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

Website: http://nces.ed.gov/ecls/kindergarten.asp
Updated: May 2017

1. OVERVIEW

EARLY CHILDHOOD LONGITUDINAL SAMPLE SURVEYS: KINDERGARTEN COHORT

ECLS-K collects data from:

- Children
- Parents/guardians
- Teachers
- School administrators

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) is following a sample of kindergartners in the 2010–11 school year through the fifth grade. The ECLS provides comprehensive and reliable datasets with information about the ways in which children are prepared for school and how children develop within 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 the associations and interactions of a wide range of family, school, community, and child characteristics with 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 a *birth* cohort study (ECLS-B)—and each of these has its own components. This chapter describes the first *kindergarten* cohort study, ECLS-K. For details on the 2011 *kindergarten* cohort study, see the handbook chapter for ECLS-K:2011. Details on the *birth* cohort study can be found in the ECLS-B handbook chapter.

The Early Childhood Longitudinal Study, Kindergarten Class of 1998–99. 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. 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 kindergarten and first grade, a 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. Beginning with the third-grade data collection, children 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. 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 questionnaires. 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 rounds 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 three 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 questionnaires. 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 for particular populations (e.g., English language learners), 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 ended. 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. In each round of data collection except eighth grade, field staff members completed a checklist providing information on school facilities. 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 checklist.

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 springs of first grade (2000), third grade (2002), fifth grade (2004), and eighth grade (2007).

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 child characteristics are associated with early success in school and later development and academic performance. The longitudinal nature of the studies enables researchers to study children's achievement and growth in reading and mathematics knowledge and skills, as well as 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.

3. KEY CONCEPTS

Item Response Theory (IRT) scale scores. The ECLS-K direct cognitive assessment employed a two-stage design. As such, within any given domain, children received a routing set of items (stage 1) and then, based on their performance on the routing items, proceeded to a second set of items of a certain difficulty level (stage 2). Because not all children received all items, the assessment scores in the ECLS-K study were modeled using Item Response Theory (IRT). Based on children's performance on the items they received, an ability estimate (theta) was derived for each domain. 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 useful in identifying cross-sectional differences among subgroups in overall achievement levels and provide a summary measure of achievement useful for correlational analysis with status variables. The IRT scale scores are also used as longitudinal measures of overall growth. Gain scores may be calculated by subtracting children's scale scores at two points in time.

Standardized scores (T-scores). These scores are IRT based and derived from the child's ability estimate theta. They provide norm-referenced measurements of achievement; that is, estimates of achievement level relative to the population as a whole. A higher mean T-score for a particular subgroup indicates that the group's performance was 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 derived from the child's ability estimate theta. They provide information on performance on clusters of items of similar difficulty along the overall scale. The scores measure the probability of correct responses in each cluster and can take on any value between 0 and 1. Because each proficiency probability score targets a particular set of skills, they can be used 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. Office of Management and Budget guidelines for collecting information on race and ethnicity were followed. A respondent could select one or more of five dichotomous race categories when reporting their own race or that of their child. Each respondent was also asked to identify whether he or she (as well as the study child if the respondent was a parent) was Hispanic. The study data files include several variables indicating race and ethnicity. There are six dichotomous race variables indicating whether a respondent or study child was of a certain race (White, Black, Asian, Native Hawaiian or Other Pacific Islander, American Indian or Alaska Native, and more than one race) as well as one dichotomous ethnicity variable indicating whether a respondent or study child was Hispanic. These variables were used to create one race/ethnicity composite variable with mutually exclusive categories: White, not Hispanic; Black, not Hispanic; Hispanic of any race; Asian, not Hispanic; Native Hawaiian or Other Pacific Islander, not Hispanic; American Indian or Alaska Native, not Hispanic; and Two or more races, not Hispanic.

Socioeconomic status (SES). Each ECLS-K data file includes a measure of SES reflecting the SES of a child's household at the time of data collection. The components used to create the SES variable are father/male guardian's education, mother/female guardian's education, father/male guardian's occupational prestige, mother/female guardian's occupational prestige, and household income. In households with two mothers or two fathers, education and occupational prestige for both mothers/fathers were used. 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.

4. SURVEY DESIGN

Target Population

Kindergarten children enrolled in school in the United States during the 1998–99 school year are the target population for the ECLS-K cohort.

Sample Design

The ECLS-K followed a nationally representative sample of children from kindergarten through the spring of 2007, when most of the children were in eighth grade. Study children were included in data collections after the kindergarten year even if they were no longer in the modal grade for children who were in kindergarten in the 1998–99 school year.

Base-year (i.e., kindergarten) collections. A nationally representative sample of 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 included both children in kindergarten for the first time and children who were repeating kindergarten. 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.

In the kindergarten year, the study can also be used to general estimates of schools educating kindergarten-age children and kindergarten teachers. After the base year, the data are only representative at the child level.

The sample design for the ECLS-K was a dual-frame, multi-stage sample. First, 100 PSUs were selected from an initial frame of approximately 1,335 PSUs, representing counties or groups of contiguous counties. The 24 PSUs with the largest measures of size (where the measure of size was the number of 5-year-olds, taking into account a factor for oversampling Asian/Pacific Islander 5-year-olds) were included in the ECLS-K sample with certainty. The remaining PSUs were partitioned into 38 strata of roughly equal measures of size. The frame of these 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 Islanders, Blacks, or Hispanics), 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 without replacement in each stratum, with probability proportional to size and with known joint probability of inclusion of the pair.

School selection occurred within the sampled PSUs. Public schools were sampled from a public school 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 the spring of 1998 to include newly opened schools that were not included in the CCD and PSS used for initial sample selection (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 the fall of 1998, approximately 23 kindergarten children were selected, on average, from each of the sampled schools. Asian/Pacific Islander children and private schools were oversampled.

A nationally representative sample of 22,782 children enrolled in 1,277 kindergarten programs during the 1998–99 school year was selected to participate in the ECLS-K.

Fall first-grade collection. The fall first-grade collection was designed to enable researchers to measure the extent of summer learning loss, examine the factors associated with such loss, and to better understand the relationships of school and home characteristics with children's learning. The fall data collection consisted of a 30 percent sample of schools containing approximately 25 percent of the base-year students eligible to participate in the second year. 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 collection. 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 the United States 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 data collection, approximately 18,080 children were eligible to participate (14,250 public school students

and 3,840 private school students). Student freshening brought 170 first-graders into the ECLS-K sample.

Spring third-grade collection. 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 (i.e., those students who transferred from an originally sampled school) flagged for following in the spring of first grade, with the addition of movers whose home language was not English (who were followed at 100 percent). For the spring third-grade data collection, approximately 16,670 children were eligible to participate (13,170 public school students and 3,500 private school students).

Spring fifth-grade collection. In fifth grade, four groups of children were not followed for data collection. 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 neither first-grade nor third-grade data had been collected.

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 maximized the amount of longitudinal data available for key analytic groups. Children whose home language was not English (language minority (LM) children) continued to be sampled at higher rates (between 15 and 50 percent for base-year LM respondents, and between 15 and 75 percent for LM children in the first-grade freshened sample).

For the spring fifth-grade data collection, approximately 12,030 children were eligible to participate (9,570 public school students and 2,460 private school students).

A sampling strategy first implemented for the fifth-grade data collection was the subsampling of eligible children for the administration of mathematics and science questionnaires. While a child-level reading teacher questionnaire was fielded for all children included in the fifth-grade data collection, half of the children were selected to have a child-level questionnaire filled out by their mathematics teachers and the other half were selected to have a child-level questionnaire filled out by their science teachers.

Spring eighth-grade collection. Children 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 were ineligible for 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). As in fifth grade, half of the children were selected to have a child-level questionnaire filled out by their mathematics teachers and the other half were selected to have a child-level questionnaire filled out by their science teachers.

For the spring eighth-grade data collection, approximately 11,930 children were eligible (9,480 in public schools and 2,450 in private schools).

Assessment Design

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 on a number of dimensions, including physical, socioemotional, 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.

- *Socioemotional development:* The ECLS-K indirect assessments of socioemotional 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. The same assessments were fielded in both kindergarten and first grade. 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 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 NAEP chapter).

  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.

- 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.

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.

Reference dates. Baseline data were collected from September through December 1998 and March through July 1999.

Data collection. 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 study 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 address some issues raised by

pilot test respondents were also made to the parent interview 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, can be used to examine children's first experiences with elementary 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 and special education teachers in the spring. For the fall 1998 and spring 1999 collections, all child assessment measures were obtained through untimed assessments, administered one-on-one to the child by an assessor using a CAPI application. 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 whose scores on the screener 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 and psychomotor assessments translated into Spanish. They also had their height and weight measured. Children whose scores on the screener 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 conducted in person 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. The teachers also were asked to complete a child-specific questionnaire for each of the sampled children they taught. 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 ended.

In the fall of 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 in first grade, third grade, fifth grade, and eighth grade included direct child assessments, parent interviews, and teacher and school administrator questionnaires. In the spring of first grade, third grade, and fifth grade student record abstracts and facilities checklists were also completed. 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 fifth grade, data collection occurred between February and June 2004. For eighth grade, data collection occurred between March and June 2007).

A continuous quality assurance process was also applied to all data collection activities. Specifically, extensive testing of the CATI and CAPI applications and the data collection contractor's Field Management System was conducted. Field procedures that maximized cooperation and thereby reduced the potential for nonresponse bias were developed. Field staff participated in trainings lasting several days in which they were instructed on proper administration of the parent interview and child assessments. During these trainings, 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. 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.

Editing. Within the CATI/CAPI instruments, the ECLS-K 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 items with hard ranges, out-of-range values (i.e., those that were not considered possible) were usually not accepted. If the respondent insisted that a response outside the hard range was correct, the interviewer could enter the information as a comment. 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, if possible. 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.

Estimation Methods

Weighting. Weights are used to adjust for disproportionate sampling at each sampling stage, 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.

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 for 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 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 interview).

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. 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 computed for each school is the inverse of the probability of selecting the PSU in which

the school was 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 the child 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. Again, these adjustments were made separately for students in public and private schools.

Weights for child-level analysis were developed for every round of data collection. Each child-level weight was developed to be used with data from specific survey components and has adjustments for nonresponse to those specific components. For example, there is a weight to be used in analysis of parent data that is the child base weight adjusted for nonresponse to the parent interview. 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.

Scaling. To maximize information on which each estimate of ability is based, the majority of the direct cognitive assessment scores computed for the study are based on item response theory (IRT). IRT uses patterns of correct and incorrect answers to compute estimates on a scale that may be compared across different assessment forms. IRT was employed in the ECLS-K to calculate ability estimates and then derive assessment scores from those ability estimates that can be compared both within a round and across rounds, regardless of which second-stage form a student was administered. 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. Socioeconomic status (SES) component variables were imputed 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). The household income variable had a higher rate of missing data (28.2 percent in the base year; 11 to 33 percent in the spring of first grade, depending on whether respondents were asked for income using a detailed set of income range categories or for exact household income; and 11.1 percent, 8.1 percent, and 7.0 percent of cases had missing data for the detailed income range in the spring of third grade, the spring of fifth grade, and the spring of eighth grade, respectively. A standard (random selection within class) hot-deck imputation methodology was used to impute for missing values of all the 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 appropriate to examine the relationship between the characteristics of respondents (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 data 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; if a respondent provided partial information about income, this information was used in the imputation process.

Imputation was also employed for composite 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 in the CCD, or left missing if they were missing in 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 the school lunch composite variables in fifth and eighth grade. First, if a school had nonmissing values for the school lunch composites in a prior round of data collection, missing values for the current round were filled in with the value from a 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 were generally transfer schools with few sample children. Imputation was done for the free- and reduced-price lunch composite variables only for children in public schools.

Future Plans

Currently, NCES does not have plans to collect any more data from the students in the ECLS-K cohort or their families. NCES is continuing its program of longitudinal studies of young children with the ECLS-K:2011. More information can be found in the ECLS-K:2011 handbook chapter.

5. DATA QUALITY AND COMPARABILITY

Sampling Errors and Weighting

The sample of children enrolled in kindergarten in the United States in 1998–99 selected for the ECLS-K is just one of many possible samples that could have been selected. Therefore, estimates produced from the ECLS-K sample may differ from estimates that would have been produced from other samples. This type of variability is called sampling error because it results from collecting data on a sample of children, rather than all children enrolled in kindergarten in the United States in 1998–99. The standard error is a measure of variability due to sampling when estimating a statistic. Standard errors can be used as a measure of the precision expected from a particular sample.

For a complex sample design such as the one employed in the ECLS-K, replication and Taylor Series methods have been developed to correctly estimate variance. These methods take into account the clustered, multistage sampling design and the use of differential sampling rates to oversample targeted subpopulations. Both replication and Taylor Series methods can be used to accurately analyze data from the studies. The paired jackknife replication method using replicate weights can be used to compute approximately unbiased estimates of the standard errors of the estimates. 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 identifiers are provided as part of the ECLS-K data files.

Design effects. An important analytic consideration is how the statistical efficiency of survey estimates from a complex sample survey such as the ECLS-K compares with estimates that would have been obtained had a simple random sample (SRS) of the same size been used. In a stratified clustered design, stratification generally leads to a gain in efficiency over simple random sampling, but clustering has the opposite effect because of the positive intracluster correlation of the units in the cluster. The basic measure of the relative efficiency of the sample is the design effect. 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 for the base year is 1.6. Design effects were not computed for items from the teacher and school administrator questionnaires in the springs of first, third, fifth, and eighth grades because no teacher or school weights were computed for any of the ECLS-K years after kindergarten.

Nonsampling Error

Nonsampling error is the term used to describe variations in the estimates that may be caused by population coverage limitations, as well as data collection, processing, and reporting procedures. The sources of nonsampling errors are typically nonresponse, differences in respondents' interpretations of the meaning of the questions, response differences related to the particular time the survey was conducted, and mistakes in data preparation. Steps are taken to reduce nonsampling error.

In order to reduce nonsampling error associated with respondents misunderstanding what was being asked of them, the survey design phase included focus groups and cognitive laboratory interviews for the purposes of assessing respondent knowledge of different topics covered in the instruments, comprehension of questions and terms, and item sensitivity. The design phase also included testing of the CAPI/CATI instruments and a field test that evaluated the implementation of the survey in order to reduce the potential for error to be introduced as a result of errors in administration.

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.

Response bias may also be present in the responses teachers provide 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. Since the base-year and first-grade surveys in the ECLS-K 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 the children's outward characteristics (e.g., sex, race, ELL status, disability status). 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, response bias may be present in the data for ECLS-K. 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 from which a sample is selected does not fully reflect the target population of inference. The potential for coverage error in the ECLS-K was reduced by using a school-level frame derived from universe surveys of all schools in the United States and master lists of all kindergartners enrolled in sampled schools.

By designing the child assessments to be both individually administered and untimed, both coverage error and bias were reduced. Untimed, individually administered exams allowed the study 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 direct child assessments were those who needed a Braille assessment, those who needed a sign language interpreter, those whose IEP clearly stated that they were not to be tested, and non-English-speaking children who lacked 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).

Nonresponse error. Bias can exist in survey data if too few sampled units responded for the data collected to

be representative of the target population or if nonresponse is significantly higher for sample entities with certain characteristics. The unit response rate is a round-specific rate in that it indicates the proportion of the eligible sample responding to a survey at a particular time point. For a longitudinal study such as the ECLS-K, it is also useful to calculate a longitudinal response rate, also called an overall unit response rate, which takes into account response for all rounds of collection. A total of 940 of the 1,280 originally sampled schools participated in at least one round of data collection during the base year of the study. This translates into a weighted school 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 times 92 percent of sampled children) and the base-year overall response rate for the parent interview was 66 percent (74 percent of schools times 89 percent of parents of sampled children). See table ECLS-K-1 for details on weighted response rates.

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.

Table ECLS-K-1. Weighted child-level response rates and overall response rates for children in ECLS-K, by collection and selected components: Various years 1998–2007

Selected components	Kindergarten, 1998-99	Grade 1, 1999-2000	Grade 3, Spring 2002	Grade 5, Spring 2004	Grade 8, Spring 2007
Response rates					
Child assessment	88.0	87.2	80.1	83.9	75.0
Parent interview	83.9	83.5	76.9	88.3	70.9
School administrator questionnaire	85.9	75.9	65.5	76.4	72.5
Teacher-level questionnaire, classroom information (Part A)	86.9	77.6	61.7	79.3	73.8
Teacher-level questionnaire, teacher information (Part B)	89.7	77.0	61.6	—	—
Child-level teacher questionnaire (Part C)	85.9	77.4	62.0	78.7	72.5
Overall response rates					
Child assessment	65.1	56.8	45.5	38.2	28.6

Parent interview	62.1	51.8	39.9	35.2	25.0
School administrator questionnaire	63.6	48.2	31.6	24.1	17.5
Teacher-level questionnaire, classroom information (Part A)	64.3	49.9	30.8	25.0	18.4
Teacher-level questionnaire, teacher information (Part B)	66.4	51.1	31.5	—	—
Child-level teacher questionnaire (Part C)	63.6	49.2	30.5	24.0	17.4

— Not available. In the grade 5 and grade 8 collection, there was only one part in the teacher-level questionnaire. The response rate is reported in the teacher-level questionnaire, classroom information (Part A) row of the table.
NOTE: Overall response rates are the product of the school-level response rate from the base year (i.e., 74.0 percent) and the completion rates from each round of data collection after the base year.
SOURCE: National Center for Education Statistics. (2009). *Eighth-Grade Methodology Report* (NCES 2009-003), table 6-14. U.S. Department of Education. Institute of Education Sciences.

Measurement error. There was a concern in the ECLS-K that the individual mode of administration might introduce 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. 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. It was found 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.

6. CONTACT INFORMATION

For content information about the ECLS project, contact:

Jill McCarroll
Phone: (202) 304-2920
E-mail: jill.mccarroll@ed.gov

Mailing Address:
National Center for Education Statistics
Institute of Education Sciences
U.S. Department of Education
Potomac Center Plaza
550 12th Street SW
Washington, DC 20202

7. METHODOLOGY AND EVALUATION REPORTS

General

Mather, N., and Woodcock, R.W. (2001). *Woodcock-Johnson III Tests of Cognitive Abilities: Examiner's Manual*. Itasca, IL: Riverside Publishing.

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.

NIH. (n.d.). *Executive Function*. Retrieved February 09, 2013 from http://www.nihtoolbox.org/WhatAndWhy/Cognition/ExecutiveFunction/Pages/default.aspx.

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.

Zelazo, P.D. (2006). The Dimensional Change Card Sort (DCCS): A Method of Assessing Executive Function in Children. *Nature Protocols, 1*(1): 297-301.