Sciences

U.S. Department of Education

Arne Duncan
Secretary

Institute of Education Sciences

John Q. Easton
Director

National Center for Education Statistics

Jack Buckley
Commissioner

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

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

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

NCES, IES, U.S. Department of Education
1990 K Street NW
Washington, DC 20006-5651

JUNE 2011

The NCES Home Page address is <http://nces.ed.gov>.

The NCES Publications and Products address is <http://nces.ed.gov/pubsearch>.

This report was prepared in part under Contract No. ED-05-CO-0044 with American Institutes for Research. Mention of trade names, commercial products, or organizations does not imply endorsement by the U.S. Government.

Suggested Citation

Burns, S., Wang, X. and Henning, A. (Eds.) (2011). *NCES Handbook of Survey Methods* (NCES 2011-609). U.S. Department of Education, National Center for Education Statistics. Washington, DC: U.S. Government Printing Office.

For ordering information on this report, write to

ED Pubs, U.S. Department of Education
P.O. Box 22207
Alexandria, VA 22304

or call toll free 1-877-4-ED-Pubs or order online at <http://www.edpubs.gov>.

Content Contact

Marilyn Seastrom
(202) 502-7303
Marilyn.Seastrom@ed.gov

Acknowledgments

This second edition of the *NCES Handbook of Survey Methods* would not have been possible without the efforts of many people. Sincere thanks to all the NCES project officers on the surveys and studies documented in the Handbook—who answered questions, provided unpublished materials, and reviewed the chapters about their surveys.

Contributing Authors

Xiaolei Wang	American Institutes for Research
Alexandra Henning	Quality Information Partners, Inc.
Weiwei Cui	National Institute of Statistical Sciences
Francis Huang	Formerly of American Institutes for Research
Susan Armstrong	American Institutes for Research
Kyo Sin Kang	American Institutes for Research
Jinae Boyer	American Institutes for Research
Simone Robers	American Institutes for Research

Contents

Acknowledgments	iii
Contributing Authors	v
List of Acronyms	xiii
Introduction	1
NCES’s Role and Organization	1
Organization of the Handbook.....	2

EARLY CHILDHOOD EDUCATION SURVEY

Chapter 1: Early Childhood Longitudinal Study (ECLS)	5
1. Overview	5
2. Uses of Data	8
3. Key Concepts	8
4. Survey Design	9
5. Data Quality and Comparability	20
6. Contact Information.....	25
7. Methodology and Evaluation Reports	25

ELEMENTARY AND SECONDARY EDUCATION SURVEYS

Chapter 2: Common Core of Data (CCD)	27
1. Overview	27
2. Uses of Data	29
3. Key Concepts	29
4. Survey Design	30
5. Data Quality and Comparability	32
6. Contact Information.....	33
7. Methodology and Evaluation Reports	34
Chapter 3: Private School Universe Survey (PSS)	35
1. Overview	35
2. Uses of Data	35
3. Key Concepts	36
4. Survey Design	36
5. Data Quality and Comparability	40
6. Contact Information.....	42
7. Methodology and Evaluation Reports	42
Chapter 4: Schools and Staffing Survey (SASS)	45
1. Overview	45
2. Uses of Data	47
3. Key Concepts	48
4. Survey Design	49
5. Data Quality and Comparability	53
6. Contact Information.....	55

7. Methodology and Evaluation Reports	55
Chapter 5: SASS Teacher Follow-up Survey (TFS)	59
1. Overview	59
2. Uses of Data	60
3. Key Concepts	60
4. Survey Design	60
5. Data Quality and Comparability	62
6. Contact Information.....	63
7. Methodology and Evaluation Reports	64
Chapter 6: National Longitudinal Study of the High School Class of 1972 (NLS:72)	65
1. Overview	65
2. Uses of Data	67
3. Key Concepts	67
4. Survey Design	68
5. Data Quality and Comparability	73
6. Contact Information.....	76
7. Methodology and Evaluation Reports	76
Chapter 7: High School and Beyond (HS&B) Longitudinal Study	79
1. Overview	79
2. Uses of Data	81
3. Key Concepts	82
4. Survey Design	82
5. Data Quality and Comparability	87
6. Contact Information.....	90
7. Methodology and Evaluation Reports	90
Chapter 8: National Education Longitudinal Study of 1988 (NELS:88)	93
1. Overview	93
2. Uses of Data	96
3. Key Concepts	97
4. Survey Design	97
5. Data Quality and Comparability	102
6. Contact Information.....	107
7. Methodology and Evaluation Reports	107
Chapter 9: Education Longitudinal Study of 2002 (ELS:2002).....	109
1. Overview	109
2. Uses of Data	113
3. Key Concepts	114
4. Survey Design	114
5. Data Quality and Comparability	120
6. Contact Information.....	125
7. Methodology and Evaluation Reports	125

LIBRARY SURVEYS

Chapter 10: SASS School Library Media Center Survey (SLS)	127
1. Overview	127
2. Uses of Data	128
3. Key Concepts	128
4. Survey Design	128
5. Data Quality and Comparability	130

6. Contact Information.....	130
7. Methodology and Evaluation Reports	130
Chapter 11: Academic Libraries Survey (ALS).....	133
1. Overview	133
2. Uses of Data	134
3. Key Concepts	134
4. Survey Design	135
5. Data Quality and Comparability.....	140
6. Contact Information.....	141
7. Methodology and Evaluation Reports	141
 POSTSECONDARY AND ADULT EDUCATION SURVEYS	
Chapter 12: Integrated Postsecondary Education Data System (IPEDS)	143
1. Overview	143
2. Uses of Data	148
3. Key Concepts	148
4. Survey Design	149
5. Data Quality and Comparability.....	153
6. Contact Information.....	158
7. Methodology and Evaluation Reports	158
Chapter 13: National Study of Postsecondary Faculty (NSOPF)	161
1. Overview	161
2. Uses of Data	162
3. Key Concepts	163
4. Survey Design	164
5. Data Quality and Comparability.....	171
6. Contact Information.....	176
7. Methodology and Evaluation Reports	176
Chapter 14: National Postsecondary Student Aid Study (NPSAS)	177
1. Overview	177
2. Uses of Data	178
3. Key Concepts	179
4. Survey Design	179
5. Data Quality and Comparability.....	186
6. Contact Information.....	189
7. Methodology and Evaluation Reports	190
Chapter 15: Beginning Postsecondary Students (BPS) Longitudinal Study	193
1. Overview	193
2. Uses of Data	194
3. Key Concepts	195
4. Survey Design	195
5. Data Quality and Comparability.....	198
6. Contact Information.....	201
7. Methodology and Evaluation Reports	201
Chapter 16: Baccalaureate and Beyond (B&B) Longitudinal Study	203
1. Overview	203
2. Uses of Data	205
3. Key Concepts	205
4. Survey Design	206

5.	Data Quality and Comparability	212
6.	Contact Information.....	215
7.	Methodology and Evaluation Reports	215
Chapter 17: Survey of Earned Doctorates (SED)		217
1.	Overview	217
2.	Uses of Data	218
3.	Key Concepts	218
4.	Survey Design	219
5.	Data Quality and Comparability	221
6.	Contact Information.....	224
7.	Methodology and Evaluation Reports	224

EDUCATIONAL ASSESSMENT SURVEYS

Chapter 18: National Assessment of Educational Progress (NAEP).....		227
1.	Overview	227
2.	Uses of Data	232
3.	Key Concepts	232
4.	Survey Design	232
5.	Data Quality and Comparability	242
6.	Contact Information.....	248
7.	Methodology and Evaluation Reports	248

Chapter 19: National Adult Literacy Survey (NALS).....		253
1.	Overview	253
2.	Uses of Data	254
3.	Key Concepts	255
4.	Survey Design	255
5.	Data Quality and Comparability	261
6.	Contact Information.....	263
7.	Methodology and Evaluation Reports	264

Chapter 20: National Assessment of Adult Literacy (NAAL)		265
1.	Overview	265
2.	Uses of Data	268
3.	Key Concepts	268
4.	Survey Design	269
5.	Data Quality and Comparability	277
6.	Contact Information.....	279
7.	Methodology and Evaluation Reports	279

Chapter 21: Trends In International Mathematics and Science Study (TIMSS)		281
1.	Overview	281
2.	Uses of Data	284
3.	Key Concepts	284
4.	Survey Design	285
5.	Data Quality and Comparability	290
6.	Contact Information.....	292
7.	Methodology and Evaluation Reports	292

Chapter 22: Program For International Student Assessment (PISA)		295
1.	Overview	295
2.	Uses of Data	296
3.	Key Concepts	296

4.	Survey Design	296
5.	Data Quality and Comparability	303
6.	Contact Information.....	305
7.	Methodology and Evaluation Reports	305
Chapter 23: International Adult Literacy Survey (IALS)		309
1.	Overview	309
2.	Uses of Data	310
3.	Key Concepts	310
4.	Survey Design	311
5.	Data Quality and Comparability.....	316
6.	Contact Information.....	319
7.	Methodology and Evaluation Reports	319
Chapter 24: Adult Literacy and Lifeskills Survey (ALL)		321
1.	Overview	321
2.	Uses of Data	322
3.	Key Concepts	322
4.	Survey Design	323
5.	Data Quality and Comparability.....	327
6.	Contact Information.....	329
7.	Methodology and Evaluation Reports	329
Chapter 25: Progress In International Reading Literacy Study (PIRLS).....		331
1.	Overview	331
2.	Uses of Data	333
3.	Key Concepts	333
4.	Survey Design	334
5.	Data Quality and Comparability.....	338
6.	Contact Information.....	342
7.	Methodology and Evaluation Reports	342
 HOUSEHOLD SURVEYS		
Chapter 26: National Household Education Surveys Program (NHES)		345
1.	Overview	345
2.	Uses of Data	348
3.	Key Concepts	348
4.	Survey Design	348
5.	Data Quality and Comparability.....	359
6.	Contact Information.....	263
7.	Methodology and Evaluation Reports	263
Chapter 27: Current Population Survey (CPS) – October Supplement.....		367
1.	Overview	367
2.	Uses of Data	368
3.	Key Concepts	368
4.	Survey Design	369
5.	Data Quality and Comparability.....	371
6.	Contact Information.....	373
7.	Methodology and Evaluation Reports	373

SMALL SPECIAL-PURPOSE NCES SURVEYS

Chapter 28: Crime and Safety Surveys	375
1. School Crime Supplement (SCS)	375
2. School Survey On Crime and Safety (SSOCS)	380
Chapter 29: High School Transcript (HST) Studies.....	387
1. NAEP High School Transcript Studies.....	388
2. Longitudinal Surveys High School Transcript Studies.....	395
Chapter 30: Quick Response Information System	403
1. Fast Response Survey System (FRSS)	403
2. Postsecondary Education Quick Information System (PEQIS).....	410

APPENDICES

Appendix A: Glossary of Statistical Terms	A-1
Appendix B: Ordering NCES Publications and Data Files.....	B-1
Appendix C: Web-based and Standalone Tools for Use with NCES Survey Data	C-1
Appendix D: NCES Survey Web Site Addresses	D-1
Appendix E: Index.....	E-1

List of Acronyms

2PL	Two-parameter logistic [model]
3PL	Three-parameter logistic [model]
ACA	American Correctional Association
ACER	Australian Council for Educational Research
AE	Adult Education [survey]
AELL	Adult Education and Lifelong Learning Survey
AEWR	Adult Education for Work-Related Reasons [survey]
ACG	Academic Competitiveness Grant
AI/AN	American Indian and Alaska Native
ALA-ORS	American Library Association's Office of Research and Statistics
ALL	Adult Literacy and Lifeskills Survey
ALS	Academic Libraries Survey
ALSA	Adult Literacy Supplement Assessment
AM	analytical software
AP	Advanced Placement
AQS	Attachment Q-Sort
ASCII	American Standard Code for Information Interchange (character-encoding scheme used for representing data as text)
ASPA	Before- and After-School Programs and Activities Survey
B&B	Baccalaureate and Beyond Longitudinal Study
BHR	Balanced half-sample replication [method of variance estimation] (also referred to as BRR)
BHS	Balanced half-sample
BIB	Balanced Incomplete Block Spiraling (a type of matrix sampling)
BIE	Bureau of Indian Education. In 2006, the Office of the Assistant Secretary for the Department of the Interior established the Bureau of Indian Education (BIE). This organization is responsible for prekindergarten, elementary, secondary, and postsecondary schools and educational activities of the Bureau of Indian Affairs (BIA). Prior to 2006, these schools were referred to as BIA schools.
BILOG	computer software used in scaling
BJS	Bureau of Justice Statistics, U.S. Department of Justice
BLS	Bureau of Labor Statistics, U.S. Department of Labor
BPS	Beginning Postsecondary Students Longitudinal Study
BRR	Balanced Repeated Replication [method of variance estimation] (also referred to as BHR)
BSF-R	Bayley Short Form-Research edition [scale]
BSID	Bayley Scales for Infant Development
BYI	Base Year Ineligible [study]
C	Completions [survey component]
CA	Chief Administrator
CACE	Computer-Assisted Coding and Editing
CADE	Computer-Assisted Data Entry

CAPI	Computer-Assisted Personal Interview
CATI	Computer-Assisted Telephone Interview
CCD	Common Core of Data
CCO	Child Care Observation
CCP	Child Care Provider
CHAID	Chi-squared Automatic Interaction Detector
CI	Civic Involvement [survey]
CIP	Classification of Instructional Programs
CN	Consolidated Form [survey component]
CN-F	Consolidated Form, “finance” part [survey component]
COSLA	Chief Officers of State Library Agencies
CSSO	Chief State School Officer
CPS	Current Population Survey
CSSC	Classification of Secondary School Courses
CSV	Comma-separated values (file format)

DAS	Data Analysis System
DIF	Differential Item Functioning
DoDEA	Department of Defense Education Activity [schools]
DOL	U.S. Department of Labor
DPC	Data Processing Center
DRF	Doctorate Records File

E12	12-Month Enrollment [survey component]
EAP	Employees by Assigned Position
ECB	Electronic Codebook
ECE	Early Childhood Education [survey]
ECEP	Early Care and Education Provider
ECLS	Early Childhood Longitudinal Study
ECLS-B	Early Childhood Longitudinal Study, Birth Cohort
ECLS-K	Early Childhood Longitudinal Study, Kindergarten Cohort
ECPP	Early Childhood Program Participation [survey]
EDAT	Education Data Analysis Tool
ED	Enumeration District
ED Pubs	U.S. Department of Education Publications Center
EDEN	Education Data Exchange Network
EDI	Electronic Data Interchange
EEO-6	Higher Education Staff Information [report]
EEOC	U.S. Equal Employment Opportunity Commission
EF	Fall Enrollment
EFC	Expected Family Contribution
ELS: 2002	Education Longitudinal Study of 2002
EP	Fall Enrollment in Occupationally Specific Programs [survey component]
ERIC	Education Resources Information Center
ESL	English as a Second Language

ETS	Educational Testing Service
-----	-----------------------------

F	Finance [survey component]
FAN	Fluency Addition to NAAL
FASB	Financial Accounting Standards Board
FICE	Federal Interagency Committee on Education
FIMS	First International Mathematics Study
FIPS	Federal Information Processing Standards
FISS	First International Science Study
FRSS	Fast Response Survey System
FSES	Followback Study of Excluded Students
FTB	First-time beginning/beginners
FTE	Full-time Equivalency
FY	Fiscal year

GASB	Governmental Accounting Standards Board
GED	General Equivalency Diploma
GEM	Generalized Exponential Model
GPA	Grade Point Average
GPCM	Generalized Partial Credit Model
GPO	Government Printing Office
GR200	Graduation Rates 200
GRL	Graded response logistic [model]
GRS	Graduation Rate Survey
GSS	General Social Survey
GVF	Generalized Variance Function

HBCU	Historically Black Colleges and Universities
HEGIS	Higher Education General Information System
HEOS	Higher Education Opportunity Act
HHL	Household and Library Use [survey]
HLM	Hierarchical Linear Modeling
HR	Human Resources [survey component]
HS&B	High School and Beyond
HSES	High School Effectiveness Study
HSI	Hispanic-Serving Institutions
HST	High School Transcript [studies]

IAEP	International Assessment of Educational Progress
IALS	International Adult Literacy Survey
IB	International Baccalaureate
IC	Institutional Characteristics [survey component]
ICS	Integrated Control System
ICT	Information and communication technology
IDE	International Data Explorer
IDEALS	Input and Data Editing for Academic Library Statistics
IEA	International Association for the Evaluation of Educational Achievement

IEP	Individualized Education Program
IES	Institute of Education Sciences
IHE	Institution of Higher Education
IMS	Integrated Management System
IPEDS	Integrated Postsecondary Education Data System
IRT	Item Response Theory
ISCED	International Standard Classification of Education
ISCO	International Standard Classification of Occupations
ISIC	International Standard Industrial Classification
ISR	Interview status recode

JRR	Jackknife Repeated Replication
-----	--------------------------------

LEA	Local Education Agency
LEP	Limited English Proficiency

MDR	Market Data Retrieval, a Dun & Bradstreet (D&B) Company
MEP	Migrant Education Program
MIL	Missing Information Letter
MOL	Mathematics Online [study]
MOS	Measure of size
MML	Marginal maximum likelihood
MSA	Metropolitan Statistical Area

NAAL	National Assessment of Adult Literacy
NAEP	National Assessment of Educational Progress
NAGB	National Assessment Governing Board
NALS	National Adult Literacy Survey
NCATS	Nursing Child Assessment Teaching Scale
NCEA	National Catholic Educational Association
NCES	National Center for Education Statistics
NCLIS	U.S. National Commission on Libraries and Information Science
NCOA	National Change of Address (data base)
NCR	National Research Council
NCVS	National Crime Victimization Survey
NDE	NAEP Data Explorer
NELS:88	National Education Longitudinal Study of 1988
NHES	National Household Education Survey
NIES	National Indian Education Study
NIH	National Institutes of Health
NLS:72	National Longitudinal Study of the High School Class of 1972
NORC	National Opinion Research Center
NPEFS	National Public Education Financial Survey
NPM	National Project Manager
NPS	New Participant Supplement
NPSAS	National Postsecondary Student Aid Study

NPSQ	New Participant Supplement Questionnaire
NRC	National Research Council (chapters 17 and 20); National Research Coordinator (chapter 25)
NSF	National Science Foundation
NSOPF	National Study of Postsecondary Faculty
NTID	National Technical Institute for the Deaf
OBE	Office of Business Economics
OCR	[U.S. Department of Education's] Office for Civil Rights
OECD	Organization for Economic Cooperation and Development
OLDS	Oral Language Development Scale
OMB	Office of Management and Budget
OPAC	Online Public Access Catalog
OPEID	An 8-digit school identification code developed by OPE
OREALC	Regional Office of Education for Latin America and the Caribbean [UNESCO] [Acronym is based on foreign phrase]
OPE	Office of Postsecondary Education
PC CARP	Analytical software
PEPS	Postsecondary Education Participants System
PEQIS	Postsecondary Education Quick Information System
PETS	Postsecondary Education Transcript Study (chapter 6); Postsecondary Education Telephone System (chapter 12)
PFI	Parent and Family Involvement in Education [survey]
PIAAC	Program for the International Assessment of Adult Competencies
PIRLS	Progress in International Reading Literacy Study
PISA	Program for International Student Assessment
PLS	Public Libraries Survey
PPA	Program Participation Agreement
PPS	Probability proportional to size (a sampling method)
PROC IMPUTE	A regression-based imputation method
PSS	Private School Universe Survey
PSU	Primary Sampling Unit
QED	Quality Education Data, Inc.
QCM	Quality Control Monitor
R	analytical software
RA	Research assistantship
RCS	Receipt Control System
RDD	Random digit dialing (telephone survey technique)
ROTC	Reserve Officers Training Corps
RTD	Registered time-to-degree
RTI	Research Triangle Institute
S	Fall Staff [survey component]
SA	Survey administrator (chapter 9); Salaries [survey component] (chapter 12)
SAA	Survey administrator assistant
SAAL	State Assessment of Adult Literacy

SAQ	Self-administered questionnaire
SAS	analytical software
SASS	Schools and Staffing Survey
SAVD	School Associated Violent Death Study
SCS	School Crime Supplement
SD	Students with Disabilities
SDR	Students with Disabilities
SEA	State Education Agency
SED	Survey of Earned Doctorates
SEOG	Supplemental Educational Opportunity Grants
SES	Socioeconomic status
SFA	Student Financial Aid
SHR	Supplementary Homicide Reports
SIF	School Information Form
SIMS	Second International Mathematics Study
SIPP	Survey of Income and Program Participation
SISS	Second International Science Study
SLS	School Library Survey
SMART	Science and Mathematics Access to Retain Talent [Grant]
S-Plus	analytical software
SPSS	analytical software
SQL	Structured Query Language
SR	School Readiness [survey]
SRIF	Student Record Information Form
SS&D	School Safety and Discipline [survey]
SSOCS	School Survey on Crime and Safety
SSP	Stratification Search [analytical] Program
SST	Secondary School Taxonomy
STATA	analytical software
STEM	Science, Technology, Engineering, and Mathematics
SQL	Structured Query Language
SUDAAN	analytical software
SUREG	a command in computer program STATA

TAS-45	Toddler Attachment Sort-45
TBA	Technology-Based Assessment [project]
TFS	Teacher Follow-up Survey
TIGER	Topologically Integrated Geographical Encoding and Referencing
TIMSS	Trends in International Mathematics and Science Study
TRE	Technology-Rich Environments [study]
TSA	Trial state assessment
TTD	Total time-to-degree
TUDA	Trial Urban District Assessment

UCR	Uniform Crime Reporting
-----	-------------------------

UH	Une Heure (one-hour) [booklet]
UIS	Institute for Statistics
UNESCO	United Nations Educational, Scientific, and Cultural Organ
USDA	U. S. Department of Agriculture
USU	Ultimate Sampling Unit

WesVar	analytical software
WISQARS	Web-based Injury Statistics Query and Reporting System
WOL	Writing Online [study]

YALS	Young Adult Literacy Survey
YCI	Youth Civic Involvement [survey]
YRBSS	Youth Risk Behavior Surveillance System

Introduction

Since its inception, the National Center for Education Statistics (NCES) has been committed to the practice of documenting its statistical methods for its customers and of seeking to avoid misinterpretation of its published data. The reason for this policy is to assure customers that proper statistical standards and techniques have been observed, to guide them in the appropriate use of information from NCES, and to make them aware of the known limitations of NCES data. This second edition of the *NCES Handbook of Survey Methods* continues this commitment by presenting descriptions of how each survey program in NCES obtains and prepares the data it publishes.

NCES statistics are used for many purposes. This handbook aims to provide users of NCES data with the most current information necessary to evaluate the suitability of the statistics for their needs, with a focus on the methodologies for survey design, data collection, and data processing. It is intended to be used as a companion report to *Programs and Plans of the National Center for Education Statistics*, which provides a summary description of the type of data collected by each program at the Center.

NCES's Role and Organization

Among federal agencies collecting and issuing statistics, NCES is the primary federal entity for collecting and analyzing data related to education. The Center's data serve the needs of Congress, other federal agencies, national education associations, academic education researchers, public and private education institutions, tutors, education administration bodies, business, and the general public. NCES is a component of the Institute of Education Sciences (IES) within the U.S. Department of Education.

Within NCES, the Statistical Standards Program, under the direction of the NCES Chief Statistician, provides expertise in statistical standards and methodology, technology, and customer service activities across subject-matter lines. The specific survey programs of NCES, however, have developed around subject-matter areas. As a result, except for the Statistical Standards Program, NCES is organized according to these subject-matter areas, with each survey program falling under one of the following four NCES divisions:

- Assessment
- Early Childhood, International, and Crosscutting Studies
- Elementary/Secondary and Libraries Studies
- Postsecondary Studies

LAYOUT OF HANDBOOK CHAPTERS

- Overview
- Uses of Data
- Key Concepts
- Survey Design
- Data Quality and Comparability
- Contact Information
- Methodology and Evaluation Reports

Organization of the Handbook

The handbook contains 30 chapters. Chapters 1 to 27 each focus on one of the 27 major NCES survey programs. To facilitate locating similar information for the various programs, the information in each of these chapters is presented in a uniform format with the following standard sections and headings:

1. *Overview.* This section includes a description of the purpose of the survey, the type of information collected in the survey, and the periodicity of the survey.
2. *Uses of Data.* This section summarizes the range of issues addressed by the data collected in the survey.
3. *Key Concepts.* This section provides the definitions of a few important concepts specific to the survey.
4. *Survey Design.* This section describes the target population, sample design, data collection and processing procedures, estimation methods, and future plans for the survey. Note that the handbook does not include a list of the data elements collected by each survey. That information can be found in the survey questionnaires, electronic codebooks, data analysis systems, or technical documentation, many available through the NCES website (<http://nces.ed.gov>). However, some general remarks about the data collected can be made here:
 - All race/ethnicity data are collected according to Office of Management and Budget (OMB) standards. For all surveys, data on individuals can be disaggregated by “Black,” “White,” “Hispanic”, and “Other”; for some surveys, data can also be disaggregated by “Asian/Pacific Islander” And “American Native or Alaska Native”.
 - All data on individuals can be disaggregated by sex.
 - All elementary/secondary student-level data collections include information on limited English proficiency and student disability.
 - School-level data collections include information on programs and services offered.

5. *Data Quality and Comparability.* This section describes the appropriate method to use for

estimating sampling error for sample surveys and presents important findings related to different types of nonsampling error (such as coverage error, unit and item nonresponse error, and measurement error). In addition, this section provides summary descriptions of recent design and/or questionnaire changes as well as information on the comparability of similar data collected in other studies.

6. *Contact Information.* This section lists the name of the main contact person for each survey along with a telephone number, e-mail address, and mailing address. Note that at NCES, telephone numbers are assigned according to survey program; staff members leaving one survey program for another have to change telephone numbers.

To find out the current number for a particular staff member, see the NCES Staff Directory (<http://nces.ed.gov/ncestaff>). To find out the current contacts for a particular survey program, please check the program’s website. (NCES survey website addresses are listed in appendix D.)

7. *Methodology and Evaluation Reports.* This section lists the primary recent methodological reports for the survey. Use the NCES number provided to find a particular report through the NCES Electronic Catalog (<http://nces.ed.gov/pubsearch>). Each NCES survey website also contains a list of that survey’s publications.

Note that some of the chapters include cautions to data users. The cautions usually appear in Section 5: Data Quality and Comparability. For example, in chapter 12, section 5, caution is urged when comparing institutions for which data have been imputed for the Integrated Postsecondary Education Data System (IPEDS), since these data are intended for computing national totals and not intended to be an accurate portrayal of an institution’s data.

The first 27 chapters are organized under the following subject-matter rubrics:

- Early Childhood Education Survey
 - Chapter 1: Early Childhood Longitudinal Study (ECLS)
- Elementary and Secondary Education Surveys
 - Chapter 2: Common Core of Data (CCD)
 - Chapter 3: Private School Universe Survey (PSS)

- Chapter 4: Schools and Staffing Survey (SASS)
 - Chapter 5: SASS Teacher Follow-up Survey (TFS)
 - Chapter 6: National Longitudinal Study of the High School Class of 1972 (NLS:72)
 - Chapter 7: High School and Beyond (HS&B) Longitudinal Study
 - Chapter 8: National Education Longitudinal Study of 1988 (NELS:88)
 - Chapter 9: Education Longitudinal Study of 2002 (ELS:2002)
 - Library Surveys
 - Chapter 10: SASS School Library Survey (SLS)
 - Chapter 11: Academic Libraries Survey (ALS)
 - Postsecondary and Adult Education Surveys
 - Chapter 12: Integrated Postsecondary Education Data System (IPEDS)
 - Chapter 13: National Study of Postsecondary Faculty (NSOPF)
 - Chapter 14: National Postsecondary Student Aid Study (NPSAS)
 - Chapter 15: Beginning Postsecondary Students (BPS) Longitudinal Study
 - Chapter 16: Baccalaureate and Beyond (B&B) Longitudinal Study
 - Chapter 17: Survey of Earned Doctorates (SED)
 - Educational Assessment Surveys
 - Chapter 18: National Assessment of Educational Progress (NAEP)
 - Chapter 19: National Adult Literacy Survey (NALS)
 - Chapter 20: National Assessment of Adult Literacy (NAAL)
 - Chapter 21: Trends in International Mathematics and Science Study (TIMSS)
 - Chapter 22: Program for International Student Assessment (PISA)
 - Chapter 23: International Adult Literacy Survey (IALS)
 - Chapter 24: Adult Literacy and Lifeskills (ALL)
 - Chapter 25: Progress in International Reading Literacy Study (PIRLS)
 - Household Surveys
 - Chapter 26: National Household Education Surveys (NHES) Program
 - Chapter 27: Current Population Survey (CPS)—October Supplement
- Chapters 28 through 30 cover multiple surveys or survey systems. The format is similar to that for chapters 1 to 27, but is somewhat abbreviated to allow adequate coverage of multiple surveys within each chapter.
- Small Special-Purpose NCES Surveys
 - Chapter 28: Crime and Safety Surveys: School Crime Supplement (SCS) and School Survey on Crime and Safety (SSOCS)
 - Chapter 29: High School Transcript (HST) Studies
 - Chapter 30: Quick Response Information System
- Details of three surveys are not available at the time of publication, and thus not included in this version of Handbook. The High School Longitudinal Survey (HSLs:09) is a nationally representative, longitudinal study of more than 21,000 ninth graders in 940 schools who will be followed through their secondary and postsecondary years. The study focuses on understanding students' trajectories from the beginning of high school into postsecondary education, the workforce, and beyond. What students decide to pursue when, why, and how are crucial questions for HSLs:09, especially, but not solely, in regards to science, technology, engineering, and math (STEM) courses, majors, and careers. This study includes a student assessment in algebraic skills, reasoning, and problem solving, and surveys of students, their parents, math and science teachers, school administrators, as well as school counselors. The first wave of data collection for HSLs:09 began

in the fall of 2009. The next data collection will occur in the spring of 2012.

The Beginning Teacher Longitudinal Study (BTLS) follows a cohort of beginning public school teachers, who were initially interviewed as part of the 2007–08 Schools and Staffing Survey, over a decade as they continue in pre-K-12 teaching or change careers. In the 2007-08 school year, approximately 2,000 beginning public school teachers responded to a variety of questions about themselves, their schools, their preparation, struggles and future plans. The second year of data collection was 2008–09 and was included in the Teacher Follow-up Survey (TFS). Of the two questionnaires for teachers who began teaching in 2007, one was for teachers who left teaching since the previous SASS and the other for those who were teaching either in the same school as the previous year or in a different school. The topics for the Current Teacher questionnaire included teaching status and assignments, ratings of various aspects of teaching, reasons for moving to a new school, information on having had a mentor teacher in the previous year, and earnings. The topics for the Former Teacher questionnaire included employment status, ratings of various aspects of teaching and their current jobs, information on decisions to leave teaching, whether they had applied for a teaching position, and information on having had a mentor teacher in the previous year. The third year of data collection covered the 2009–10 school year. Current teachers were asked questions regarding teaching status and assignments, their opinions of various aspects of teaching, reasons for moving to a new school, reasons for returning to teaching (if they left after the 07–08 school year but returned for the 2009–10 school year), earnings, and information on having and serving as a mentor. Former teachers were surveyed on current employment status, their opinions on various aspects of teaching and their current jobs, information on decisions to leave teaching (if they left

after the 08–09 school year), and whether they had applied for a new teaching position.

The Program for the International Assessment of Adult Competencies (PIAAC) is a cyclical, large-scale, direct household assessment under the auspices of the Organization for Economic Cooperation and Development (OECD). The assessment will be first administered in 2011 to approximately 5,000 individuals between the ages of 16 and 65 in each of the 27 participating countries. The goal of PIAAC is to assess and compare the basic skills and competencies of adults around the world. The assessment focuses on cognitive and workplace skills needed for successful participation in 21st-century society and the global economy. Specifically, PIAAC measures relationships between individuals' educational background, workplace experiences and skills, occupational attainment, use of information and communications technology, and cognitive skills in the areas of literacy, numeracy, and problem solving.

To avoid repetition within the handbook, some of the statistical terms and procedures that are referred to in multiple chapters of the handbook are defined in **Appendix A. Glossary of Statistical Terms.**

Appendix B describes the various ways in which NCES publications and data files may be obtained. It also provides the reader with information on how to obtain a license for restricted-use data files.

Appendix C provides a list of the web-based and standalone tools for use with each of the NCES surveys.

Appendix D contains a list of the website addresses for each of the NCES surveys.

Appendix E contains an index.

Chapter 1: Early Childhood Longitudinal Study (ECLS)

1. OVERVIEW

The Early Childhood Longitudinal Study (ECLS) program is one of the active longitudinal surveys sponsored by NCES. The ECLS program includes three cohorts: a birth cohort and two kindergarten cohorts (the kindergarten class of 1998–99 and the kindergarten class of 2010–11). The birth cohort study (ECLS-B) followed a sample of children born in 2001 from birth through kindergarten; the first kindergarten study (ECLS-K) followed a sample of children who were in kindergarten in the 1998–99 school year through the eighth grade; and the second kindergarten study (ECLS-K:2011) will follow a sample of kindergartners in the 2010–11 school year through the fifth grade. The ECLS provides a comprehensive and reliable dataset with information about the ways in which children are prepared for school and how children develop in relation to their family, early childhood and school environments.

Purpose

The ECLS provides national data on (1) children’s status at birth and at various points thereafter; (2) children’s transitions to nonparental care, early education programs, and school; and (3) children’s experiences and growth through the eighth grade. These data enable researchers to test hypotheses about associations and interactions of a wide range of family, school, community, and individual variables on children’s development, early learning, and performance in school.

Components

The ECLS has three cohort studies—two *kindergarten* cohort studies (ECLS-K and ECLS-K:2011) and the *birth* cohort study (ECLS-B)—and each of these has its own components.

The Early Childhood Longitudinal Study, Kindergarten Class of 2010–11 (ECLS-K:2011). The ECLS-K:2011 will collect data from children, their families, classroom teachers, special education teachers, school administrators, and care providers on children’s cognitive, social, emotional, and physical development. Information also will be collected on children’s home environment, home educational activities, school environment, classroom environment, classroom curriculum, teacher background, and before- and after-school care.

The Early Childhood Longitudinal Study, Kindergarten Class of 1998–99 (ECLS-K). The ECLS-K collected data from children, their families, classroom teachers, special education teachers, school administrators, and student records. The various components are described below.

Direct child assessments. The direct child assessments covered several cognitive domains (reading and mathematics in kindergarten through eighth grade; general knowledge, consisting of science and social studies questions, in kindergarten and first grade; and science in third, fifth, and eighth grades); a psychomotor assessment (fall kindergarten only), including fine and gross motor skills; and height and weight measurements. Beginning with the third-grade data collection, children

EARLY CHILDHOOD LONGITUDINAL SAMPLE SURVEY: BIRTH COHORT AND KINDERGARTEN COHORT

ECLS collects data from:

- Children
- Parents/guardians
- Child care providers and preschool teachers
- Teachers
- School administrators

reported on their own perceptions of their abilities and achievement, as well as their interest in and enjoyment of reading, math, and other school subjects. An English language proficiency screener, the Oral Language Development Scale (OLDS), was administered to children if school records indicated that the child's home language was not English. The child had to demonstrate a certain level of English proficiency on the OLDS to be administered the ECLS-K cognitive assessment in English. If a child spoke Spanish at home and did not have the English skills required for the ECLS-K battery, the child was administered a Spanish version of the OLDS, and the mathematics and psychomotor assessments were administered in Spanish. . The assessment for each cognitive domain included a routing test (to determine a child's approximate skill level) and second-stage tests that were tailored to different skill levels. In the eighth-grade data collection, children completed a student questionnaire after completing the routing test. The student questionnaire covered many topics about the child's school experiences, school-sponsored and out-of-school activities, self-perceptions of social and academic competence and interests, weight and exercise, and diet.

Parent interviews. Parents/guardians were asked to provide key information about their children and their families, such as the demographics of household members (e.g., age, relation to child, race/ethnicity), family structure (household members and composition), parent/guardian involvement at the school and with children's schoolwork, home educational activities, children's child care experiences, child health, parental/guardian education and employment status, and their children's social skills and behaviors.

Classroom teacher questionnaire. In the kindergarten collections, all kindergarten teachers with ECLS-K-sampled children were asked to provide information on their educational backgrounds, teaching practices, teaching experiences, and the classroom settings in which they taught. They also were asked to complete a child-specific questionnaire that collected information on each sample child's social skills and approaches to learning, academic skills, and education placements. This procedure continued in later waves of the study. However, modifications were made beginning with the spring-fifth grade data collection, where the teachers who were most knowledgeable about the child's performance in each of the core academic subjects (i.e., reading/language arts, mathematics, and science) provided the data pertinent to each child's classroom environment and instruction for the academic subject about which they were most knowledgeable. Teachers

also provided information about their professional background.

Special education teacher questionnaire. In each spring data collection, the primary special education teachers of and special education staff (e.g., speech pathologists, reading instructors, audiologists) who worked with sample children receiving special education services in school were asked to complete questionnaires about the children's experiences in special education, as well as their own professional background. Items in the special education teacher questionnaires addressed topics such as the child's disability, Individualized Education Program (IEP) goals, the amount and type of services sampled children received, and communication with parents and general education teachers about the child's special education program and progress.

School administrator questionnaire. School administrators were asked about school characteristics (e.g., school type, enrollment, and student body composition), school facilities and resources, community characteristics and school safety, school policies and practices, school-family-community connections, school programs programs for particular populations (e.g., limited English proficient students), staffing and teacher characteristics, school governance and climate, and their own characteristics.

Student records abstract. . In each round of data collection except eighth grade, school staff members were asked to complete a student records abstract form for each sampled child after the school year closed. These forms were used to obtain information about the child's attendance record, the presence of an IEP, the type of language or English proficiency screening that the school used, and (in the kindergarten year collection) whether the child participated in Head Start prior to kindergarten. A copy of each child's report card was also requested.

School facilities checklist. This checklist was used to collect information about the (1) availability and condition of the selected school's facilities, such as classrooms, gymnasiums, and toilets; (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 was added to the spring first-grade data collection.

The Early Childhood Longitudinal Study, Birth Cohort (ECLS-B). The ECLS-B, which began in October 2001, was designed to study children's early learning and development from birth through the fall of

the kindergarten year. Over the course of the study, data were collected from multiple sources, including birth certificates, children, parents, nonparental care providers, teachers, and school administrators. These components are described below.

Birth certificates. These records provided information on the date of birth, child's sex, parents' education, parents' race and ethnicity (including Hispanic origin), mother's marital status, mother's pregnancy history, prenatal care, medical and other risk factors during this pregnancy and complications during labor and birth, and child's health characteristics at birth (such as congenital anomalies and abnormal conditions of the baby and the baby's Apgar score).

Parent/guardian interviews. A parent/guardian interview was conducted in the children's home at each data collection point to capture information about the children's early health and development, their experiences with family members and other significant people in their lives, the parents/guardians as caregivers, the home environment, and the neighborhood in which they lived. In most cases, the parent/guardian interviewed was the child's mother or female guardian.

Child assessments. Beginning at 9 months, children participated in activities designed to measure important developmental skills in the cognitive, socioemotional, and physical domains.

Cognitive domain. The cognitive assessments at the 9-month and 2-year data collections assessed general mental ability, including problem solving and language acquisition. The Bayley Short Form-Research Edition (BSF-R), designed specifically for the ECLS-B, was utilized in the 9-month and 2-year data collections and consists of selected items from the Bayley Scales for Infant Development (BSID-II).

The cognitive assessments at the preschool, kindergarten 2006, and kindergarten 2007 data collections assessed early reading and mathematics and consisted of items from the ECLS-K as well as other studies and instruments. Color knowledge also was assessed in the preschool data collection.

Socioemotional domain. The Nursing Child Assessment Teaching Scale (NCATS) was used in the 9-month collection to assess child-parent interactions. An attachment rating, the Toddler Attachment Sort-45 (TAS-45), was used in the second wave of data collection. A videotaped parent-child interaction (Two Bags Task) was also used in the second and third waves of data collection.

Physical domain. In the 9-month data collection, children's height, weight, and middle upper arm circumference were assessed; additionally, a measure of head circumference was taken for children born with very low birth weight. These physical measures were taken again at all follow-up data collections. Additionally, children's fine motor skills and gross motor skills were assessed at all data collections (using the BSF-R motor scale in the 9-month and 2-year data collections and the ECLS-K Bruininks-Oseretsky Test of Motor Proficiency and Movement Assessment Battery for Children in the preschool, kindergarten 2006, and kindergarten 2007 data collections).

Nonparental care and education providers. Individuals and organizations that provided regular care for a child were interviewed with the permission of the child's parents. They were asked about their backgrounds, teaching practices and experience, the children in their care, and children's learning environments. This information was collected from the 2-year data collection on. In the kindergarten 2006 and 2007 collections, a wrap-around care provider interview was used for those children who were in kindergarten and had a before- or after-school care arrangement.

Teacher questionnaires and school data. Once the children entered kindergarten, teachers provided information on their classrooms and on children's cognitive and social development. Information for the school each child attended was obtained from NCES's school universe data files—the Common Core of Data (CCD) for public schools and the Private School Universe Survey (PSS) for private schools.

Father questionnaires. Fathers (both resident and nonresident fathers) completed a self-administered questionnaire, which asked questions about the particular role fathers play in their children's lives; the questionnaire provided information about children's well-being, the activities fathers engage in with their children, and key information about fathers as caregivers. Both resident and nonresident father questionnaires were included in the collections when the children were 9 months old and 2 years old. The resident father questionnaire was included in the preschool collection. No father questionnaires were included in the kindergarten collections.

Periodicity

The ECLS-K collected data in the fall and spring of kindergarten (1998–99), the fall of first grade (1999) (data were collected from a 30 percent subsample in this round), and in the spring of first grade (2000), third grade (2002), fifth grade (2004), and eighth grade (2007).

As currently planned, the ECLS-K:2011 will collect data in the fall and the spring of kindergarten (2010–11), the fall and the spring of first grade (2011–12), and the springs of second grade (2013), third grade (2014), fourth grade (2015), and fifth grade (2016).

The ECLS-B collected data when the children were about 9 months old (2001–02), about 2 years old (2003), about 4 years old (the preschool collection) (2005), and in the fall of kindergarten (2006 and 2007). Note that because of age requirements for school entry, children sampled in the ECLS-B entered kindergarten in two different years. All study children were included in the kindergarten 2006 collection, regardless of their enrollment status or grade in school. The kindergarten 2007 collection included just a portion of the total ECLS-B sample: children who were not yet in kindergarten in the 2006 collection, children who were in kindergarten in the 2006 collection and were repeating kindergarten in the 2007 collection, and twins of children in these groups. The ECLS-B study ended with the kindergarten 2007 wave of collection.

2. USES OF DATA

The ECLS-K provides information critical to informing policies that can respond sensitively and creatively to diverse learning environments. In addition, the ECLS-K enables researchers to study how a wide range of family, school, community, and individual variables are associated with early success in school and later development. The longitudinal nature of the study enables researchers to study children's reading achievement, growth in mathematics, and knowledge of the physical and social worlds in which they live. It also permits researchers to relate trajectories of growth and change to variations in children's school experiences in kindergarten and the early grades.

Like the kindergarten cohort study, the ECLS-B has two goals, descriptive and analytic. The study provides descriptive data on children's health status at birth; children's experiences in the home, nonparental care, and school; and children's development and growth through first grade. The data collected in the ECLS-B can be used to explore the relationships between children's developmental outcomes and their family, health care, nonparental care, school, and community.

The longitudinal nature of the study enables researchers to study children's physical, social, and emotional growth and to relate trajectories of growth and change to variations in children's experience.

3. KEY CONCEPTS

Number right scores. These scores are the counts of raw number of items a child answered correctly. These scores are useful for descriptive purposes only for assessments that are the same for all children. They are not comparable across grades. In the ECLS-K, some assessment items were not included as part of the set of proficiency scores (see details below) because they did not follow a hierarchical pattern. For these items, several item cluster scores were reported for the reading (kindergarten through fifth grade) and science assessments (third and fifth grades). These are simple counts of the number right on small subsets of items linked to particular skills. Because they are based on very few items, their reliability is relatively low.

Item Response Theory (IRT) scale scores. The ECLS direct cognitive assessments employ a two-stage design. As such, within any given domain, children receive a routing set of items (stage 1) and then based on their performance proceed to a certain difficulty level (stage 2). Because not all children receive all items, the assessment scores in the ECLS studies are modeled using Item Response Theory (IRT). Based on children's performance on the items they received, an ability estimate (θ) is derived for each domain. The θ is used to derive other scores, such as scale scores, T-scores, and proficiency probability scores. The IRT scale scores represent estimates of the number of items children would have answered correctly if they had received all of the scored questions in a given content domain. They are reported in both the ECLS-K and ECLS-B. They are useful in identifying cross-sectional differences among subgroups in overall achievement levels and provide a summary measure of achievement useful for correlations analysis with status variables. The IRT scale scores are also used as longitudinal measures of overall growth. Gain scores may be obtained by subtracting children's scale scores at two points in time.

Standardized scores (T-scores). These scores are also IRT based. They 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 that of other groups. A change in mean T-scores over time reflects a change in the group's status with respect to that of other groups.

Proficiency probability scores. These scores are IRT-based and provide information on proficiency in clusters of items of similar difficulty along the overall

scale. The scores measure the probability of mastery of each level and can take on any value between 0 and 1. Because each proficiency probability score targets a particular set of skills, they are ideal for studying the details of achievement. They are useful as longitudinal measures of change because they show not only the extent of gains, but also where on the achievement (or development) scale the gains are taking place.

Race/ethnicity. In the ECLS, new Office of Management and Budget guidelines were followed under which a respondent could select one or more of five dichotomous race categories. In addition, a sixth dichotomous variable was created for those who simply indicated that they were multiracial without specifying the race. Each respondent additionally had to identify whether the child was Hispanic. Using the six dichotomous race variables and the Hispanic ethnicity variable, a race/ethnicity composite variable was created. The categories were White, non-Hispanic; Black or African-American, non-Hispanic; Hispanic, race specified; Hispanic, no race specified; Asian; Native Hawaiian or other Pacific Islander; American Indian or Alaska Native; and more than one race specified, non-Hispanic.

Socioeconomic status (SES). The SES variable reflects the SES of the household at the time of data collection. The components used to create the SES variable were father/male guardian's education, mother/female guardian's education, father/male guardian's occupation, mother/female guardian's occupation, and household income. In the ECLS-K, each parent's occupation was scored using the average of the 1989 General Social Survey (GSS) prestige scores for the 1980 census occupational category codes that correspond to the ECLS-K occupation code. In the ECLS-B, each parent's occupation was scored using the average of the 1989 GSS prestige scores for the 2000 census occupational category codes covered by the ECLS-B occupation.

4. SURVEY DESIGN

Target Population

Representative samples of kindergartners and babies are studied longitudinally for 6 or more years. Kindergarten children enrolled during the 1998–99 school year are the baseline for the ECLS-K cohort; babies born during 2001 are the baseline for the ECLS-B cohort.¹ Kindergarten children enrolled in the

2010–11 school year are the baseline for the ECLS-K:2011 cohort.

Sample Design

The sample design is discussed separately for the kindergarten and birth cohorts.

Kindergarten cohort (ECLS-K). The ECLS-K followed a nationally representative cohort of children from kindergarten through eighth grade.

Base-year (i.e., kindergarten) survey. A nationally representative sample children enrolled in kindergarten programs during the 1998–99 school year was sampled for participation in the study. These children were selected from both public and private schools, offering both full-day and part-day kindergarten programs. The sample was designed to support separate estimates of public and private school kindergartners; Black, Hispanic, White, and Asian/Pacific Islander children; and children grouped by SES.

The sample design for the ECLS-K was a dual-frame, multi-stage sample. First, 100 primary sampling units (PSUs) were selected from an initial frame of 1,400 PSUs, representing counties or groups of contiguous counties. The 24 PSUs with the largest measures of size (where the measure of size is the number of 5-year-olds, taking into account a factor for oversampling Asian/Pacific Islander 5-year-olds) were designated as certainty selections and were set aside. The remaining PSUs were partitioned into 38 strata of roughly equal measures of size. The frame of noncertainty PSUs was first sorted into eight superstrata by metropolitan statistical area (MSA) status and by census region resulting in four MSA superstrata and four non-MSA superstrata. Within the four MSA superstrata, the variables used for further stratification were race/ethnicity (high concentration of Asian/Pacific Islander, Black, or Hispanic), size of class, and 1988 per capita income. Within the four non-MSA superstrata, the stratification variables were race/ethnicity and per capita income. Two PSUs were selected from each noncertainty stratum using Durbin's method. This method selects two first-stage units per stratum without replacement, with probability proportional to size and a known probability of inclusion. The Durbin method was used because it allows variances to be estimated as if the units were selected with replacement.

School selection occurred within the sampled PSUs. Public schools were sampled from a public school

¹ The ECLS-B target population excludes children who were born to mothers younger than age 15 and children who died or were adopted

prior to the 9-month home visit. Over time, the target population excludes children who died or moved abroad permanently.

frame (the 1995–96 CCD), and private schools were sampled from a private school frame (the 1995–96 PSS). The school frame was freshened in spring 1998 to include newly opened schools that were not included in the CCD and PSS (as well as schools that were included in the CCD and PSS but that did not offer kindergarten, according to these sources). A school sample supplement was selected from the freshened frame. In fall 1998, approximately 23 kindergarten children were selected, on average, from each of the sampled schools. Asian/Pacific Islander children and private schools were oversampled.

For the base year of the ECLS-K, 22,670 children were eligible (17,780 in public schools and 2,890 in private schools).

Fall first grade. The fall first grade collection was designed to enable researchers to measure the extent of summer learning loss and the factors associated with such loss and to better disentangle the relationships of school and home characteristics with children's learning. Data collection was limited to 26.7 percent of the base-year children in 30 percent of the originally sampled ECLS-K schools; that is, a total of 5,650 (4,450 public school and 1,200 private school) children. Data collection was attempted for every eligible child (i.e., a base-year respondent) still attending the school in which he or she had been sampled during kindergarten. To contain the cost of collecting data for a child who transferred from the school in which he or she was originally sampled, a random 50 percent of movers (i.e., children who changed schools) were flagged to be followed for the fall first-grade data collection.

Spring first grade. This 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. For the spring first grade, 18,080 children were eligible (14,250 public school and 3,840 private school children). Student freshening brought 170 first-graders into the ECLS-K sample.

Spring third grade. The sample of children for the spring third-grade data collection consisted of all children who were base-year respondents and children who were brought into the sample in the spring of first grade through sample freshening. Sample freshening

was not implemented in third grade. While all students still enrolled in their base-year schools were recontacted, slightly more than 50 percent of the base-year sampled students who had transferred from their kindergarten school were followed for data collection. This subsample of students was the same 50 percent subsample of base-year movers flagged for following in the spring of first grade, with the addition of movers whose home language was not English (followed at 100 percent). For the spring third grade, 16,670 children were eligible² (13,170 public schools and 3,500 private school children).

Spring fifth grade. In fifth grade, four groups of children were not followed, irrespective of other subsampling procedures that were implemented. These were (1) children who became ineligible in an earlier round (because they had died or moved out of the country), (2) children who were subsampled out in previous rounds because they had moved out of their original schools and were not followed, (3) children whose parents emphatically refused to cooperate in any of the data collection rounds since the spring of kindergarten, and (4) children eligible for the third-grade data collection for whom there were neither first-grade nor third-grade data.

Of the remaining children, those who moved from their original schools during fifth grade or earlier were subsampled for follow-up. In order to contain the cost of data collection, the rate of subsampling was lower in fifth grade than it had been in previous years. The subsampling rates maximize the amount of longitudinal data available for key analytic groups. Children whose home language was not English (English language learners or ELLs) continued to be sampled at higher rates (between 15 and 50 percent for base-year ELL respondents, and between 15 and 75 percent for ELL children freshened in first grade).

For the spring fifth grade, 12,030 children were eligible³ (9,570 in public schools and 2,460 in private schools).

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

² This number reflects the longitudinal sample and excludes the 170 first grade freshened cases.

³ This number reflects the longitudinal sample and excludes the 170 first grade freshened cases.

child-level questionnaires filled out by their science teachers.

Spring eighth grade. In the eighth-grade sample, the ineligible children were those who had moved out of the country, were deceased, or had moved to another school and were not subsampled for follow-up in an earlier grade. In the eighth-grade data collection, there was no subsampling of movers for follow-up as in previous rounds, since the majority of children did not remain in the same school from fifth grade to eighth grade (having moved out of elementary school into middle school).

For the spring eighth grade, 11,930 children were eligible⁴ (9,480 in public schools and 2,450 in private schools).

Birth cohort (ECLS-B). The ECLS-B followed a nationally representative sample of children born in 2001 from the time the children were 9 months old through their kindergarten year.

Base-year (i.e., 9-month) survey. The ECLS-B sampled approximately 14,000 babies born in 2001, yielding approximately 10,700 completed cases in the 9-month collection. The sample included children from different racial/ethnic and socioeconomic backgrounds. Chinese children, other Asian/Pacific Islander children, children born with moderately low birth-weight (1,500–2,500 grams), children born with very low birth-weight (under 1,500 grams), and twins were oversampled. There was also a special supplemental component to oversample American Indian children.

The ECLS-B sample design consisted of a two-stage sample of PSUs and children born in the year 2001 within sampled PSUs. The PSUs were MSAs, counties, or groups of counties. Among the 96 sampled PSUs, 24 were large enough to be selected with certainty. The remaining PSUs were selected from groups of PSUs that were stratified by census region; MSA status; minority status (high/low); median income (high/low); and a composite measure of size, which was the expected number of births in 2001 in the PSU. Two PSUs were selected per stratum with probability proportional to size, a function of the expected number of births occurring within the PSU in 2001.

Births were sampled by place of occurrence, rather than by place of current residence. As a result, a different PSU sample than the PSU sample used in the ECLS-K, which uses residence-based population data,

had to be selected. Within the sampled PSUs, children born in the year 2001 were selected by systematic sampling from birth certificates using the National Center for Health Statistics vital statistics record system. The sample was selected on a flow basis, beginning with January 2001 births (who were first assessed 9 months later, in October 2001). Approximately equal numbers of infants were sampled in each month of 2001. Different sampling rates were used for births in different subgroups, as defined by race/ethnicity, birth weight, and plurality (that is, whether or not the sampled newborn was a twin).

The sample of American Indian/Alaska Native (AI/AN) newborns drew from an additional 18 PSUs selected from a supplemental frame consisting of areas where the population has a higher proportion of AI/AN births. These PSUs were located in the western region of the United States. Six of the PSUs were selected with certainty. The noncertainty PSUs were selected independently of the core sample PSUs, with probability proportional to the number of AI/AN births.

Due to state-imposed operational restrictions and passive and active consent procedures, certain sampled PSUs had low expected response rates. For states where expected response rates were only slightly lower than planned, a larger sample was selected in order to achieve adequate numbers of respondents. Substitutions were made for PSUs in states where very low response rates were expected. The original PSU was matched with potential substitute PSUs on the criteria of median income; percentage of newborns in poverty; percentage of newborn Black, Hispanic, and other race/ethnicity children; population density; and birth rate. (AI/AN PSUs also were matched on tribal similarity. A Mahalanobis distance measure of similarity was used to create initial rankings.) Sampling rates from the original PSU were applied within the substitute PSU to obtain the original expected yield. A total of seven PSUs were used as substitutes for the original ECLS-B PSUs. Also, in two instances, an alternative frame was used to draw a sample of births occurring within PSUs with enrollment restrictions. Specifically, birth records were selected directly from hospital lists of births in counties that defined these original PSUs.

For the 9-month collection, approximately 14,200 children were eligible,⁵ and approximately 10,700 participated.

⁴ This number reflects the longitudinal sample and excludes the 170 first grade freshened cases.

⁵ Because the ECLS-B data are restricted-use only, the numbers provided in this section for the ECLS-B are all rounded to the nearest 50.

Two-year collection. Only cases with a completed 9-month parent interview (about 10,700) were eligible for inclusion in the 2-year data collection. However, from that 10,700, about 100 cases where the child had died or moved abroad permanently between the 9-month and 2-year rounds were considered ineligible. There was no further sampling of cases. For the 2-year round of the ECLS-B approximately 9,850 cases participated (i.e. had a completed parent survey).

Preschool collection. All 9,850 cases with a complete 2-year parent interview and an additional 50 AI/AN cases were fielded and considered eligible for the preschool data collection, with the exception of approximately 100 cases in which children had died or moved permanently abroad between the 2-year interview and the preschool wave. For the preschool round of the ECLS-B approximately 8,950 cases participated (i.e., had a completed parent survey).

Kindergarten 2006 collection. For budgetary reasons, the kindergarten 2006 data collection followed a reduced sample (approximately 85 percent) of children who were eligible for the wave. The subsample was allocated disproportionately to the race/ethnicity, birth weight, and plurality domains to maintain larger sample sizes for the smaller domains. AI/AN children and Chinese children who were eligible were included with certainty in the kindergarten 2006 subsample. Eligible children were those with a parent response at all of the prior waves (9 months, 2 years, and preschool) and children sampled in the AI/AN domain with a parent response to the 9-month wave and at least one of the 2-year or preschool waves. AI/AN children who did not respond to either the 2-year or preschool waves were not included in the kindergarten 2006 wave. In addition, children who were identified as ineligible because they had died or moved out of the United States were not included in the kindergarten 2006 data collection.

After subsampling, approximately 7,700 children were eligible for the kindergarten 2006 wave and 7,000 participated (i.e., had a completed parent survey).

Kindergarten 2007 collection. The kindergarten 2007 data collection included a subset of the ECLS-B sample children with a completed parent interview at kindergarten 2006 and who met one of the following conditions: the child had not started kindergarten at the time of the kindergarten 2006 data collection; the child was the twin of a child who had not started kindergarten at the time of the kindergarten 2006 data collection; the child was in kindergarten during the kindergarten 2006 data collection and repeating kindergarten in school year 2007–08; or the child was

the twin of a child who was repeating kindergarten in school year 2007–08.

Of the 7,000 cases from the kindergarten 2006 collection, based on the aforementioned criteria, 2,050 were eligible for the kindergarten 2007 (1,770 as first time entering school and 280 as likely repeating kindergarten). For the kindergarten 2007 wave, approximately 1,900 participated (i.e., had a completed parent survey).

Assessment Design

The design of the ECLS assessments is discussed separately for the kindergarten and birth cohorts.

Kindergarten cohort (ECLS-K). The design of the ECLS-K assessment was guided by the domain assessment framework proposed by the National Education Goals Panel's Resource Group on School Readiness. A critical component of the ECLS-K is the assessment of children along a number of dimensions, such as physical development, social and emotional development, and cognitive development. These domains were chosen because of their importance to success in school. The ECLS-K monitored the status and growth of its children along these domains:

- *Physical and psychomotor development:* Children's height and weight were measured at each data collection point in the ECLS-K. The psychomotor component was included only in the fall kindergarten collection. In that collection, kindergartners were asked to demonstrate their fine and gross motor skills through activities such as building a structure using blocks, copying shapes, drawing figures, balancing, hopping, skipping, and walking backward. Parents and teachers reported on other related issues, such as general health, nutrition, and physical activity. Beginning in third grade, the children also were asked to provide information about their eating habits and physical activity.
- *Social and emotional development:* The ECLS-K assessments of social and emotional development focused on the skills and behaviors that contribute to social competence. Aspects of social competence include social skills (e.g., cooperation, assertion, responsibility, self-control) and problem behaviors (e.g., impulsive reactions, verbal and physical aggression). Parents and teachers were the primary sources of information on children's social competence and skills in kindergarten and first grade. The measurement of children's social and emotional

development at grades three, five, and eight included instruments completed by the children themselves along with data reported by parents and teachers.

- *Cognitive development:* In kindergarten and first grade, the ECLS-K focused on three broad areas of competence: language and literacy, mathematics, and general knowledge of the social and physical worlds. Starting in third grade, a science assessment replaced the general knowledge assessment. In the higher grades, children's cognitive skills were expected to have advanced beyond the levels covered by the kindergarten and first-grade assessments; for this reason, a new set of assessment instruments was developed for third grade, for fifth grade, and again for eighth grade. Some of the assessment items were retained from one round to the next to support the development of longitudinal score scales in each subject area. The skills measured in each of these domains are a sample of the typical and important skills that are taught in American elementary schools and that children are expected to learn in school. The ECLS-K was developed to describe the behaviors, skills, and knowledge within broad cognitive domains that are most relevant to school curricula at each grade level and to measure children's growth from kindergarten to eighth grade. The ECLS-K assessment framework was based on current curricular domain frameworks for reading, mathematics, science, and social studies, as well as on existing assessment frameworks, such as those used in the National Assessment of Educational Progress. (See chapter 18.)

The cognitive assessments were developed through extensive field testing and analysis of item performance. The final items were selected based on their psychometric properties and content relevance. Children's knowledge and skills in the natural and social sciences were measured in the general knowledge subdomain in kindergarten and first grade. The contents of this subtest, classified as science and social sciences, surveyed children's knowledge and understanding of relevant concepts. The science assessment used from third grade on measured children's knowledge in life science, physical science, and Earth science.

- Each direct cognitive domain subtest consisted of a routing test and second-stage tests that were tailored to different skill levels. All children

were first administered a short routing test of domain-specific items having a broad range of complexity or difficulty levels. Performance on the routing test was used to determine the appropriate second-stage assessment form to be administered next to the child. The use of multilevel forms for each domain subtest minimized the chances of administering items that were all very easy or all very difficult for a given child. The assessments were administered in one-on-one, untimed sessions with a trained child assessor. If necessary, the session could take place over multiple periods.

Birth cohort (ECLS-B). The ECLS-B direct child assessment relied on instruments considered "gold standards" in the field. However, adaptations were necessary to take these instruments from a laboratory or clinic setting to a home setting. The ECLS-B child assessment was designed for ease of and flexibility in administration while at the same time being psychometrically and substantively sound. The key instruments used in the study were a shortened research edition of the BSID-II, called the Bayley Short Form-Research Edition (BSF-R), the NCATS, the Two Bags Task, an attachment measure—the TAS-45, and Bruininks-Oseretsky Test of Motor Proficiency and Movement Assessment Battery for Children.

- *Cognitive development and fine and gross motor skills:* The BSID-II is considered the gold standard for assessing early childhood development (ages 1 to 42 months). In the 9-month and 2-year collections, children's cognitive development, as well as their receptive and expressive language skills, were assessed using an adaptation of the mental scale of the BSID-II. Children retrieved hidden toys and looked at picture books, and their production of vowel-consonant combinations was noted. Fine and gross motor skills were assessed using an adaptation of the motor scale of the BSID-II. Children grasped small objects and were observed crawling and walking. The study had intended to field the entire Bayley assessment, as it was originally expected to take about 20 minutes to complete. However, a field test of the 9-month ECLS-B data collection revealed that it actually required an average of 40 minutes to complete. As a result, modifications were implemented to the original BSID-II. The ECLS-B contractor, Westat, worked with experts to identify a reduced-item set that could be administered in less time and could produce reliable, valid scores equivalent to the full set of Bayley items. The BSF-R took

approximately 25 minutes to administer. Because the BSF-R was not appropriate for children older than 42 months of age, a new direct child cognitive assessment was developed for use in the preschool and kindergarten collections. These assessments were patterned after the ECLS-K assessments and incorporated items from the ECLS-K, as well as other published assessments, such as the preLas 2000, Test of Early Mathematics Ability, Third Edition (TEMA 3), and the Peabody Picture Vocabulary Test, Third Edition (PPVT-III). The cognitive domains covered in the preschool-kindergarten assessments were early reading and mathematics skills. The preschool collection also included a measure of children's color knowledge, which involved asking the children to name the colors of each bear presented to them in picture format. Children's fine and gross motor skills were measured using the Bruininks-Oseretsky Test of Motor Proficiency and Movement Assessment Battery for Children. To assess fine motor skills, children were asked to copy a series of forms (e.g., circle, triangle, square) that were first drawn by an assessor and to build a structure with blocks that was first demonstrated by the assessor. To assess gross motor skills, children were asked to hop, skip, jump backwards, and balance on one foot.

Because the NCATS is only appropriate for children up to 36 months of age, the Two Bags Task was used in the 2-year and preschool data collections. The Two Bags Task is a simplified version of the Three Bags Task that was used successfully in such large-scale studies as the Early Head Start Research and Evaluation Project and is intended to capture children's socioemotional functioning. It is a semistructured activity completed by the parent and child in interaction. During this 10-minute task, the parent-child dyad is asked to play with two different sets of toys, each placed within a separate numbered bag. In the 2-year collection, bag number 1 contained a children's picture book and bag number 2 contained a set of dishes. In the preschool collection, bag number 1 also contained a children's picture book but bag number 2 contained PlayDoh. The rating scales provide information on parents' behaviors during the interaction (parental sensitivity, intrusiveness, simulation of cognitive development, positive regard, negative regard, and detachment) and children's behaviors during the interaction (child

engagement of parent, sustained attention, and negativity toward parent).

In the preschool and kindergarten collections, information on children's socioemotional functioning was collected indirectly through questions asked of parents and teachers.

- *Children's security of attachment:* The TAS-45 is a modified version of the Attachment Q-Sort (AQS), a widely used observational measure of children's security of attachment. It includes 45 items describing children's behaviors. After being in the home with the child and parent for several hours, the ECLS-B assessors completed a task in which they indicated whether each of the 45 behaviors applied to the child and how strongly the behavior either applied or did not apply, based upon their observations of the child in the home. These items/behaviors cluster around common attachment-related constructs, such as "cooperativeness," "independence," or "attention-seeking." Nine clusters, or "hot spots," were identified in the data. These hot spots, along with a traditional attachment classification (Avoidant, Secure, Ambivalent, and Disorganized) and traditional security and dependency scores were developed from the TAS-45. The TAS-45 was only administered in the 2-year data collection.

Data Collection and Processing

The ECLS-K compiled data from four primary sources: children, children's parents/guardians, teachers, and school administrators. Data collection began in fall 1998 and continued through spring 2007. Self-administered questionnaires, one-on-one assessments, and telephone or in-person interviews were used to collect the data. Westat conducted all rounds of data collection from kindergarten through eighth grade.

The ECLS-B compiled data from multiple sources, including administrative records, children, parents, nonparental care providers, teachers, and NCES school universe files. Data collection began in 2001 and continued through 2008. The primary modes of data collection were an in-person home visit during which parent respondents were interviewed and children were directly assessed. Self-administered questionnaires and telephone interviews also were used to collect data. Westat was the 9-month and 2-year data collection contractor. RTI International conducted the preschool and kindergarten data collections.

Reference dates. For the ECLS-K, baseline data for the fall were collected from September through December

1998. For the ECLS-B, baseline data were collected from October 2001 through December 2002.

Data collection. The ECLS-K and the ECLS-B are discussed separately.

Kindergarten cohort (ECLS-K). The data collection schedule for the ECLS-K was based on a desire to capture information about children as critical events and transitions were occurring rather than measuring these events retrospectively. A large-scale field test of the kindergarten and first-grade assessment instruments and questionnaires was conducted in 1995–96. This field test was used primarily to collect psychometric data on the ECLS-K assessment item pool and to evaluate questions in the different survey instruments. Data from this field test were used to develop the routing and second-stage tests for the ECLS-K kindergarten and first-grade direct cognitive assessment battery and to finalize the parent, teacher, and school administrator instruments. A pilot test of the systems and procedures, including field supervisor and assessor training, was conducted in April and May 1998 with 12 elementary schools in the Washington, DC, metropolitan area. Modifications to the data collection procedures, training programs, and systems were made to improve efficiency and reduce respondent burden. Modifications to the parent interview to address some issues raised by pilot test respondents were also made at this time.

Data on the kindergarten cohort were collected twice during the base year of the study—once in the beginning (fall) and once near the end (spring) of the 1998–99 school year. The fall 1998 data collection obtained baseline data on children at the very beginning of their exposure to the influences of school, providing measures of the characteristics and attributes of children as they entered formal school for the first time. The data collected in spring 1999, together with the data from the beginning of the school year, are used to examine children’s first encounter with school. Data were collected from the child, the child’s parents/guardians, and teachers in both fall and spring. Data were collected from school administrators in the spring. For the fall 1998 and spring 1999 collections, all child assessment measures were obtained through untimed CAPI, administered one-on-one by the assessor to the child. The assessment was normally conducted in a school classroom or library and took approximately 50 to 70 minutes per child. Children with a primary home language other than English (according to school records) were first administered an English language screener (OLDS) to determine whether their English language skills were sufficient enough to take the cognitive assessments in English.

Children who fell below the cut score for the OLDS and whose language was Spanish were administered a Spanish-language version of the OLDS and the ECLS-K mathematics assessment translated into Spanish, and they had their height and weight measured. Children who fell below the cut score and whose language was neither English nor Spanish had only their height and weight measured. (A child was administered the OLDS in each round of data collection until he or she passed it; the OLDS was no longer used after the spring first grade data collection because by then most children demonstrated sufficient English language skills to be assessed in English.) Most of the parent data were collected by computer-assisted telephone interviewing (CATI), though some of the interviews were collected through CAPI when respondents did not have a telephone or were reluctant to be interviewed by telephone. All kindergarten teachers with sampled children were asked to fill out self-administered questionnaires providing information on themselves and their teaching practices. For each of the sampled children they taught, the teachers also completed a child-specific questionnaire. In the spring, school administrators were asked to complete a self-administered questionnaire that included questions on the school characteristics and environment, as well the administrator’s own background. Also, in the spring, the special education teachers or service providers of children in special education were asked to complete a self-administered questionnaire about the children’s experiences in special education and about their own background. In addition, school staff members were asked to complete a student record abstract after the school year closed.

In fall 1999, when most of the kindergarten cohort had moved on to first grade, data were collected from a 30 percent subsample of the cohort. The direct child assessment was administered during a 12-week field period (September–November 1999). The parent interview was administered between early September and mid-November 1999; it averaged 35 minutes, and was conducted primarily by telephone.

Spring data collections (first grade, third grade, fifth grade, and eighth grade) included direct child assessments, parent interviews, teacher and school questionnaires, student record abstracts, and facilities checklists. As in other rounds, the child assessments were administered with CAPI (March–June 2000 for the first-grade collection, March–June 2002 for the third-grade collection, February–June 2004 for the fifth-grade collection, and March–June 2007 for the eighth-grade collection), while both CATI and CAPI were used for the parent interview (March–July 2000 for first grade, March–July 2002 for third grade, February–June 2004 for fifth grade, and March–June

2007 for eighth grade). Self-administered questionnaires were used to gather information from teachers, school administrators, and student records (March–June 2000 for first grade and March–June 2002 for third grade, but field staff prompted by telephone for the return of these materials through October 2000 and October 2002, respectively. For the fifth grade, data collection was between February and June 2004. For the eighth grade, data collection was between March and June 2007.).

A continuous quality assurance process was applied to all data collection activities. Data collection quality control efforts began with the development and testing of the CATI and CAPI applications and the contractor's Field Management System. As these applications were programmed, extensive testing of the system was conducted. Quality control processes continued with the development of field procedures that maximized cooperation and thereby reduced the potential for nonresponse bias. Quality control activities also were practiced during training and data collection. During the original assessor training, field staff practiced conducting the parent interview in pairs and practiced the direct child assessment with kindergarten children brought to the training site for this purpose. In later data collection periods, experienced staff used a home study training package while new staff were trained in classroom sessions. After data collection began, field supervisors observed each assessor conducting child assessments and made telephone calls to parents to validate the interview. Field managers also made telephone calls to the schools to collect information on the school activities for validation purposes.

Birth cohort (ECLS-B). A field test of the ECLS-B instruments and procedures was conducted in the fall of 1999. The design featured many different tasks. For example, while in the home, a field staff member had to complete approximately 11 discrete tasks, and each task had special skill requirements. Early in the field test, NCES and the ECLS-B contractor found several problems regarding the complexity of the home visit: while separately no one task was difficult, the total data collection protocol was complex, so it was necessary to simplify these tasks in order to reduce the burden on field staff and to ensure the reliable and valid administration of all tasks. As a result, several modifications were made to the original data collection design.

A second field test of the ECLS-B instruments and procedures began in September 2000. A field test sample was drawn consisting of 1,060 children born between January and April 2000. Home visits were

conducted when the children were 9 months old and again when they were 18 months old. Results from this field test indicated that the changes to the design that resulted from the first field test were successful.

The ECLS-B schedule called for information to be gathered on the children and from the parents during an in-home visit. The children's mother or primary caregiver was the respondent for the parent interview at each round of data collection. Child assessments were conducted in the child's home by the trained ECLS-B assessors at every round of data collection as well. Resident fathers (defined as the spouse or partner of the female parent respondent) were asked to complete a self-administered questionnaire with questions regarding their involvement in their children's lives in the 9-month, 2-year, and preschool data collections. Biological, non-resident fathers were asked to complete a self-administered questionnaire in the 9-month and 2-year data collections if the mother gave permission for him to be contacted. In the 2-year and preschool data collections, information was collected from children's primary nonparental care providers through a telephone interview. Direct observations to assess child care quality also were conducted by trained observers for a subsample of children with regular nonparental care. In the kindergarten 2006 collection, the child care provider telephone interview used in the preschool collection was again fielded for children who had not yet entered kindergarten. A wrap-around care and education provider telephone interview (WECEP) was introduced in this collection to obtain information on children's before- and after-school care arrangements for those children who were in kindergarten. The WECEP was used in the kindergarten 2007 collection as well. Observations of care settings were not conducted in the kindergarten collections. Teachers of children in kindergarten in 2006 and 2007 were asked to complete a self-administered questionnaire similar to those used in the ECLS-K that asked about the child's classroom, the child's behaviors and performance in the classroom, and their own background. Although the ECLS-B did not include a school administrator questionnaire, information on children's schools was obtained from the NCES school universe files, the Common Core of Data (CCD) for public schools and the Private School Survey (PSS) for private schools.

The ECLS-B 9-month data collection began in October 2001 and continued through December 2002. The 2-year data collection began in January 2003 and continued through April 2004. While the 9-month and 2-year data collection schedules were designed to collect information on children as close as possible to the date on which they turned the age of interest for the collection (i.e., 9 months and 2 years), the collection

schedules for the preschool and kindergarten rounds were changed to correspond with an academic calendar. Thus, the preschool wave of data collection began in late August 2005 and ended in mid-July 2006. The kindergarten 2006 collection began in fall 2006 through spring 2007. The kindergarten 2007 collection began in fall 2007 through spring 2008. In all collections, CAPI was the principal mode of data collection for the parent interview. Self-administered questionnaires were used to gather information from the resident father, nonresident father, and teacher. A self-administered questionnaire was used to obtain information on potentially sensitive topics from the parent respondent at 9 months and 2 years; starting with the preschool collection, potentially sensitive items were administered using audio computer-assisted self-interviewing technology (ACASI). Data were collected from the child by several means: a series of structured, standardized activities were scored in the home by the field interviewer; structured interactions with the parent were videotaped for later coding; physical measurements were obtained; and behavior was observed throughout the home visit.

Child-parent interactions were assessed by NCATS at the 9-month data collection, and again by the Two Bags Task at the 2-year and preschool data collections. In all cases, the ECLS-B videotaped these structured interactions. Although it is more typical for a health or social service professional to complete NCATS via live coding (i.e., while the interaction is occurring), the ECLS-B field staff needed to observe and score 73 items of parent and child behavior. Given the other tasks the field staff had to learn and complete, live coding would have limited the number of scales that could realistically be used, thereby reducing the amount of information that could be gathered. The videotapes were coded along all scales.

Data were collected from child care providers by means of CATI. A subset of child care providers was sampled for on-site observations in the 2-year and preschool collections; observers recorded data in booklets, and child care center directors completed a self-administered paper questionnaire.

Editing. Within the CATI/CAPI instruments, the ECLS-K and ECLS-B respondent answers were subjected to both “hard” and “soft” range edits during the interviewing process. Responses outside the soft range of reasonably expected values were confirmed with the respondent and entered a second time. For hard-range items, out-of-range values were usually not accepted. If the respondent insisted that a response outside the hard range was correct, the assessor could enter the information in a comments data file. Data

preparation and project staff reviewed these comments. Out-of-range values were accepted if the comments supported the response.

Consistency checks were also built into the CATI/CAPI data collection. When a logical error occurred during an interview, the assessor saw a message requesting verification of the last response and a resolution of the discrepancy. In some instances, if the verified response still resulted in a logical error, the assessor recorded the problem either in a comment or in a problem report.

The overall data editing process consisted of running range edits for soft and hard ranges, running consistency edits, and reviewing frequencies of the results. Where applicable, these steps also were implemented for hard-copy questionnaire instruments, videotaped instruments, and observational instruments.

Estimation Methods

Data were weighted to account for differential probabilities of selection at each sampling stage and to adjust for the effects of nonresponse. A hot-deck imputation methodology was used to impute missing values for all components of SES in the ECLS-K and ECLS-B. Imputation also was implemented for child assessment proficiency-level variables and free/reduced-price school lunch data in the ECLS-K.

Weighting. Weighting in the ECLS-K and ECLS-B is discussed separately.

Kindergarten cohort (ECLS-K). Several sets of weights were computed for each of the seven rounds of data collection (fall kindergarten, spring kindergarten, fall first grade, spring first grade, spring third grade, spring fifth grade, and spring eighth grade). These weights include cross-sectional weights for analyses of data from one time point, as well as longitudinal weights for analyses of data from multiple rounds of the study. Unlike surveys that have only one type of survey instrument aimed at one type of sampling unit, the ECLS-K is a complex study with multiple types of sampling units, each having its own survey instrument. Each type of unit was selected into the sample through a different mechanism: children were sampled directly through a sample of schools; parents of the sampled children were automatically included in the survey; all kindergarten teachers and administrators in the sampled schools were included; and special education teachers were included in the sample if they taught any of the sampled children. Each sampled unit had its own survey instrument: children were assessed directly using a series of cognitive and physical assessments; parents were interviewed with a parent instrument;

teachers filled out at least two different types of questionnaires, depending on the round of data collection and on whether they were regular or special education teachers; and school principals reported their school characteristics using the school administrator questionnaire. The stages of sampling, in conjunction with different nonresponse levels at each stage and the diversity of survey instruments, required that multiple sampling weights be computed for use in analyzing the ECLS-K data.

Weight development was driven by three factors: (1) how many points in time would be used in analysis (i.e., whether the analysis would be longitudinal or cross-sectional); (2) what level of analysis would be conducted (e.g., child, teacher, or school); and (3) what source of data would be used (e.g., child assessment, teacher questionnaire, parent questionnaire).

For the kindergarten rounds of data collection, weights were computed in two stages. In the first stage, base weights were computed. The base weights are the inverse of the probability of selecting the unit. In the second stage, base weights were adjusted for nonresponse. Nonresponse adjustment cells were generated using variables with known values for both respondents and nonrespondents. Chi-squared Automatic Interaction Detector (CHAID) analyses were conducted to identify the variables most highly related to nonresponse. Once the nonresponse cells were determined, the nonresponse adjustment factors were calculated as the reciprocals of the response rates within the selected nonresponse cells. Beginning with the first grade round of data collection, a third stage called raking was introduced into the weight development process to remove the variability due to the subsampling of schools and children who changed schools (i.e., movers). In this stage, child weights were raked to sample-based control totals computed using the base year child weights adjusted for nonresponse.

The base weight for each school is the inverse of the probability of selecting the PSU in which the school is located multiplied by the inverse of the probability of selecting the school within the PSU. The base weights for eligible schools were adjusted for nonresponse; this was done separately for public and private schools.

The base weight for each child in the sample is the school nonresponse-adjusted weight for the school attended multiplied by a poststratified within-school student weight (total number of students in the school divided by the number of students sampled in the school). The poststratified within-school weight was calculated separately for Asian/Pacific Islander and non-Asian/Pacific Islander children because different

sampling rates were used for these two groups. Within a school, all Asian/Pacific Islander children have the same base weights and all non-Asian/Pacific Islander children have the same base weights. The parent weight, for use with analysis of parent data, is the base child weight adjusted for nonresponse to the parent interview. Again, these adjustments were made separately for students in public and private schools. The teacher weight, for use with child-level analysis that includes teacher data from the child-level questionnaire specific to the sample child, is the base child weight adjusted for nonresponse to the teacher child-level questionnaire. Weights for child-level analysis were developed for every round of data collection. Weights for analysis at the school and teacher levels (i.e., weights that allow for the generation of national estimates of schools educating kindergarten-age children and kindergarten teachers) were developed only for the kindergarten data collections. The sample is not representative of schools or teachers after the kindergarten year,

Birth cohort (ECLS-B). Several sets of weights were computed for each round of data collection. Weights are used to adjust for disproportionate sampling, survey nonresponse, and noncoverage of the target population when analyzing complex survey data. The weights are designed to eliminate or reduce bias that would otherwise occur with analyses of unweighted data. The ECLS-B weights were developed in three steps: First, base weights were calculated using the overall selection probabilities; next, weights were adjusted for survey nonresponse; finally, raking was used to adjust for undercoverage and to improve the precision of survey estimates.

The base weight gives the approximate representation of each sampled birth record. The base weight for a given birth record was calculated as the reciprocal of the overall probability of selection, computed as the product of each stage's probability of selection. These overall probabilities of selection and base weights are used to compute analysis weights for all ECLS-B children in each round of data collection.

Next, base weights were adjusted for survey nonresponse. A selected set of variables related to child and family characteristics was used to construct nonresponse adjustment cells for each set of weights. Respondents and nonrespondents were compared on the characteristics selected based on analyses using segmentation modeling via CHAID. In the first round of data collection, data from the birth certificate were used to compare respondents and nonrespondents, because these data were available for all sampled cases regardless of participation status. In later collections,

respondents and nonrespondents were compared on both birth certificate data and data collected in prior rounds. A nonresponse adjustment factor was calculated for each cell as the ratio of the sum of weights for eligible cases in the cell to the sum of weights for eligible and responding cases in the cell. Finally, the nonresponse-adjusted weights were raked to 11 dimensions to ensure that sums of weights matched known population totals, thus correcting for survey undercoverage. The 11 dimensions were selected because of their substantive interest as well as their relationship to response propensity, as indicated by the CHAID modeling and also some preliminary logistic regression analyses.

The development of the ECLS-B weights was a sequential process. The 9-month weights were developed first, starting with the base weights; the 2-year weights were developed as adjustments to the 9-month weights; the preschool weights started with the 2-year weights, the kindergarten 2006 weights started with the preschool weights, and the kindergarten 2007 weights started with the kindergarten 2006 weights. A set of weights also was developed to allow for analysis of children in their first year of kindergarten, whether that year was in the 2006 collection or the 2007 collection. These weights were developed as adjustments to the preschool weights. As there are three main components in the 9-month round (parent interview data, child assessment data, and father data) and five or more components in each of the following rounds (parent interview data, child assessment data, father data, child care provider data, child care observation data, teacher data, and/or school data, depending on the round), several sets of weights were developed, taking into account the level of nonresponse for the different components and combinations of completed components that would be of most analytic interest. For example, the 9-month parent-father-child weight is valid for cases for which all three components are complete and adjusts for nonresponse to these components, whereas the 9-month parent weight is valid for all cases for which the parent component is complete, regardless of whether the child or father components are complete, and adjusts for nonresponse to the parent interview. Both cross-sectional weights for analysis of data at one round and longitudinal weights for analysis of data from multiple rounds of the study were computed.

Scaling. IRT was employed in the ECLS-K and ECLS-B to calculate scores that could be compared both within a round and across rounds, regardless of which second-stage form a student took. 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.

Imputation.

Kindergarten cohort (ECLS-K). In the ECLS-K, SES component variables were computed for the base-year, spring first-grade, spring third-grade, spring fifth-grade, and spring eighth-grade rounds. The percentages of missing data for the education and occupation variables were small (2 to 11 percent in the base year, 4 to 8 percent in the spring of first grade, 2 to 3 percent in the spring of third grade, 1 to 2 percent in the spring of fifth grade; and 3 percent in the spring of eighth grade); however, the household income variable had a higher rate of missing data (28.2 percent in the base year and 11 to 33 percent in the spring of first grade, depending on whether a detailed income range or the exact household income was requested; in the spring of third grade, 11.1 percent of cases had missing data for the detailed income range; this percentage was 8.1 percent of cases in the spring of fifth grade and 7.0 percent of cases in the spring of eighth grade). A standard (random selection within class) hot-deck imputation methodology was used to impute for missing values of all SES components in all years. From the spring of first grade on, the initial step in the imputation procedure was to fill in missing values from information gathered during an earlier interview with a parent, if one had taken place. If no prior data were available, standard hot-deck imputation was used.

The SES component variables were highly correlated, so a multivariate analysis was more appropriate to examine the relationship between the characteristics of donors and nonrespondents. For the base year, CHAID was used to divide the data into cells based on the distribution of the variable to be imputed, as well as to analyze the data and determine the best predictors. These relationships were used for imputation in later rounds of the ECLS-K.

The variables were imputed in sequential order and separately by type of household. For households with both parents present, the mother's and father's variables were imputed separately. If this was not the case, an "unknown" or missing category was created as an additional level for the CHAID analysis. As a rule, no imputed value was used as a donor. In addition, the same donor was not used more than two times. The order of the imputation for all the variables was from the lowest percentage missing to the highest.

Imputation for occupation involved two steps. First, the labor force status of the parent was imputed, whether the parent was employed or not. Then the parent's occupation was imputed only for those parents whose

status was identified as employed, either through the parent interview or the first imputation step. The variable for income was imputed last using a three-stage procedure; if a respondent provided partial information about income, this was used in the imputation process.

Imputation was also employed for variables related to the percentage of children in a school who received free or reduced-price lunch. Not all school principals answered all three questions that were used to derive the composite variables indicating the percentage of students in the school who received free lunch and the percentage who received reduced-price lunch: total school enrollment, number of children eligible for free lunch, and number of children eligible for reduced-price lunch. Prior to the fifth grade, if these three source variables had missing values, the composites were filled in with values computed using the most recent CCD data if they were not missing from the CCD, or left missing if they were missing from the CCD. Beginning in fifth grade, missing values in the composite variables were imputed. Missing values in the source variables, however, were not imputed.

A two-stage procedure was used for imputing school lunch composites. First, if a school had nonmissing values for the school lunch composites in kindergarten, first grade, and third grade, missing values for the fifth grade were filled in with values from previous years. A similar procedure was employed for eighth grade, which was first if a school had nonmissing values for a prior round, eighth grade was filled with the value from the previous year. Second, data still missing after this initial step were imputed using a hot-deck methodology. Imputation cells were created using the Title I status of the school and school longitude and latitude. School data that were imputed by hot deck are generally transfer schools with few sample children.

Birth cohort (ECLS-B). As in the ECLS-K, variables used to derive the SES composite variable were imputed using a hot-deck methodology. These variables include mother's and father's education, mother's and father's occupation, and income range. Imputation cells were defined by respondent characteristics that were the best predictors of the variables to be imputed, as determined using a CHAID analysis. Hot-deck imputation was done in a sequential order, separately, by type of household (female single parent, male single parent, and both parents present). As with the ECLS-K, missing data from a later round were first filled with data obtained in a prior round, if available. For households with both parents present, the mother's and father's variables were imputed separately. Imputed as well as reported

values were used to define imputation cells; missing values for donor characteristics were treated as a separate category. No imputed value was used as a donor. No donor was used more than once. The order of hot-deck imputation for all variables was from the lowest percentage missing to the highest.

Future Plans

The ECLS-K:2011 will follow students from kindergarten in 2010 through fifth grade in 2015. Because it is designed to allow for comparisons between the 2010-11 cohort and the cohort of kindergartners included in the ECLS-K, by design the ECLS-K:2011 is very similar to the ECLS-K and includes most of the same components. Some changes of note are the introduction of a basic reading skills assessment to be administered to all children, regardless of primary home language; a Spanish basic reading skills assessment to be administered to Spanish-speaking children who do not pass an English language screener; and the replacement of the final and gross motor skills assessments with an assessment of children's executive functions, a set of interdependent processes that work together to accomplish purposeful, goal-directed activities and include working memory, attention, inhibitory control, and other self-regulatory processes.

5. DATA QUALITY AND COMPARABILITY

Sampling Error

The estimators of sampling variances for the ECLS statistics take the ECLS complex sample design into account. Both replication and Taylor Series methods can be used to accurately analyze data from the studies. The paired jackknife replication method using 90 replicate weights can be used to compute approximately unbiased estimates of the standard errors of the estimates. (The fall first-grade subsample in the ECLS-K uses 40 replicate weights.) When using the Taylor Series method, a different set of stratum and first-stage unit (i.e., PSU) identifiers should be used for each set of weights. Both replicate weights and Taylor series and identifiers are provided as part of the ECLS-K and ECLS-B data files.

Design effects.

Kindergarten cohort (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. The median child-level design effect is 4.7 for fall kindergarten and 4.1 for spring kindergarten. The median child-level

design effect for spring third grade, spring fifth grade, and spring eighth grade is 3.3, 4.0, and 3.1, respectively.

The size of the ECLS-K design effects is largely a function of the number of children sampled per school. With about 20 children sampled per school, an intraclass correlation of 0.2 might result in a design effect of about 5. The median design effect is 3.4 for the panel of students common to both the fall and spring of kindergarten, and the lower median design effect is due to the smaller cluster size in the panel. The ECLS-K design effects are slightly higher than the average of 3.8 (with the exception of the spring third-grade collection and spring eighth-grade collection design effect) that was anticipated during the design phase of the study, both for estimates for proportions and for score estimates.

The median teacher-level design effect is 2.5 for both the fall and spring of kindergarten. This design effect is lower than the child-level design effects because the number of responding teachers per school is relatively small. The design effect for teachers is largely a result of selecting a sample using the most effective design for child-level statistics, rather than a design that would be most effective for producing teacher-level statistics.

The median school-level design effect is 1.6. Design effects were not computed for items from the teacher and school administrator questionnaires in the spring of first, third, fifth, and eighth grades because no teacher or school weights were computed for any of the ECLS-K years after kindergarten.

A multilevel analysis was carried out to estimate components of variance in the fall- and spring-kindergarten cognitive scores associated with (1) the student, (2) the school, (3) the data collection team leader, and (4) the individual test administrator. This secondary analysis was motivated by Westat's earlier finding of larger-than-expected design effects. In addition, the impact of parent's education on the above sources of variance was also estimated.

Birth cohort (ECLS-B) As noted above, several sets of weights were developed for use with different combinations of survey components that are of analytic interest. Design effects were computed for different survey estimates produced using these different weights. Using the parent weights, the median parent-level design effect is 2.1 for the 9-month data collection, 2.4 for the 2-year collection, 2.1 for the preschool collection, 2.0 for the kindergarten 2006 collection, and 2.2 for the kindergarten 2007 collection.

The median design effects for other weights across all components and all rounds of collections ranges from a low of 1.2 for the 2-year weight connected to response to the child care observation (W22PO) weight and a high of 4.2 for the 9-month weight connected to response to the 9-month child assessment (W1CO) weight.

It is noted that the design effects for assessment estimates are higher than the design effects for some other types of estimates. This can be due to either naturally occurring higher intracluster correlations for assessment estimate items or interviewer effects. In the ECLS-B, where the general relationship between interviewer and cluster is one-to-one, the two are difficult, if not impossible, to disentangle. Similar observations about the design effects for assessment estimates were made in the ECLS-K data.

Nonsampling Error

In order to reduce nonsampling error, the survey design phase included focus groups and cognitive laboratory interviews for the purposes of assessing respondent knowledge topics, comprehension of questions and terms, and item sensitivity. The design phase also entailed testing of the CAPI instrument and a field test that evaluated the implementation of the survey.

Another potential source of nonsampling error is respondent bias that occurs when respondents systematically misreport (intentionally or unintentionally) information in a study. One potential source of respondent bias in the ECLS surveys is social desirability bias. If there are no systematic differences among specific groups under study in their tendency to give socially desirable responses, then comparisons of the different groups will accurately reflect differences among the groups. An associated error occurs when respondents give unduly positive assessments about those close to them. For example, parents may give more positive assessments of their children's experiences than might be obtained from institutional records or from the teachers.

Potentially, response bias may also be introduced in the responses of teachers about each individual student. For example, each teacher filled out a survey for each of the sampled children they taught in which they answered questions on the child's socioemotional development in the ECLS-K and ECLS-B. Since the base-year and first-grade surveys in the ECLS-K and the kindergarten surveys in the ECLS-B were first conducted in the fall, it is possible that the teachers did not have adequate time to observe the children, and thus some of their responses may be influenced by their expectations based on which groups (e.g., sex, race,

ELL status, disability) the children belonged to. In order to minimize bias, all items were subjected to multiple cognitive interviews and field tests, and actual teachers were involved in the design of the cognitive assessment battery and questionnaires. NCES also followed the criteria recommended in a working paper on the accuracy of teachers' judgments of students' academic performances (see Perry and Meisels 1996).

As in any survey, respondent bias may be present in the ECLS-K and ECLS-B. It is not possible to state precisely how such bias may affect the results. NCES has tried to minimize some of these biases by conducting one-on-one, untimed assessments, and by asking some of the same questions about the sampled child of both teachers and parents.

Coverage error. Undercoverage occurs when the sampling frame used does not fully reflect the target population of inference. By designing the ECLS-K child assessment to be both individually administered and untimed, both coverage error and bias were reduced. Individual administration decreases problems associated with group administration, such as children slowing down and not staying with the group or simply getting distracted. The advantage of having untimed exams was that the study was able to include most children with special needs and/or who needed some type of accommodation, such as children with a learning disability, with hearing aids, etc. The only children who were excluded from the study were those who were blind, those who were deaf, those whose IEP clearly stated that they were not to be tested, and non-English-speaking children who were determined to lack adequate English or Spanish language skills to meaningfully participate in the ECLS-K battery. Exclusion from the direct child assessment did not exclude children from other parts of the study (e.g., teacher questionnaire, parent interview).

For the ECLS-B, the 9-month target population is all infants born in the United States in 2001 to mothers 15 years of age and older who were not adopted prior to, and who were alive during, the 9-month data collection period. The target population for later rounds of collection also excludes children who died or moved abroad permanently. Concern about noncoverage in the ECLS-B relates mainly to a few PSUs where births were sampled from hospital frames. In addition, the main sampling frame consisted of birth certificates available from state registrars. This sampling frame failed to cover unregistered births, but the number of these was thought to be negligible, according to the National Center for Health Statistics.

Nonresponse error.

Kindergarten cohort (ECLS-K). Overall, 880 of the 1,280 eligible schools (69.4 percent weighted) agreed to participate in the fall kindergarten study. Due to the lower-than-expected cooperation rate for public schools in the fall of the base year, 74 additional public schools were included in the sample as substitutes for schools that did not participate. These schools were included in order to meet the target sample sizes for students. Substitute schools are not included in the school response rate calculations.

A nonresponse bias analysis was conducted to determine if substantial bias was introduced due to school nonresponse in the ECLS-K. Five different approaches were used to examine the possibility of bias in the ECLS-K sample. First, weighted and unweighted response rates for schools, children, parents, teachers, and school administrators were examined to see whether there were large response rate differences by characteristics of schools (e.g., urbanicity, region, school size, percent Black, Hispanic, and other race/ethnicity students, grade range) and children (e.g., sex, age, race/ethnicity). Second, estimates based on the ECLS-K respondents were compared to estimates based on the full sample. The distributions of schools by school type, urbanicity, and region, and the distributions of enrollment by kindergarten type (public vs. private), race/ethnicity, urbanicity, region, and eligibility for free and reduced-price lunch were compared for the responding schools and all the schools in the sampling frame. Third, estimates from the ECLS-K were compared with estimates from other data sources (e.g., Current Population Survey, National Household Education Surveys Program, Survey of Income and Program Participation). Fourth, estimates using the ECLS-K unadjusted weights were compared with estimates using the ECLS-K weights adjusted for nonresponse. Large differences in the estimates produced with these two different weights would indicate the potential for bias. Fifth, and last, simulations of nonresponse were conducted. The results of these analyses are summarized in the ECLS-K user's manuals. Findings from these analyses suggest that there is no bias due to school nonresponse.

A total of 940 of the 1,280 originally sampled schools participated during the base year of the study. This translates into a weighted response rate (weighted by the base weight) of 74 percent for the base year of the study. The weighted child base-year survey response rate was 92 percent (i.e., 92 percent of the children were assessed at least once during kindergarten). The weighted parent base-year unit response rate was 89 percent (i.e., a parent interview was completed at least once during kindergarten). Thus, the overall base-year

response rate for children was 68 percent (74 percent of schools x 92 percent of sampled children) and the base-year overall response rate for the parent interview was 66 percent (74 percent of schools x 89 percent of parents of sampled children). About 76 percent of children and 72 percent of parents eligible for the eighth grade data collection (spring 2007) participated.

Birth cohort (ECLS-B). Response rates for all rounds of data collection are determined first and foremost by completion of the corresponding round's parent CAPI instrument. The parent CAPI instrument was chosen as the primary vehicle for determining the overall response rate because there were very few cases (e.g., 0.3 percent at 9 months and 0.06 percent at 2 years) in which other components of the study (e.g., direct child assessments or father questionnaires) were completed but the parent interview was not. All response rates are computed at the child level. In the 9-month data collection, all sampled children were eligible except those children who died before the home visit occurred, children born to mothers younger than 15 years old, children who were adopted before the age of 9 months, and children who were removed from the sample as part of a cost reduction process in February 2002. Response rates for subsequent rounds are conditioned on the completion of a prior round parent interview. For example, the 2-year-round response rate is conditioned on the completion of the 9-month parent interview; all sampled children whose parents completed the 9-month parent component were eligible except those children who had died before the 2-year home visit occurred and children who had moved abroad permanently. For the preschool-year data collection, approximately 9,850 cases with completed 2-year parent interviews, and an additional 50 AI/AN cases with completed 9-month parent interviews, were fielded and considered eligible (approximately 100 children were removed from the sample because they had died or moved abroad permanently). For the kindergarten 2006 collection, there were about 7,000 parent interviews. For the kindergarten 2007 collection, there were about 1,900 parent interviews.

Response rates are also calculated for the other components of the ECLS-B: the child assessments; the resident and nonresident father questionnaires; the care provider interview; the child care observation; the teacher questionnaire; and the school data. Response rates for these other components are conditioned on the completion of the parent interview in all rounds of the

ECLS-B. Only cases with completed parent interviews were assigned weights for the other components of the study.

Table 1. Weighted unit response rates for all children and children sampled in kindergarten, by questionnaire and data collection: Various years 1998–2004

Data collection	All children		Children sampled in kindergarten	
	Child assessment	Parent interview	Child assessment	Parent interview
Fall kindergarten	89.9	85.3	†	†
Spring kindergarten	88.0	83.9	†	†
Spring first grade	87.2	83.5	88.0	84.5
Spring third grade	80.1	76.9	80.8	77.8
Spring fifth grade	83.9	88.3	84.7	89.1

† Not applicable.

SOURCE: Tourangeau, K., Burke, J., Le, T., Wan, S., Weant, M., Brown, E., Vaden-Kiernan, N., Rinker, E., Dulaney, R., Ellingsen, K., Barrett, B., Flores-Cervantes, I., Zill, N., Pollack, J., Rock, D., Atkins-Burnett, S., Meisels, S., Bose, J., West, J., Denton, K., Rathbun, A., and Walston, J. (2001). *ECLS-K, Base Year Public-Use Data File, Kindergarten Class of 1998-99: Data Files and Electronic Code Book (Child, Teacher, School Files), and User's Manual* (NCES 2001-029REV). National Center for Education Statistics, U.S. Department of Education. Washington, DC. Tourangeau, K., Burke, J., Le, T., Wan, S., Weant, M., Nord, C., Vaden-Kiernan, N., Bissett, E., Dulaney, R., Fields, A., Byrne, L., Flores-Cervantes, I., Fowler, J., Pollack, J., Rock, D., Atkins-Burnett, S., Meisels, S., Bose, J., West, J., Denton, K., Rathbun, A., and Walston, J. (2002). *User's Manual for the ECLS-K First-Grade Public-Use Data Files and Electronic Codebook* (NCES 2002-135). National Center for Education Statistics, U.S. Department of Education. Washington, DC. Tourangeau, K., Brick, M., Le, T., Wan, S., Weant, M., Nord, C., Vaden-Kiernan, N., Hagedorn, M., Bissett, E., Dulaney, R., Fowler, J., Pollack, J., Rock, D., Weiss, M.J., Atkins-Burnett, S., Hausken, E.G., West, J., Rathbun, A., and Walston, J. (2004). *User's Manual for the ECLS-K Third-Grade Public-Use Data Files and Electronic Codebook* (NCES 2004-001). National Center for Education Statistics, Institute of Education Sciences, U.S. Department of Education. Washington, DC. Tourangeau, K., Nord, C., Lê T., Pollack, J.M., and Atkins-Burnett, S. (2006). *Early Childhood Longitudinal Study Kindergarten Class of 1998–99 (ECLS-K), Combined User's Manual for the ECLS-K Fifth-Grade Data Files and Electronic Codebooks* (NCES 2006-032). National Center for Education Statistics, Institute of Education Sciences, U.S. Department of Education. Washington, DC.

In the 9-month data collection, the weighted completion rate for the parent CAPI instrument was 74.1 percent (table 2). The weighted completion rates for the child assessment, resident father questionnaires,

and nonresident father questionnaires were 95.6, 76.1, and 50.0 percent, respectively.

In the 2-year data collection, the weighted completion rate for the parent CAPI instrument was 93.1 percent. The weighted completion rates for the child assessment, resident father questionnaires, nonresident father questionnaires, child care provider interview, and child care observation (CCO) component were 94.2, 77.7, 39.8, 70.0, and 51.3 percent, respectively. The longitudinal weighted response rates for the parent CAPI instrument, child assessment, and all father questionnaires were 69.0, 65.0, and 48.7 percent, respectively.

In the preschool data collection, the weighted completion rate for the parent CAPI instrument was 91.3 percent. The weighted completion rates for the child assessment, resident father questionnaires, child care provider interview, and CCO component were 98.3, 87.7, 87.4, and 56.8 percent, respectively. The longitudinal weighted response rates for the parent instrument, child assessment, resident father questionnaires, child care provider interview, and CCO component were 63.1, 62.0, 55.3, 55.1, and 35.8, respectively.

In the kindergarten 2006 data collection, the weighted response rate for the parent instrument was 91.8 percent. The weighted unit response rate for the kindergarten 2006 child assessment was 98.6 percent. The weighted unit response rate for the teacher survey for ECLS-B children with a completed parent interview who were enrolled in kindergarten or higher in 2006-07 and were not homeschooled was 75.6 percent; the weighted unit response rate for school data for these same children was 95.9 percent. The overall weighted unit response rate for the parent component after the kindergarten 2006 data collection was 58.0 percent. The longitudinal weighted unit response rates for the parent, child, teacher, and school components after the kindergarten 2006 collection were 58.0, 57.2, 43.8, and 55.6 percent, respectively.

The weighted unit response rate for the kindergarten 2007 parent interview was 92.5 percent. The weighted unit response rate for the kindergarten 2007 child assessment was 99.4 percent. The weighted unit response rate for the teacher survey for ECLS-B children with a completed parent interview who were enrolled in kindergarten or higher in 2007-08 and were not homeschooled was 77.4 percent; the weighted unit response rate for school data for these same children was 96.9 percent. The longitudinal weighted unit response rate for the parent component after the kindergarten 2007 data collection was 53.7 percent.

The overall weighted unit response rates for the child, teacher, and school components after the kindergarten 2007 collection were 53.3, 41.5, and 52.0 percent, respectively.

Table 2. Weighted unit response rates for all children in the ECLS-B, by survey and component: Various years 2001–2007

Component	Kinder- Kindergarten				
	9-month	2-year	Pre-school	2006	2007
Parent CAPI	74.1	93.1	91.3	91.8	92.5
Child assessment	95.6	94.2	98.3	98.6	99.4
Resident father	76.1	77.7	87.7	†	†
Nonresident father	50.0	39.8	†	†	†
Child care provider	†	70.0	87.4	†	†
Child care observation	†	51.3	56.8	†	†

† Not applicable.

SOURCE: Denton Flanagan, K., and McPhee, C. (2009). *The Children Born in 2001 at Kindergarten Entry: First Findings From the Kindergarten Data Collections of the Early Childhood Longitudinal Study, Birth Cohort (ECLS-B)* (NCES 2010-05). National Center for Education Statistics, Institute of Education Sciences, U.S. Department of Education. Washington, DC. Jacobson Chernoff, J., Flanagan, K. D., McPhee, C., and Park, J. (2007). *Preschool: First Findings From the Preschool Follow-up of the Early Childhood Longitudinal Study, Birth Cohort (ECLS-B)* (NCES 2008-025). National Center for Education Statistics, Institute of Education Sciences, U.S. Department of Education. Washington, DC. Nord, C., Edwards, B., Andreassen, C., Green, J. L., and Wallner-Allen, K. (2006). *Early Childhood Longitudinal Study, Birth Cohort (ECLS-B), User's Manual for the ECLS-B Longitudinal 9-Month–2-Year Data File and Electronic Codebook (NCES 2006–046)*. National Center for Education Statistics, Institute of Education Sciences, U.S. Department of Education. Washington, DC. Nord, C., Edwards, B., Hilpert, R., Branden, L., Andreassen, C., Elmore, A., Sesay, D., Fletcher, P., Green, J.L., Saunders, R., Dulaney, R., Reaney, L., and Flanagan, K.D. (2004). *User's Manual for the ECLS-B Nine-Month Restricted-Use Data File and Electronic Codebook (NCES 2004-092)*. National Center for Education Statistics, Institute of Education Sciences, U.S. Department of Education. Washington, DC.

An analysis was conducted to assess the potential bias in survey estimates due to unit or item nonresponse for the various components of the survey. This evaluation consisted of several types of comparisons. First, data obtained from children's birth certificates were compared between cases in the sampling frame and sample respondents; data for sample respondents were

weighted first using base weights and then using final weights. These comparisons were made for respondents to the parent CAPI interview, the father questionnaires, the child care provider interview, and the CCO component. In another analysis, birth certificate and survey data were compared between 9-month respondents (using final 9-month weights) and 2-year respondents (using both final 9-month weights and final 2-year weights). These comparisons were done for respondents to the parent CAPI interview, the child assessments, the father questionnaires, and the child care provider interview. The analysis found little or no evidence of potential for bias due to unit nonresponse. Differences between sample respondents and sample frame data were generally small and largely corrected by nonresponse corrections and other adjustments to the base weights. An evaluation comparing the demographic characteristics of respondents and nonrespondents for selected items with less than an 85 percent response rate found no evidence of potential for bias due to item nonresponse. Similar analyses of nonresponse bias were conducted for later rounds of data collection, with no evidence found for bias due to item nonresponse.

Measurement error. In addition to the potential clustering effects related to shared parent SES within schools (described in “Design effects,” above), there was a concern in the ECLS-K that the individual mode of administration might inject additional and unwanted variance into both the individual and 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 high design effects if the individual assessors differed systematically in their modes of administration. It was found, however, that the component of variance associated with the individual test administration effect was negligible in all cognitive areas and thus had little or no impact on the design effects.

A potential area for measurement error occurs with the NCATS and Two Bags Task components of the ECLS-B home visit. The parent-child interactions for these two components of the study were videotaped and coded later. The process of coding the tapes, however, is not problem-free. The videotape of the interaction must be of high quality to ensure valid coding. For example, field staff needed to tape the very beginning of the interaction and should not interrupt it. The task of coding is further complicated by the coding staff’s experience. Like the ECLS-B home visit field staff, the NCATS and Two Bags Task coders did not, for the most part, possess an extensive background in child development. Training the coding staff to reach

90 percent reliability proved difficult at times and often required additional training.

6. CONTACT INFORMATION

For content information about the ECLS project, contact:

Gail M. Mulligan
Phone: (202) 502-7491
E-mail: gail.mulligan@ed.gov

Mailing Address:

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

7. METHODOLOGY AND EVALUATION REPORTS

General

Anderson, C., Fletcher, P., and Park, J. (2007). *Early Childhood Longitudinal Study, Birth Cohort (ECLS-B) Psychometric Report for the 2-year Data Collection*. (NCES 2007-084). Washington, DC: National Center for Education Statistics.

Najarian, M., Snow, K., Lennon, J., and Kinsey, S. (2010). *Early Childhood Longitudinal Study, Birth Cohort (ECLS-B), Preschool–Kindergarten 2007 Psychometric Report* (NCES 2010-009). National Center for Education Statistics, Institute of Education Sciences, U.S. Department of Education. Washington, DC.

Najarian, M., Pollack, J.M., and Sorongon, A.G. (2009). *Early Childhood Longitudinal Study, Kindergarten Class of 1998-99 (ECLS-K), Psychometric Report for the Eighth Grade*. (NCES 2009-002). Washington, DC: National Center for Education Statistics.

Pollack, J., Atkins-Burnett, S., Najarian, M., and Rock, D. (2005). *Early Childhood Longitudinal Study, Kindergarten Class of 1998-99 (ECLS-K), Psychometric Report for the Fifth Grade*. (NCES 2006-036rev). Washington, DC: National Center for Education Statistics.

- Pollack, J., Atkins-Burnett, S., Rock, D., and Weiss, M. (2005). *Early Childhood Longitudinal Study, Kindergarten Class of 1998-99 (ECLS-K) Psychometric Report for the Third Grade*. (NCES 2005-062). Washington, DC: National Center for Education Statistics.
- Rock, D., and Pollack, J. (2002). *Early Childhood Longitudinal Study, Kindergarten Class of 1998-99 (ECLS-K) Psychometric Report for Kindergarten through First Grade*. (NCES 2002-05). Washington, DC: National Center for Education Statistics.
- Tourangeau, K., Lê, T., Nord, C., and Sorongon, A.G. (2009). *Early Childhood Longitudinal Study, Kindergarten Class of 1998-99 (ECLS-K) Eighth-Grade Methodology Report*. (NCES 2009-003). Washington, DC: National Center for Education Statistics.
- Tourangeau, K., Lê, T., and Nord, C. (2006). *Early Childhood Longitudinal Study, Kindergarten Class of 1998-99 (ECLS-K), Fifth Grade Methodology Report*. (NCES 2006-037). Washington, DC: National Center for Education Statistics.
- Tourangeau, K., Brick, M., Byrne, L., Lê, T., Nord, C., West, J., and Germino Hausken, E. (2005). *Early Childhood Longitudinal Study, Kindergarten Class of 1998-99 (ECLS-K) Third Grade Methodology Report*. (NCES 2005-018). Washington, DC: National Center for Education Statistics.
- Tourangeau, K., Brick, M., Le, T., Wan, S., Weant, M., Nord, C., Vaden-Kiernan, N., Hagedorn, M., Bissett, E., Dulaney, R., Fowler, J., Pollack, J., Rock, D., Weiss, M.J., Atkins-Burnett, S., Hausken, E.G., West, J., Rathbun, A., and Walston, J. (2004). *User's Manual for the ECLS-K Third-Grade Public-Use Data Files and Electronic Codebook* (NCES 2004-001). National Center for Education Statistics, Institute of Education Sciences, U.S. Department of Education. Washington, DC.
- Tourangeau, K., Nord, C., Le, T., Wan, S., Bose, J., and West, J. (2002). *User's Guide to the Longitudinal Kindergarten-First Grade Public-Use Data File* (NCES 2002-149). National Center for Education Statistics, Institute of Education Sciences, U.S. Department of Education. Washington, DC.
- Tourangeau, K., Burke, J., Le, T., Wan, S., Weant, M., Nord, C., Vaden-Kiernan, N., Bissett, E., Dulaney, R., Fields, A., Byrne, L., Flores-Cervantes, I., Fowler, J., Pollack, J., Rock, D., Atkins-Burnett, S., Meisels, S., Bose, J., West, J., Denton, K., Rathbun, A., and Walston, J. (2002). *User's Manual for the ECLS-K First Grade Public-Use Data Files and Electronic Code Book* (NCES 2002-135). National Center for Education Statistics, Institute of Education Sciences, U.S. Department of Education. Washington, DC.