Early Childhood Longitudinal Study, Kindergarten Class of 2010–11 (ECLS-K:2011)

Website: http://nces.ed.gov/ecls/kindergarten2011.asp
Updated: May 2017

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) 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.

EARLY CHILDHOOD LONGITUDINAL SAMPLE SURVEYS: KINDERGARTEN COHORTS

ECLS collects data from:

- Children
- Parents/guardians
- Before- and after-school care providers and preschool teachers
- Teachers
- School administrators

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 2011 *kindergarten* class of 2010–11 study ECLS-K:2011. For details on the first *kindergarten* cohort study, see the handbook chapter for ECLS-K. Details on the *birth* cohort study can be found in the ECLS-B handbook chapter.

The Early Childhood Longitudinal Study, Kindergarten Class of 2010–11. The ECLS-K:2011 collects data on children's cognitive, social, emotional, and physical development from the children, their families, classroom teachers, special education teachers, school administrators, and before- and after-school care providers. Information is also collected on the children's home environment, including home educational activities; their school environment, their classroom environment, including classroom curriculum; their teachers' background; and before- and after-school care in kindergarten.

Direct child assessments. The child assessment measures reading, mathematics, and science knowledge and skills, as well as executive function—executive function is "the capacity to plan, organize, and monitor the execution of behaviors that are strategically directed in a goal-oriented manner" (NIH, n.d.). The kindergarten science assessment was only administered in the spring. Also in the kindergarten and first-grade years, Spanish-

speaking English language learner (ELL) children who did not achieve a minimum score on assessment items measuring their basic English language skills had their Spanish early reading skills assessed.

In addition to the cognitive components, all children had their height and weight measured in all rounds of data collection. With the exception of the one-stage kindergarten science assessment, all direct cognitive assessments were two-stage assessments. For these assessments, the first stage was a routing section that included items covering a broad range of difficulty. A child's performance on the routing section in reading, mathematics, or science determined which one of three second-stage tests (low, middle, or high difficulty) the child was administered in that domain. The second-stage tests varied by level of difficulty so that a child would be administered questions appropriate to his or her demonstrated level of ability for each of these cognitive domains.

Parent interviews. Information is collected from parents/guardians at each round of data collection using computer-assisted interviews (CAIs). The parent interviews at each round ask about a variety of topics including family structure, family literacy practices, parental involvement in the child's education, nonparental care arrangements, household composition, family income, parent education level and employment, and other demographic indicators. Parents are also asked to report on their children's health, socioemotional well-being, and disability status.

Classroom teacher questionnaires. Teachers are asked to complete self-administered questionnaires at each round of data collection to provide information about the children they teach, the children's learning environment, and themselves. More specifically, they are asked about their own backgrounds, teaching practices, and experience. They are also asked to provide information on the classroom experiences for the sampled children they teach and to evaluate each sampled child on a number of critical cognitive and noncognitive dimensions.

Special education teacher questionnaires. Special education teachers and related service providers of sampled children who have an Individualized Education Program (IEP) are asked to provide information on the nature and types of services provided to the children, as well as on their own background and experience. Information is collected from special education teachers via self-administered questionnaires during spring data collections.

School administrator questionnaires. School administrators are asked to provide information on the physical, organizational, and fiscal characteristics of their schools, and on the schools' learning environment and programs. School administrators also provide information on their own background and experience. Information is collected from school administrators via self-administered questionnaires during spring data collections.

Before- and after-school care provider questionnaires. The kindergarten before- and after-school care (BASC) component collected important information about children's environments and experiences in nonparental care with regular before- or after-school care providers. Adults other than the child's parents/guardians who cared for the study child for at least 5 hours per week were asked to provide information such as the location where care was provided, children's activities while in care, characteristics of other children in care, and their own background and experience. The BASC component was only included during the spring kindergarten round.

Periodicity

The ECLS-K:2011 data have been collected in the fall and the spring of kindergarten (2010–11), the fall and the spring of first grade (2011–12), the fall and spring of second grade (2012–13), the spring of third grade (2014), and the spring of fourth grade (2015), with the final data collection planned for the spring of fifth grade (2015).

The fall direct child assessments are conducted from August through December and the spring direct child assessments are conducted from late March through June.

2. USES OF DATA

The ECLS-K:2011 provides information critical to informing policies that can respond sensitively and creatively to diverse learning environments. In addition, the ECLS-K:2011 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. The longitudinal nature of the study enables researchers to study children's achievement and growth in reading, mathematics, and science knowledge and skills. 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:2011 direct cognitive assessment employs 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 on the routing items, proceed to a second set of items of a certain difficulty level (stage 2). Because not all children received all items, the assessment scores were modeled using IRT. Based on children's performance on the items they received, an ability estimate (theta) is 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.

Race/ethnicity. Office of Management and Budget guidelines for collecting information on race and ethnicity were followed under which a respondent could select one or more of five dichotomous race categories when reporting their own race or that of their child. Each respondent additionally had 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. In later rounds, more detailed information was collected if the child was Hispanic, Native Hawaiian or Other Pacific Islanders, or another Asian subgroup.

Socioeconomic status (SES). The ECLS-K:2011 data file provides 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 was 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 the 2010–11 school year are the baseline for the ECLS-K:2011 cohort.

Sample Design

The ECLS-K:2011 is following a nationally representative cohort of children from kindergarten through the spring of 2016, when most of the children are expected to be in fifth grade.

Base-year (i.e., kindergarten) collections. Approximately 20,250 children in 1,320 schools (1,035 public and 285 private) were sampled and eligible for the base-year data collections of the ECLS-K:2011. These children were selected from both public and private schools. The sample includes children from different racial/ethnic and socioeconomic backgrounds. Asian/Pacific Islander (API) students were oversampled to assure that the sample included enough students of this race/ethnicity to be able to make accurate estimates for these students as a group.

The ECLS-K:2011 cohort was sampled using a multistage sampling design. The first-stage sampling frame for the ECLS-K:2011 was a list of the 3,141 counties in the United States. The county-level frame was used to form a list of primary sampling units (PSUs) from which a subset of PSUs was sampled. Ten PSUs with a large measure of size (defined as the number of 5-year-old children in the PSU) were included in the ECLS-K:2011 sample with certainty. The remaining PSUs were sampled using a stratified sampling procedure. They were grouped into 40 strata defined by MSA status, census geographic region, size class (defined using the measure of size), per capita income, and the race/ethnicity of 5-year-old children residing in the PSU (specifically the percent of 5-year-old APIs, the percent of 5-year-old Blacks, and the percent of 5-year-old Hispanics). Two PSUs were selected without replacement in each stratum, with probability proportional to size and with known joint probability of inclusion of the pair.

The second stage of sampling involved selecting samples of public and private schools that have kindergarten programs or that educate children of kindergarten age in an ungraded setting from within the sampled PSUs. The target for the number of schools participating in the base year of the study was 180 private and 720 public schools, for a total of 900 schools. In order to achieve this target number, approximately 280 private schools and 1,030 public schools were initially sampled from a frame of public schools and a frame of private schools constructed for the 2010 National Assessment of Education Progress (NAEP). The NAEP frame had not yet been updated and, therefore, was not final at the time it was obtained for use in the ECLS-K:2011. For this reason, a supllmental frame of newly opened schools and kindergarten programs was developed in the spring of 2010, and a supplemental sample of schools selected from that frame was added to the main sample of study schools. Schools were selected with probability proportional to size. The measure of size for schools was kindergarten enrollment adjusted to take into account the desired oversampling of APIs.

In the third stage of sampling, approximately 23 kindergartners were selected from a list of all enrolled kindergartners or students of kindergarten age being educated in an ungraded classroom in each of the sampled schools. As noted above, Asian/Pacific Islander students were oversampled to assure that the sample included enough students of this race/ethnicity to be able to make accurate estimates for these students as a group.

For the base year of the ECLS-K:2011, approximately 18,200 children enrolled in the originally sampled 970 schools participated.

First-grade collections. Two data collections were conducted in the 2011–12 school year, when the majority of the children were in first grade: one in the fall and one in the spring. The fall first-grade data collection was conducted with a subsample of 30 PSUs (out of the 90 PSUs selected for the base year of the study). This data collection included base-year respondents—those students in the base year who had a completed assessment or parent interview in at least one of the two rounds of kindergarten data collection—who attended the sample schools in those 30 PSUs during their kindergarten year. The spring first-grade data collection included base-year respondents in all 90 sampled PSUs. Due to the increased data collection costs associated with following students who transferred from their original sample school, in each round of data collection only a subsample of these students were followed into their new schools. About 5,230 children from about 350 sampled schools participated in the fall first-grade data collection, and about 15,130 children from about 990 sampled schools participated in the spring first-grade data collection.

Second-grade collections. The fall second-grade data collection included base-year respondents—those students who had a completed assessment or a completed parent interview in at least one of the two rounds of the kindergarten data collection—who attended schools within a subsample of 30 PSUs during their kindergarten year. This is the same subgroup of students who were included in the fall first-grade data collection. The spring second-grade data collection included base-year respondents who attended schools within all 90 sampled PSUs. Due to the increased data collection costs associated with following students who transferred from their original sample school, in each round of data collection only a subsample of these students were followed into their new schools. About 4,740 children from about 890 schools participated in the fall second-grade data collection, and about 13,850 children from about 2,330 schools participated in the spring second-grade data collection.

Assessments

A critical component of the ECLS-K:2011 is the assessment of children on a number of dimensions, including cognitive, physical, and socioemotional development. These domains were chosen because of their importance to success in school.

- *Cognitive development:* The ECLS-K:2011 direct cognitive assessment battery measured kindergartners' knowledge and skills in reading, mathematics, and science, as well as executive function. Because the ECLS-K:2011 is a longitudinal study, the assessments also were designed to allow for the measurement of growth in these domains across time. The longitudinal design of the ECLS-K:2011 required that the cognitive assessments be developed to support the measurement of change in knowledge and skills demonstrated by children from kindergarten entry through the spring of fifth grade.

  The ECLS-K:2011 reading and math specifications were based on the frameworks developed for the National Assessment of Educational Progress (NAEP). Although the NAEP assessments are administered starting in fourth grade, the specifications were extrapolated down to kindergarten based on current curriculum standards from several states and, for math, the

National Council of Teachers of Mathematics *Principles and Standards for School Mathematics*. The frameworks necessarily cover content strands applicable to a range of content at different grade levels, for example from number sense (i.e., basic knowledge of numbers) to algebra in mathematics. Content appropriate for most students in the targeted grade level was included in the assessments used in that grade. For example, in the kindergarten math assessment, the "algebra" content strand was assessed through children's recognition of patterns. While the assessments were designed to contain mostly items that assessed knowledge and skills at a kindergarten level, easier and more difficult items were included to measure the abilities of students performing below or above grade level.

The cognitive assessments were individually administered by trained assessors using computer-assisted technology and small easel test books containing the assessment items. The reading and mathematics assessments were administered in both the fall and spring data collections using two-stage adaptive tests. For each assessment, the first-stage was a routing section that included items covering a broad range of difficulty. A child's performance on the routing section determined which one of three second-stage tests (low, middle, or high difficulty) the child was administered. The second-stage tests varied by level of difficulty so that a child would be administered questions appropriate to his or her demonstrated level of ability for each of these cognitive domains. The purpose of this adaptive assessment design was to maximize accuracy of measurement while minimizing administration time.

Kindergarten science knowledge and skills were measured using a 20-item assessment that was administered only in the spring data collection. All students were administered the entire assessment. A two-stage design was not needed for science because the length of the test was relatively short with respect to both time (approximately 10 minutes) and the number of items. In all later rounds of data collection, science was administered using a 2-stage assessment, such as is described for reading and mathematics above.

- *Executive function.* Measures of executive function were included in the direct child assessment batteries to assess children's cognitive flexibility, working memory, and inhibition.

  The Dimensional Change Card Sort (Zelazo, 2006) is used to collect information on children's cognitive flexibility. In the version of this task used in the kindergarten and first-grade collections, children are asked to sort a series of 22 picture cards according to different rules. Each card has a picture of either a red rabbit or a blue boat. The children are asked to sort each card into one of two trays depending on the sorting rule they have been told. Beginning in the fall second-grade collection, the DCCS was no longer administered using the picture cards. Instead, a computer version of the DCCS that also captures children's reaction time was employed.

  The Numbers Reversed subtest of the Woodcock-Johnson III Tests of Cognitive Abilities (Mather and Woodcock 2001) assesses the child's working memory. It is a backward digit span task that requires the child to repeat an orally presented sequence of numbers in the reverse order in which the numbers are presented. Children are given sequences of increasing length (up to a maximum of eight numbers) until the child gets three consecutive number sequences incorrect or completes all number sequences.

- *Physical development:* Children's height and weight are being measured and body mass index (BMI) is being calculated at each data collection point in the ECLS-K:2011.

- *Socioemotional development:* The ECLS-K:2011 indirect assessments of socioemotional development focus 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 are the primary sources of information on children's social competence and skills in kindergarten.

Data Collection and Processing

The ECLS-K:2011 includes data from five primary sources: the students, their parents/guardians, their teachers, their schools, and their before-and after-school care providers. Data collection began in fall 2010 and will continue through spring 2016. Self-administered questionnaires, one-on-one assessments, and telephone or in-person interviews are being used to collect the data. Westat is the data collection contractor for all rounds of data collection.

Reference dates. For the ECLS-K:2011, baseline data were collected from September through December

2010 for the fall and from late March through June 2011 for the spring.

Data collection. Fall and spring data collections included direct child assessments, parent interviews, and teacher questionnaires. The spring kindergarten round included the before- and after-school care provider questionnaires. The fall second-grade and spring third-grade rounds included the hearing evaluation component, which is also planned for the spring fifth-grade round. Beginning in the spring of third grade, the child questionnaire was also included. Development of the ECLS-K:2011 survey instruments built upon those from the earlier ECLS studies and carried forward much of the same content and approaches. Development of the Before and After-School Care (BASC) questionnaire was based on the Wrap-Around Early Care and Education Provider (WECEP) interview from the ECLS-B. Development of the other survey instruments (i.e., direct child assessment, parent interview, and school staff questionnaires) was based on the instruments from the ECLS-K. Exceptions were the hearing evaluations and executive function components, which are new to the ECLS-K:2011 study.

In the fall of 2009, two field tests were conducted. These field tests served as the primary vehicle for (1) estimating the psychometric parameters of all items in the assessment battery item pool, (2) producing psychometrically sound and valid direct and indirect cognitive assessment instruments, (3) assessing the feasibility of screening children's vision and hearing during the national collection, and (4) obtaining valid assessments for both an English reading score for Spanish-speaking children not being assessed fully in English and an assessment of these children's early reading skills (e.g., letter recognition and sounds) in Spanish. Development of the survey instruments was also guided by advice given by the ECLS-K:2011 Technical Review Panel (TRP), the ECLS-K:2011 Content Review Panels (CRP), and other experts and consultants. Another field to test items for inclusion in the third-, fourth-, and fifth-grade assessments as well as the child questionnaire was conducted in the spring of 2013.

Two data collections were conducted in the 2010–11 school year, one in the fall and one in the spring. The fall and spring kindergarten rounds of the ECLS-K:2011 data collection included a direct child assessment with cognitive and physical measurement components. The components of the ECLS-K:2011 assessment administered to children who spoke a language other than English at home depended on the children's performance on a language screener used in the fall and spring data collections. The screener consisted of two tasks from the Preschool Language Assessment Scale (*pre*LAS 2000). All children also received the first 18 items of the reading assessment in English, regardless of their home language or performance on the *pre*LAS tasks. These items, plus two items from the *pre*LAS task (a total of 20 items), make up the section of the reading assessment referred to as the English basic reading skills (EBRS) section because they measure such skills. Once the EBRS items were administered, the cognitive assessments in English ended for children whose home language was not English and who did not achieve at least a minimum score on the language screener. Spanish-speaking children who did not achieve at least the minimum score on the screener were then administered a short reading assessment in Spanish that measured Spanish early reading skills (SERS), as well as the mathematics and executive function assessments that had been translated into Spanish. Children whose home language was one other than English or Spanish and who did not achieve at least the minimum score on the screener were not administered any of the remaining cognitive assessments beyond the EBRS. All children had their height and weight measured.

Parent interviews were conducted mostly by telephone, though the interview was conducted in-person for parents who did not have telephones or who preferred an in-person interview. The respondent to the parent interview was usually a parent or guardian in the household who identified himself or herself as the person who knew the most about the child's care, education, and health. During the later data collection rounds, interviewers attempted to complete the parent interview with the same respondent who answered the parent interview in the previous round, though another parent or guardian in the household who knew about the child's care, education, and health was selected if the prior-round respondent was not available.

The parent interviews were fully translated into Spanish before data collection began and could be administered by bilingual interviewers if parent respondents preferred to speak in Spanish. Because it was cost prohibitive to do so, the parent interviews were not translated into other languages. However, interviews could be completed with parents who spoke other languages by using an interpreter who translated from the English during the interview.

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 related 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. Before- and after-school caregivers identified in the fall kindergarten parent interview were asked to complete self-administered hard-copy questionnaires for the before- and after-school care (BASC) component of the ECLS-K:2011 during the spring kindergarten round. The BASC instruments asked about the characteristics of the child's care arrangement, as well as the provider's background and professional development activities. The provider with whom the child spent the most time on a weekly basis was the respondent for the care provider questionnaire, as well as for a child-level questionnaire with questions specifically about the study child. There were two versions of the care provider questionnaire, one for providers in center-based arrangements and one for providers in home-based arrangements. The administration of the different survey instruments in later grades was similar to the administration of those instruments in kindergarten, though the BASC questionnaires were not fielded again.

A continuous quality assurance process was applied to all data collection activities at all rounds. Data collection quality control efforts began with the development and testing of the CATI and CAPI applications and the data collection 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. 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:2011 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. The ECLS-K:2011 data are weighted to compensate for unequal probabilities of selection at each sampling stage and to adjust for the effects of school, teacher, before- and after-school care provider, child, and parent nonresponse. The sample weights to be used in the ECLS-K:2011 analyses were developed in several stages. The first stage of the weighting process assigned weights to the sampled primary sampling units that are equal to the inverse of the PSU probability of selection. The second stage of the weighting process assigned weights to the schools sampled within selected PSUs. The base weight for each sampled school is the PSU weight multiplied by the inverse of the probability of selecting the school from the PSU. The base weights of responding schools were adjusted to compensate for nonresponse among the set of eligible schools. These adjustments were made separately for public and private schools.

To compute the base weight for each student in the sample, the school nonresponse-adjusted weight for the school the student attended was multiplied by the within-school student weight. The within-school student weight was calculated separately for API

students and non-API students to account for the oversampling of API students. For API students, the within-school student weight is the total number of API kindergarten students in the school divided by the number of API kindergarten students sampled in the school. For non-API students, the within-school student weight is the total number of non-API kindergarten students in the school divided by the number of non-API kindergarten students sampled in the school. The student-level base weight was adjusted for nonresponse to produce each of the final student-level weights created for each round of the ECLS-K:2011 data collection. For each weight, a response status was defined based on the presence of data for particular components. The response status was used to adjust the base weight for nonresponse to arrive at the final full sample weight. Nonresponse classes were formed separately for each school type (public/Catholic/non-Catholic private). Within school type, analysis of child response propensity was conducted using child characteristics such as date of birth and race/ethnicity to form nonresponse classes. The child-level nonresponse adjustment was computed as the sum of the weights for all the eligible (responding and nonresponding) children in a nonresponse class divided by the sum of the weights of the eligible responding children in that nonresponse class.

A sample weight could be produced for use with data from every component of the study (e.g., data from the fall child assessment, from the fall parent interview, from the spring child assessment, from the spring parent interview, etc.) and for every combination of components for the study (e.g., data from the fall child assessment with data from the fall parent interview or data from the spring child assessment with data from the school administrator questionnaire). However, creating all possible weights for a study with as many components as the ECLS-K:2011 has would be impractical. In order to determine which weights would be most useful for researchers analyzing data, completion rates for each component at each round (e.g., response to the child assessment or the parent interview in fall kindergarten) were reviewed, and consideration was given to how analysts are likely to use the data (i.e., which weights will have greatest analytic utility).

Scaling. To maximize information on which each estimate of ability derived from the direct child assessments is based, the majority of the direct cognitive assessment scores computed for the study are based on IRT. IRT uses patterns of correct and incorrect answers to compute estimates on a scale that may be compared across different assessment forms within a given domain. IRT was employed in the ECLS-K:2011 to calculate ability estimates and then derive assessment scores from those ability estimates that can be compared both within a round and across rounds.

Imputation. Not all parent respondents provided complete education, occupation, and household income information. Therefore, it was necessary to impute missing values for these components of the socioeconomic status (SES) composite variable before computing the composite. The percentages of missing data for the education and occupation variables were small (for example, 2 to 3 percent in the base year). However, the household income variable generally has a higher rate of missing data (for example, 15.3 percent in the base year). Imputation was done separately for each component using the hot deck method. In this method, similar respondents and nonrespondents are grouped, or assigned to "imputation cells," and a respondent's value is randomly "donated" to a nonrespondent within the same cell. Cells were defined using demographic characteristics that are the best predictors of the component. Characteristics such as census region, school type (public/Catholic/non-Catholic religious private/other private), school locale (city/suburb/town/rural), household type (female single parent/male single parent/two parents present), parents' race/ethnicity, and parents' age were used to form the cells. Chi-square automatic interaction detector (CHAID) analyses were used to determine the predictors. Imputed as well as reported values were used to create imputation cells, but imputed values were not donated. No donor was used more than once.

For households with both parents present, each parent's variables were imputed separately. The order of imputation was parent 1's education, parent 2's education, parent 1's labor force status, parent 1's occupation, parent 2's labor force status, parent 2's occupation, and then household income.

Composites indicating the percent of students in the school who were approved for free school meals and the percent of students in a school who were approved for reduced-price school meals were derived from information collected from the school administrator during the spring data collection. Some school administrators did not complete the school administrator questionnaire, and among those who did, not all responded to all three questions needed to compute these composites related to approval for free or reduced-price meals. If school administrator data for public schools were missing, data were taken from the CCD (Common Core of Data). No external source

data were available for private schools. Hot-deck imputation was then conducted for cases from public schools for which data were not available in the CCD. Imputation cells were created using a measure of district poverty and whether the school received Title I funding. Within each imputation cell, the schools were sorted by longitude and latitude. Hand imputation was used for a small number of private schools.

Future Plans

The ECLS-K:2011 followed students through the spring of 2016, when most of them were expected to be in fifth grade.

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 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:2011 data files.

Design effects. An important analytic procedure is to compare the statistical efficiency of survey estimates from a complex sample survey such as the ECLS-K:2011 compares with estimates that would have been obtained had a simple random sample (SRS) of the same size. 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, defined as the ratio, for a given statistic, of the variance estimate under the actual sample design to the variance estimate that would be obtained with an SRS of the same sample size. In the ECLS-K:2011, a large number of data items were collected from children, parents, teachers, school administrators, and before- and after-school care providers. Each item has its own design effect that can be estimated from the survey data. For example, the median child-level design effect is 3.2 for fall kindergarten and 4.0 for spring 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 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 data were collected in falls of the base-year, first-grade, and second-grade, it is possible that the teachers did not have adequate time to observe the children, and thus some of their responses (especially at these rounds) 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, the ECLS-K:2011 used items that were previously used in the ECLS-K. Actual teachers were involved in the design of the

cognitive assessment battery and questionnaires for the ECLS-K. 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 the ECLS-K:2011. 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:2011 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 studies 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 were blind, those who were deaf, and those whose IEP clearly stated that they were not to be tested. Exclusion from the direct child assessment did not exclude children from other parts of the study (e.g., teacher questionnaire, parent interview).

Nonresponse error. A total of approximately 780 of the 1,320 originally sampled schools participated during the base year of the study. This translates into a weighted unit response rate (weighted by the base weight) of 63 percent for the base year. Due to the lower-than-expected cooperation rate for public schools in the fall of the base year, 85 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.

The weighted student unit response rates were 87 percent for the fall data collection and 85 percent for the spring data collection. The weighted student unit response rate for participation in the fall or spring data collections was 89 percent (i.e., a child assessment was completed at least once during kindergarten). The weighted student unit response rate for participation in both the fall and spring data collections was 76 percent (i.e., a child assessment was completed in both the fall and spring of kindergarten). The weighted parent unit response rates were 74 percent for the fall data collection and 67 percent for the spring data collection. The weighted parent unit response rate for participation in the fall or spring data collections was 80 percent (i.e., a parent interview was completed at least once during kindergarten). The weighted parent unit response rate for participation in both the fall and spring data collections was 55 percent (i.e., a parent interview was completed in both the fall and spring of kindergarten). The overall base-year response rate for students (with a complete assessment in either fall or spring) was 56 percent (63 percent of schools x 89 percent of sampled children) and the overall base-year response rate for the parent interview (i.e., a complete parent interview in either fall or spring) was 50 percent (63 percent of schools x 80 percent of parents of sampled children).

For the first-grade follow-up, the weighted child assessment unit response rates were 89 percent for the fall and 88 percent for the spring. The weighted parent unit response rates were 87 percent for the fall first-grade data collection, and 76 percent for the spring. Overall response rates for the child assessment, which take into account the base-year school-level response rate (63 percent), were 56 percent for the fall and 55 percent for the spring. Overall parent interview response rates, which also take into account school-level response, were 54 percent for the fall first-grade data collection and 48 percent for the spring first-grade data collection.

For the second-grade follow-up, the weighted child assessment unit response rates were 84 percent for the fall and 83 percent for the spring. The overall response rates for the child assessment were 53 percent for the fall collection and 52 percent for the spring. No parent interviews were conducted during this round.

A nonresponse bias analysis was conducted to determine if substantial bias was introduced as a result of nonresponse. To examine the effect of school nonresponse, estimates from the ECLS-K:2011 schools were compared to those produced using frame data (i.e., data from the Common Core of Data and the Private School Universe Survey). The differences in the two sets of estimates are very small, suggesting there is not significant nonresponse bias present in the data. To examine the effect of nonresponse for data collected through instruments that have a response rate

Table ECLS-K:2011-1. Weighted unit response rates, by instrument: School years 2010–11 through 2012–13

Instrument	Kindergarten (2010–11)	First Grade (2011–12)	Second Grade (2012–13)
School	63	—	—
Child Assessment			
Fall	87	89	84
Spring	85	88	83
Overall Child Assessment[1]			
Fall	55	56	53
Spring	53	55	52
Parent Interview			
Fall	74	87	—
Spring	67	76	—
Overall Parent Interview[1]			
Fall	47	54	—
Spring	42	48	—

— Not available.

[1] The overall response rates take into account the base-year school-level response rate (63 percent).

NOTE: The weighted unit response rates for the child assessment and parent interview were calculated using the student base weight, which is the product of the school base weight and the within-school student weight.

SOURCE: Mulligan, G.M., McCarroll, J.C., Flanagan, K.D., and Potter, D. (2014). *Findings From the First-Grade Rounds of the Early Childhood Longitudinal Study, Kindergarten Class of 2010–11 (ECLS-K:2011)* (NCES 2015-109). National Center for Education Statistics, Institute of Education Sciences, U.S. Department of Education. Washington, DC. Mulligan, G.M., McCarroll, J.C., Flanagan, K.D., and Potter, D. (2015). *Findings From the Second-Grade Rounds of the Early Childhood Longitudinal Study, Kindergarten Class of 2010–11 (ECLS-K:2011)* (NCES 2015-077). U.S. Department of Education, Washington, DC: National Center for Education Statistics.

lower than 85 percent, in this case the kindergarten parent interviews and the second-grade child assessment, estimates produced using weights that include adjustments for nonresponse were compared to estimates produced using weights without nonresponse adjustments. Additionally, for the parent interview data, estimates from the ECLS-K:2011 were compared to those from other data sources (for example, the National Household Education Surveys Program). The results of these nonresponse bias analyses also suggest that there is not a substantial bias due to nonresponse after adjusting for that nonresponse.

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.

Mulligan, G.M., Hastedt, S., and McCarroll, J.C. (2012). *First-Time Kindergartners in 2010–11: First Findings From the Kindergarten Rounds of the Early Childhood Longitudinal Study, Kindergarten Class of 2010–11 (ECLS-K:2011)* (NCES 2012-049). U.S. Department of Education. Washington, DC: National Center for Education Statistics.

http://nces.ed.gov/pubsearch/pubsinfo.asp?pubid=2012049.

Mulligan, G.M., McCarroll, J.C., Flanagan, K.D., and Potter, D. (2014). *Findings From the First-Grade Rounds of the Early Childhood Longitudinal Study, Kindergarten Class of 2010–11 (ECLS-K:2011)* (NCES 2015-109). National Center for Education Statistics, Institute of Education Sciences, U.S. Department of Education. Washington, DC. http://nces.ed.gov/pubsearch/pubsinfo.asp?pubid=2015109.

Mulligan, G.M., McCarroll, J.C., Flanagan, K.D., and Potter, D. (2015). *Findings From the Second-Grade Rounds of the Early Childhood Longitudinal Study, Kindergarten Class of 2010–11 (ECLS-K: 2011)* (NCES 2015-077). U.S. Department of Education, Washington, DC: National Center for Education Statistics. https://nces.ed.gov/pubsearch/pubsinfo.asp?pubid=2015077.

NIH. (n.d.). *Executive Function*. http://www.nihtoolbox.org/WhatAndWhy/Cognition/ExecutiveFunction/Pages/default.aspx.

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