Early Childhood Longitudinal Study (ECLS)

Website: http://nces.ed.gov/ecls/
Updated: December 2012

1. OVERVIEW

T he 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 in relation to their family, early childhood, and school environments.

Purpose

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

The Early Childhood Longitudinal Study, Kindergarten Class of 2010–11 (ECLS-K:2011). 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 also is collected on children's home environment, home educational activities, school environment, classroom environment, classroom curriculum, teacher background, and before- and after-school care.

Direct child assessments. Children are administered assessments containing age- and grade-appropriate items measuring important cognitive skills and knowledge in each round of data collection. The untimed assessments are administered directly to the sampled children, one-on-one, by a trained assessor. The kindergarten child assessment measured reading (fall and spring), mathematics (fall and spring), and science (spring) knowledge and skills, as well as executive function (fall and spring)—*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.).

EARLY CHILDHOOD LONGITUDINAL SAMPLE SURVEYS: BIRTH COHORT AND KINDERGARTEN COHORTS

ECLS collects data from:

➢ Children

➢ Parents/guardians

➢ Child care providers and preschool teachers

➢ Teachers

➢ School administrators

Also in the kindergarten year, 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 the fall and spring. The kindergarten direct cognitive assessment included two-stage assessments for reading and mathematics. 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 kindergarten science assessment included 20 items that all children received; a two-stage assessment was not used in this domain during kindergarten.

Parent interviews. Information is collected from parents/guardians in each data collection round using computer-assisted interviews (CAIs). The parent interview asks about family structure, family literacy practices, parental involvement in school, 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 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. Information is collected from classroom teachers via self-administered questionnaires during the fall and spring data collections.

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 collection.

School administrator questionnaire. 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 collection.

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

Direct child assessments. The direct child assessments covered several cognitive domains (reading and mathematics in kindergarten through eighth grade; general knowledge, consisting of science and social studies questions, in kindergarten and first grade; and science in third, fifth, and eighth grades); a psychomotor assessment (fall kindergarten only), including fine and gross motor skills; and height and weight measurements. Beginning with the third-grade data collection, children reported on their own perceptions of their abilities and achievement, as well as their interest in and enjoyment of reading, math, and other school subjects. An English language proficiency screener, the Oral Language Development Scale (OLDS), was administered to children if school records indicated that the child's home language was not English. The child had to demonstrate a certain level of English proficiency on the OLDS to be administered the ECLS-K cognitive assessment in English. If a child spoke Spanish at home and did not have the English skills required for the ECLS-K battery, the child was administered a Spanish version of the OLDS, and the mathematics and psychomotor assessments were administered in Spanish. The assessment for each cognitive domain included a routing test (to determine a child's approximate skill level) and second-stage tests that were tailored to different skill levels. In the eighth-grade data collection, children completed a student questionnaire after completing the routing test. The student questionnaire covered many topics about the

child's school experiences, school-sponsored and out-of-school activities, self-perceptions of social and academic competence and interests, weight and exercise, and diet.

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

Classroom teacher 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 the core academic subjects (i.e., reading/language arts, mathematics, and science) provided the data pertinent to each child's classroom environment and instruction for the academic subject about which they were most knowledgeable. Teachers also provided information about their professional background.

Special education teacher 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. 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.

The Early Childhood Longitudinal Study, Birth Cohort (ECLS-B). The ECLS-B, which began in October 2001, was designed to study children's early learning and development from birth through the fall of the kindergarten year. Over the course of the study, data were collected from multiple sources, including birth certificates, children, parents/guardians, nonparental care providers, and teachers. Data from the Common Core of Data (CCD) and the Private School Survey (PSS) were linked to provide administrative information for the schools where the children attended kindergarten. These components are described below.

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

Parent/guardian interviews. A parent/guardian interview was conducted in the children's home at each

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

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

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

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

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

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

Nonparental care and education providers. Individuals and organizations that provided regular care for a child were interviewed with the permission of the child's parents. They were asked about their backgrounds,

teaching practices and experience, the children in their care, and children's learning environments. This information was collected from the 2-year data collection on. In the kindergarten 2006 and 2007 collections, a wrap-around care provider interview was used for those children who were in kindergarten and had a before- or after-school care arrangement.

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

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

Periodicity

The ECLS-K:2011 data were collected in the fall and the spring of kindergarten (2010–11), the fall and the spring of first grade (2011–12), and the fall of second grade (2012) (data were collected from a 30 percent subsample in the fall of first grade and the fall of second grade).The study will continue to collect data in the springs of second grade (2013), third grade (2014), fourth grade (2015), and fifth grade (2016).

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).

The ECLS-B collected data when the children were about 9 months old (2001–02), about 2 years old (2003), about 4 years old (the preschool collection) (2005), and in the fall of kindergarten (2006 and 2007). Note that because of age requirements for school entry, children sampled in the ECLS-B entered kindergarten in two different school years. All study children were included in the kindergarten 2006

collection, regardless of their enrollment status or grade in school. The kindergarten 2007 collection included just a portion of the total ECLS-B sample: children who were not yet in kindergarten in the 2006 collection, children who were in kindergarten in the 2006 collection and were repeating kindergarten in the 2007 collection, and twins of children in these groups. The ECLS-B study ended with the kindergarten 2007 round of collection.

2. USES OF DATA

The ECLS-K and ECLS-K:2011 provide information critical to informing policies that can respond sensitively and creatively to diverse learning environments. In addition, the ECLS-K and ECLS-K:2011 enable 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 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.

The ECLS-B provides descriptive data on children's health status at birth; children's experiences in the home, nonparental care, and school; and children's development and growth through kindergarten. The data collected in the ECLS-B can be used to explore the relationships between children's developmental outcomes and characteristics of their family, health care, nonparental care, school, and community. The longitudinal nature of the study enables researchers to study children's physical, social, and emotional growth and to relate trajectories of growth and change to variations in children's experience.

3. KEY CONCEPTS

Item Response Theory (IRT) scale scores. The ECLS direct cognitive assessments employ a two-stage design. As such, within any given domain, children receive a routing set of items (stage 1) and then based on their performance proceed to a certain difficulty level (stage 2). Because not all children receive all items, the assessment scores in the ECLS studies are modeled using Item Response Theory (IRT). Based on children's performance on the items they received, an ability estimate (theta) is derived for each domain. The theta is used to derive other scores, such as scale scores, T-scores, and proficiency probability scores.

The IRT scale scores represent estimates of the number of items children would have answered correctly if they had received all of the scored questions in a given content domain. They are reported in the ECLS-K:2011, ECLS-K and ECLS-B. They are useful in identifying cross-sectional differences among subgroups in overall achievement levels and provide a summary measure of achievement useful for 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, available for the ECLS-K and ECLS-B, are also IRT based. They provide norm-referenced measurements of achievement; that is, estimates of achievement level relative to the population as a whole. A high mean T-score for a particular subgroup indicates that the group's performance is high in comparison to that of other groups. A change in mean T-scores over time reflects a change in the group's status with respect to that of other groups.

Proficiency probability scores. These scores, available for the ECLS-K, are IRT-based and 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. In the ECLS, Office of Management and Budget guidelines 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.

Socioeconomic status (SES). Each ECLS 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 was used. In the ECLS-K and ECLS-K:2011, each parent's occupation was scored using the average of the 1989 General Social Survey (GSS) prestige scores for the 1980 census occupational category codes that correspond to the ECLS-K occupation code. In the ECLS-B, each parent's occupation was scored using the average of the 1989 GSS prestige scores for the 2000 census occupational category codes covered by the ECLS-B occupation.

4. SURVEY DESIGN

Target Population
Representative samples of kindergartners and babies are studied longitudinally for 6 or more years. Kindergarten children enrolled during the 1998–99 school year are the baseline for the ECLS-K cohort; babies born during 2001 are the baseline for the ECLS-B cohort.[1] Kindergarten children enrolled in the 2010–11 school year are the baseline for the ECLS-K:2011 cohort.

Sample Design
The sample design is discussed separately for ECLS-K:2011, ECLS-K, and ECLS-B.

Kindergarten cohort of 2010-11 (ECLS-K:2011). 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

---

[1] The ECLS-B target population excludes children who were born to mothers younger than age 15 and children who died or were adopted prior to the 9-month home visit. Over time, the target population excludes children who died or moved abroad permanently.

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 sources for the 2010 NAEP school frames were the most recent Common Core of Data (2006–07 CCD) and Private School Survey (2007–08 PSS) available at the time the NAEP frame was developed. 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 spring of 2010, after the original sample of schools was selected, procedures were used to update the school frame and select a supplemental sample of newly opened schools and existing schools that added kindergarten programs that were not included in the original frames.

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 970 schools participated.

Fall first-grade collection. The fall first grade collection took place in fall 2012. At the time of publication, documentation for this data collection was not yet available; however, the design of the collection mirrored the fall first grade collection of ECLS-K (see below for its details).

Kindergarten cohort of 1998-99 (ECLS-K). The ECLS-K followed a nationally representative cohort of children from kindergarten through the spring of 2007, when most of the children were in eighth grade.

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 was designed to support separate estimates of public and private school kindergartners; Black, Hispanic, White, and Asian/Pacific Islander children; and children grouped by SES.

The sample design for the ECLS-K was a dual-frame, multi-stage sample. First, 100 PSUs were selected from an initial frame of 1,400 PSUs, representing counties or groups of contiguous counties. The 24 PSUs with the largest measures of size (where the measure of size 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 spring 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 fall 1998, approximately 23 kindergarten children were selected, on average, from each of the sampled schools. Asian/Pacific Islander children and private schools were oversampled.

For the base year of the ECLS-K, 22,000 children were sampled and eligible.

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

Spring first-grade 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, 18,080 children were eligible (14,250 public school and 3,840

private school children). Student freshening brought 170 first-graders into the ECLS-K sample.

Spring third-grade 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, 16,670 children were eligible (13,170 public schools and 3,500 private school children).

Spring fifth-grade collection. In fifth grade, four groups of children were not followed. 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, 12,030 children were eligible (9,570 in public schools and 2,460 in private schools).

A new feature of the fifth-grade sample was the subsampling of eligible children for the administration of mathematics and science questionnaires. While a child-level reading teacher questionnaire was fielded for all children retained for 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).

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

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

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

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

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

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

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

For the 9-month collection, approximately 14,200 children were eligible,[2] and approximately 10,700 participated.

Two-year collection. Only cases with a completed 9-month parent interview (about 10,700) were eligible

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

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

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

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

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

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

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

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

Assessment Design

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

Kindergarten cohort of 2010-11 (ECLS-K:2011). 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 kindergarten 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 kindergartners was included in the kindergarten assessments. 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.

> ➤ *Executive function:* Measures of executive function were included in the kindergarten direct child assessment battery to assess children's cognitive flexibility and working memory.

> The Dimensional Change Card Sort (Zelazo, 2006) is used to collect information on children's cognitive flexibility. In this task, 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.

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.

Kindergarten cohort of 1998-99 (ECLS-K). The design of the ECLS-K assessment was guided by the domain assessment framework proposed by the National Education Goals Panel's Resource Group on School Readiness. A critical component of the ECLS-K is the assessment of children 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.

Birth cohort (ECLS-B). The ECLS-B direct child assessment in the 9-month and 2-year data collections relied on instruments considered "gold standards" in the field. However, adaptations were necessary to take these instruments from a laboratory or clinical setting to a home setting. The preschool and kindergarten direct cognitive assessments, which primarily focused on early academic skills and knowledge in reading and mathematics, drew items from the ECLS-K cognitive assessments, as well as well-established standardized assessments such as the Peabody Picture Vocabulary Test, Third Edition (PPVT-III). All of the ECLS-B child assessments were designed for ease of and flexibility in administration while at the same time being psychometrically and substantively sound.

➢ *Cognitive development and fine and gross motor skills:* The BSID-II is considered the gold standard for assessing early childhood development from ages 1 to 42 months. In the 9-month and 2-year collections, children's cognitive development, as well as their receptive and expressive language skills, were assessed using an adaptation of the mental scale of the BSID-II. Children were observed doing specific things such as retrieving hidden toys and looking at picture books, and their production of vowel-consonant combinations was noted. Fine and gross motor skills were assessed using an adaptation of the motor scale of the BSID-II. Children were observed doing specific things such as grasping small objects, crawling, and walking. The study had intended

to field the entire BSID-II assessment, as it was originally expected to take about 20 minutes to complete. However, a field test for the 9-month ECLS-B data collection revealed that it actually required an average of 40 minutes to complete in the home. As a result, modifications were made to the original BSID-II. The ECLS-B data collection contractor worked with experts to identify a reduced item set, referred to as the Bayley Short Form-Research Edition (BSF-R), that could be administered in less time and could produce reliable, valid scores equivalent to the full set of Bayley items. The BSF-R took approximately 25 minutes to administer. Because the BSF-R was not appropriate for children older than 42 months of age, a new direct child cognitive assessment was developed for use in the preschool and kindergarten collections. These assessments were patterned after the ECLS-K assessments and incorporated items from the ECLS-K, as well as other published assessments, such as the *pre*Las 2000, Test of Early Mathematics Ability, Third Edition (TEMA 3), and the Peabody Picture Vocabulary Test, Third Edition (PPVT-III). The cognitive domains covered in the preschool-kindergarten assessments were early reading and mathematics skills. The preschool collection also included a measure of children's color knowledge, which involved asking the children to name the colors of each bear presented to them in picture format. Children's fine and gross motor skills at preschool and kindergarten were measured using the Bruininks-Oseretsky Test of Motor Proficiency and Movement Assessment Battery for Children. To assess fine motor skills, children were asked to copy a series of forms (e.g., circle, triangle, square) that were first drawn by an assessor and to build a structure with blocks that was first demonstrated by the assessor. To assess gross motor skills, children were asked to hop, skip, jump forwards, walk backwards, balance on one foot, and catch a beanbag tossed to them by the assessor.

➢ *Socioemotional development:* The Two Bags Task was used in the 2-year and preschool data collections to measure socioemotional development. The Two Bags Task is a simplified version of the Three Bags Task that was used successfully in such large-scale studies as the Early Head Start Research and Evaluation Project and is intended to capture children's socioemotional functioning. It is a semistructured activity completed by the parent

and child in interaction. During this 10-minute task, the parent-child dyad was asked to play with two different sets of toys, each placed within a separate numbered bag. In the 2-year collection, bag number 1 contained a children's picture book and bag number 2 contained a set of dishes. In the preschool collection, bag number 1 also contained a children's picture book but bag number 2 contained PlayDoh. The rating scales derived from the Two Bags Task provide information on parents' behaviors during the interaction (parental sensitivity, intrusiveness, simulation of cognitive development, positive regard, negative regard, and detachment) and children's behaviors during the interaction (child engagement of parent, sustained attention, and negativity toward parent).

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

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

Data Collection and Processing
The ECLS-K: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 conducted the kindergarten rounds of data collection.

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

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

Reference dates. For the ECLS-K:2011, baseline data were collected from September through December 2010 for the fall and from late March through June 2011 for the spring. For the ECLS-K, baseline data were collected from September through December 1998. For the ECLS-B, baseline data were collected from October 2001 through December 2002.

Data collection. The ECLS-K:2011, ECLS-K, and the ECLS-B are discussed separately.

Kindergarten cohort of 2010-11 (ECLS-K:2011) Fall and spring data collections included direct child assessments, parent interviews, teacher and school administrator questionnaires, and before- and after-school care provider questionnaires. 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.

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.

Two data collections were conducted in the 2010-11 school year, one in the fall and one in the spring. The fall kindergarten direct child assessments were conducted from September through December 2010. About two-thirds of the assessments were completed by the end of October, with 25 percent completed in November and 10 percent completed in December. Fall kindergarten parent interviews were conducted from September 2010 to January 2011. The spring kindergarten direct child assessments were conducted from late March through June 2011. Approximately 60 percent of the assessments were completed by the end of April, with 35 percent completed in May and 5 percent completed in June. Spring kindergarten parent interviews were conducted from March to July 2011.

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 assessment was administered directly to the sampled children on an individual basis by trained and certified child assessors and took about 60 minutes per child. Responses were entered by the assessors into a computer-assisted interviewing (CAI) program. The kindergarten direct cognitive assessment included two-stage assessments for reading and mathematics and a one-stage, 20-item science assessment. 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 (preLAS 2000). The "Simon Says" task required children to follow simple, direct instructions given by the assessor in English. The "Art Show" task was a picture vocabulary assessment that tested children's expressive vocabulary. All children, regardless of home language, were administered the language screener as the first component of the direct cognitive assessment. For children whose home language was English, the screener primarily served as a warm-up or practice for the rest of the assessment

since the items were of low difficulty. While the screener also served as a warm-up for children whose home language was one other than English, it also determined whether the children understood English well enough to receive the full direct child assessment in English. All children also received the first 18 items of the reading assessment in English, regardless of their home language or performance on the preLAS tasks. These items, plus two items from the preLAS "Art Show" 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 spring kindergarten data collection round, interviewers attempted to complete the parent interview with the same respondent who answered the parent interview in the fall kindergarten round, though another parent or guardian in the household who knew about the child's care, education, and health was selected if the fall respondent was not available.

The parent interview was 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 interview was 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.

A continuous quality assurance process was applied to all data collection activities. Data collection quality control efforts began with the development and testing of the CATI and CAPI applications and the 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. During the assessor/interviewer training for the kindergarten year, field staff practiced conducting the parent interview in pairs and practiced the direct child assessment with kindergarten children brought to the training site for this purpose. In later data collection periods, experienced staff used a home study training package while new staff was trained in classroom sessions. After data collection began, field supervisors observed each assessor conducting child assessments and made telephone calls to parents to validate the interview. Field managers also made telephone calls to the schools to collect information on the school activities for validation purposes.

Kindergarten cohort of 1998-99 (ECLS-K). The data collection schedule for the ECLS-K was based on a desire to capture information about children as critical events and transitions were occurring rather than measuring these events retrospectively. A large-scale field test of the kindergarten and first-grade assessment instruments and questionnaires was conducted in 1995–96. This field test was used primarily to collect psychometric data on the ECLS-K assessment item pool and to evaluate questions in the different survey instruments. Data from this field test were used to develop the routing and second-stage tests for the ECLS-K kindergarten and first-grade direct cognitive assessment battery and to finalize the parent, teacher, and school administrator instruments. A pilot test of the 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 encounter with school. Data were collected from the child, the child's parents/guardians, and teachers in both fall and spring. Data were collected from school administrators in the spring. For the fall 1998 and spring 1999 collections, all child assessment measures were obtained through untimed 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 assessment 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 collected 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 fall 1999, when most of the kindergarten cohort had moved on to first grade, data were collected from a 30 percent subsample of the cohort. The direct child assessment was administered during a 12-week field period (September–November 1999). The parent interview was administered between early September and mid-November 1999; it averaged 35 minutes and was conducted primarily by telephone.

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

(March–June 2000 for first grade and March–June 2002 for third grade, but field staff prompted by telephone for the return of these materials through October 2000 and October 2002, respectively. For fifth grade, data collection occurred between February and June 2004. For eighth grade, data collection occurred between March and June 2007).

The same continuous quality assurance process that was applied to all data collection activities in ECLS-K:2011 (described above) was also applied to all data collection activities of ECLS-K. Specifically, extensive testing of the CATI and CAPI applications and the data collection contractor's Field Management System was conducted. As well, field procedures that maximized cooperation and thereby reduced the potential for nonresponse bias were developed. During the assessor/interviewer training for the kindergarten year, field staff practiced conducting the parent interview in pairs and practiced the direct child assessment with kindergarten children brought to the training site for this purpose. In later data collection periods, experienced staff used a home study training package while new staff was trained in classroom sessions. After data collection began, field supervisors observed each assessor conducting child assessments and made telephone calls to parents to validate the interview. Field managers also made telephone calls to the schools to collect information on the school activities for validation purposes.

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

A second field test of the ECLS-B instruments and procedures began in September 2000. A field test sample was drawn consisting of 1,060 children born between January and April 2000. Home visits were conducted when the children were 9 months old and again when they were 18 months old. Results from this field test indicated that the changes to the design that resulted from the first field test were successful.

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

The ECLS-B 9-month data collection began in October 2001 and continued through December 2002. The 2-year data collection began in January 2003 and continued through April 2004. While the 9-month and 2-year data collection schedules were designed to collect information on children as close as possible to the date on which they turned the age of interest for the collection (i.e., 9 months and 2 years), the collection schedules for the preschool and kindergarten rounds were changed to correspond with an academic calendar. Thus, the preschool wave of data collection began in late August 2005 and ended in mid-July 2006.

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

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

Data were collected from child care providers using CATI. A subset of child care providers was sampled for on-site observations in the 2-year and preschool collections; observers recorded data in booklets. Child care center directors were asked to complete a self-administered paper questionnaire during the observation visit.

Editing. Within the CATI/CAPI instruments, the ECLSK:2011, ECLS-K and ECLS-B respondent answers were subjected to both "hard" and "soft" range edits during the interviewing process. Responses outside the soft range of reasonably expected values were confirmed with the respondent and entered a second time. For hard-range items, out-of-range values (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, videotaped instruments, and observational instruments.

Estimation Methods

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. A hot-deck imputation methodology was used to impute missing values for all components of SES in the ECLS-K:2011, ECLS-K, and ECLS-B. Imputation also was implemented for child assessment proficiency-level variables in the ECLS-K and free/reduced-price school meals data in the ECLS-K:2011 and ECLS-K.

Weighting. Weighting in the ECLS-K:2011, ECLS-K, and ECLS-B are discussed separately.

Kindergarten cohort of 2010-11 (ECLS-K:2011). 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 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 11 final student-level weights created for the base year of the ECLS-K:2011. 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 from the base year, completion rates for each fall kindergarten and spring kindergarten component (e.g., response to the child assessment or the parent interview) were reviewed, and consideration was given to how analysts are likely to use the data (i.e., which weights will have greatest analytic utility). This process yielded the creation of 11 student-level analytic weights to be used with data from the base year of the ECLS-K:2011.

Kindergarten cohort of 1998-99 (ECLS-K). Several sets of weights were computed for each of the seven rounds of data collection (fall kindergarten, spring kindergarten, fall first grade, spring first grade, spring third grade, spring fifth grade, and spring eighth grade). These weights include cross-sectional weights for analyses of data from one time point, as well as longitudinal weights for analyses of data from multiple rounds of the study. Unlike surveys that have only one type of survey instrument 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.

Birth cohort (ECLS-B). Several sets of weights were computed for each round of data collection. The ECLS-B weights were developed in three steps: First, base weights were calculated using the overall selection probabilities; next, weights were adjusted for survey nonresponse; finally, raking was used to adjust for undercoverage and to improve the precision of survey estimates.

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

The base weights were adjusted for survey nonresponse. A selected set of variables related to child and family characteristics was used to construct nonresponse adjustment cells for each set of weights. Respondents and nonrespondents were compared on the characteristics selected based on analyses using segmentation modeling via CHAID. In the first round of data collection, data from the birth certificate were used to compare respondents and nonrespondents, because these data were available for all sampled cases regardless of participation status. In later collections, respondents and nonrespondents were compared on both birth certificate data and data collected in prior rounds. A nonresponse adjustment factor was calculated for each cell as the ratio of the sum of weights for eligible cases in the cell to the sum of weights for eligible and responding cases in the cell. Finally, the nonresponse-adjusted weights were raked to 11 dimensions to ensure that sums of weights matched known population totals, thus correcting for survey undercoverage. The 11 dimensions were selected because of their substantive interest as well as their relationship to response propensity, as indicated by the CHAID modeling and also some preliminary logistic regression analyses.

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

Scaling. IRT was employed in the ECLS-K:2011, ECLS-K, and ECLS-B to calculate assessment scores 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. Kindergarten cohort of 2010-11 (ECLS-K:2011). 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 in the base year were small (2 to 3 percent). However, the household income variable had a higher rate of missing data (15.3 percent). 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 2009-10 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.

Kindergarten cohort of 1998-99 (ECLS-K). In the ECLS-K, 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); however, the household income variable had a higher rate of missing data (28.2 percent in the base year and 11 to 33 percent in the spring of first grade, depending on whether a detailed income range or the exact household income was requested; in the spring of third grade, 11.1 percent of cases had missing data for the detailed income range; this percentage was 8.1 percent of cases in the spring of fifth grade and 7.0 percent of cases in the spring of eighth grade). A standard (random selection within class) hot-deck imputation methodology was used to impute for missing values of all 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 variables were imputed separately. If this was not the case, an "unknown" or missing category was created as an additional level for the CHAID analysis. As a rule, no imputed value was used as a donor. In addition, the same donor was not used more than two times. The order of the imputation for all the variables was from the lowest percentage missing to the highest.

Imputation for occupation involved two steps. First, the labor force status of the parent was imputed, whether the parent was employed or not. Then the parent's occupation was imputed only for those parents whose status was identified as employed, either through the parent interview or the first imputation step. The variable for income was imputed last; if a respondent provided partial information about income, this information was used in the imputation process.

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

Birth cohort (ECLS-B). Variables used to derive the SES composite variable were imputed using a hot-deck methodology. These variables include mother's and father's education, mother's and father's occupation, and household income. Imputation cells were defined by respondent characteristics that were the best predictors of the variables to be imputed, as determined using a CHAID analysis. Hot-deck imputation was done in a sequential order, separately, by type of household (female single parent, male single parent, and both parents present). Missing data were first filled with data obtained in a prior round, if available. For households with both parents present, the mother's and father's variables were imputed separately. Imputed as well as reported values were used to define imputation cells; missing values for donor characteristics were treated as a separate category. No imputed value was used as a donor. No donor was used more than once. The order of hot-deck imputation for all variables was from the lowest percentage missing to the highest.

Future Plans
The ECLS-K:2011 will continue to follow students through in the spring of 2016, when most of them are expected to be in fifth grade. All children will be followed, regardless of their actual grade level in any given round. Because the study is designed to allow for comparisons between the 2010-11 cohort and the cohort of kindergartners included in the ECLS-K, the ECLS-K:2011 is very similar to the ECLS-K and includes most of the same components.

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, ECLS-K and ECLS-B data files.

Design effects. Kindergarten cohort of 2010-11 (ECLS-K:2011). An important analytic procedure is to compare the statistical efficiency of survey estimates from a complex sample survey such as the ECLS-K:2011 with what would have been obtained in a hypothetical and usually impractical simple random sample (SRS) of the same size. In a stratified clustered design, 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. The median child-level design effect is 3.2 for fall kindergarten and 4.0 for spring kindergarten.

Kindergarten cohort of 1998-99 (ECLS-K). A large number of data items were collected from students, parents, teachers, and schools. Each item has its own design effect that can be estimated from the survey data. The median child-level design effect is 4.7 for fall kindergarten and 4.1 for spring kindergarten. The median child-level design effect for spring third grade, spring fifth grade, and spring eighth grade is 3.3, 4.0, and 3.1, respectively.

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

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

would be most effective for producing teacher-level statistics.

The median school-level design effect is 1.6. Design effects were not computed for items from the teacher and school administrator questionnaires in the 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.

Birth cohort (ECLS-B) As noted above, several sets of weights were developed for use with different combinations of survey components that are of analytic interest. Design effects were computed for different survey estimates produced using these different weights. Using the weights designed for use with data from the parent interviews, the median design effect is 2.1 for the 9-month data collection, 2.4 for the 2-year collection, 2.1 for the preschool collection, 2.0 for the kindergarten 2006 collection, and 2.2 for the kindergarten 2007 collection. The median design effects for other weights across all components and all rounds of collections range from a low of 1.2 for the 2-year weight associated with response to the child care observation (W22P0) and a high of 4.2 for the 9-month weight associated with response to the 9-month child assessment (W1C0).

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

Nonsampling Error
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:2011 and ECLS-K and the kindergarten surveys in the ECLS-B were first conducted in the fall, it is possible that the teachers did not have adequate time to observe the children, and thus some of their responses may be influenced by their expectations based on 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 the ECLS-K:2011, ECLS-K, and ECLS-B. It is not possible to state precisely how such bias may affect the results. NCES has tried to minimize some of these biases by conducting one-on-one, untimed assessments, and by asking some of the same questions about the sampled child of both teachers and parents.

Coverage error. Undercoverage occurs when the sampling frame from which a sample is selected does not fully reflect the target population of inference. The potential for coverage error in the ECLS-K and 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 study were those who were blind, those who were deaf, those whose IEP clearly stated that they were not to be tested, and, for the ECLS-K, 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).

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

Nonresponse error. *Kindergarten cohort of 2010-11 (ECLS-K:2011).* 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).

A nonresponse bias analysis was conducted to determine if substantial bias was introduced into the data from the kindergarten collections as a result of nonresponse. Three methods were used to examine the potential for nonresponse bias: (1) a comparison of estimates from the ECLS-K:2011 to those produced using frame data (i.e., data from the CCD and the PSS); (2) a comparison of estimates from the ECLS-K:2011 to other data sources (for example, the National Household Education Surveys Program); and (3) a comparison of estimates produced using weights that include adjustments for nonresponse to estimates produced using weights without nonresponse adjustments. Findings from these analyses suggest that there is not a substantial bias in the kindergarten year due to nonresponse after adjusting for that nonresponse.

Kindergarten cohort of 1998-99 (ECLS-K). Overall, 880 of the 1,280 eligible schools (69.4 percent weighted) agreed to participate in the fall kindergarten study. Due to the lower-than-expected cooperation rate for public schools in the fall of the base year, 74 additional public schools were included in the sample as substitutes for schools that did not participate. These schools were included in order to meet the target sample sizes for students. Substitute schools are not included in the school response rate calculations.

A 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 response rate (weighted by the base weight) of 74 percent for the base year of the study. The weighted child base-year survey response rate was 92 percent (i.e., 92 percent of the children were assessed at least once during kindergarten). The weighted parent base-year unit response rate was 89 percent (i.e., a parent interview was completed at least once during kindergarten). Thus, the overall base-year response rate for children was 68 percent (74 percent of schools x 92 percent of sampled children) and the base-year overall response rate for the parent interview was 66 percent (74 percent of schools x 89 percent of parents of sampled children). About 76 percent of children and 72 percent of parents eligible for the eighth grade data collection (spring 2007) participated.

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.

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

Response rates are also calculated for the other components of the ECLS-B: the child assessments; the resident and nonresident father questionnaires; the care provider interview; the child care observation; the teacher questionnaire; and the school data. In all rounds of the ECLS-B, response rates for these other components are conditioned on the completion of the parent interview. Only cases with completed parent interviews were assigned weights for the other components of the study.

In the 9-month data collection, the weighted completion rate for the parent CAPI instrument was 74.1 percent (table ECLS-1). The weighted completion rates for the child assessment, resident father questionnaires, and nonresident father questionnaires were 95.6, 76.1, and 50.0 percent, respectively.

In the 2-year data collection, the weighted completion rate for the parent CAPI instrument was 93.1 percent. The weighted completion rates for the child assessment, resident father questionnaires, nonresident

father questionnaires, child care provider interview, and child care observation (CCO) component were 94.2, 77.7, 39.8, 70.0, and 51.3 percent, respectively. The longitudinal weighted response rates for the parent CAPI instrument, child assessment, and all father questionnaires were 69.0, 65.0, and 48.7 percent, respectively.

In the preschool data collection, the weighted completion rate for the parent CAPI instrument was 91.3 percent. The weighted completion rates for the child assessment, resident father questionnaires, child

care provider interview, and CCO component were 98.3, 87.7, 87.4, and 56.8 percent, respectively. The longitudinal weighted response rates for the parent instrument, child assessment, resident father questionnaires, child care provider interview, and CCO component were 63.1, 62.0, 55.3, 55.1, and 35.8, respectively.

Table ECLS-1.   Weighted unit response rates for all children in the ECLS-B, by survey and component: Various years 2001–2007

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

† Not applicable.
SOURCE: Denton Flanagan, K., and McPhee, C. (2009). *The Children Born in 2001 at Kindergarten Entry: First Findings From the Kindergarten Data Collections of the Early Childhood Longitudinal Study, Birth Cohort (ECLS-B)* (NCES 2010-05). National Center for Education Statistics, Institute of Education Sciences, U.S. Department of Education. Washington, DC. Jacobson Chernoff, J., Flanagan, K. D., McPhee, C., and Park, J. (2007). *Preschool: First Findings From the Preschool Follow-up of the Early Childhood Longitudinal Study, Birth Cohort (ECLS-B)* (NCES 2008-025). National Center for Education Statistics, Institute of Education Sciences, U.S. Department of Education. Washington, DC. Nord, C., Edwards, B., Andreassen, C., Green, J. L., and Wallner-Allen, K. (2006). *Early Childhood Longitudinal Study, Birth Cohort (ECLS-B), User's Manual for the ECLS-B Longitudinal 9-Month–2-Year Data File and Electronic Codebook (NCES 2006–046)*. National Center for Education Statistics, Institute of Education Sciences, U.S. Department of Education. Washington, DC. Nord, C., Edwards, B., Hilpert, R., Branden, L., Andreassen, C., Elmore, A., Sesay, D., Fletcher, P., Green, J.L., Saunders, R., Dulaney, R., Reaney, L., and Flanagan, K.D. (2004). *User's Manual for the ECLS-B Nine-Month Restricted-Use Data File and Electronic Codebook* (NCES 2004-092). National Center for Education Statistics, Institute of Education Sciences, U.S. Department of Education. Washington, D.C.

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

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

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

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

Measurement error. There was a concern in the ECLS-K that the individual mode of administration might inject additional and unwanted variance into both the individual and between-school components of variance in the cognitive scores. Since it is more difficult to standardize test administrations when tests are individually administered, this source of variance could contribute to high design effects if the individual assessors differed systematically in their modes of administration. 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.

A potential area for measurement error occurs with the NCATS and Two Bags Task components of the ECLS-B home visit. The parent-child interactions for these two components of the study were videotaped and coded later. The process of coding the tapes, however, is not problem-free. The videotape of the interaction must be of high quality to ensure valid coding. For example, field staff needed to tape the very beginning of the interaction and not interrupt it. The task of coding is further complicated by the coding staff's experience. Like the ECLS-B home visit field staff, the NCATS and Two Bags Task coders did not, for the most part, possess an extensive background in child development. Training the coding staff to reach 90 percent reliability proved difficult at times and often required additional training.

6. CONTACT INFORMATION

For content information about the ECLS project, contact:

Gail M. Mulligan
Phone: (202) 245-8413
E-mail: gail.mulligan@ed.gov

Mailing Address:
National Center for Education Statistics
Potomac Center Plaza
550 12th Street, SW
Washington, DC 20202

7. METHODOLOGY AND EVALUATION REPORTS

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

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.

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.

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

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

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.