

## Considerations Regarding the ECLS-B Target Population and for Analysis of Data from the Kindergarten 2006 and Kindergarten 2007 Rounds of Data Collection

This document discusses some of the issues associated with generalizing from the ECLS-B cohort to the population of children born in the United States in 2001 and with analyzing data from the kindergarten 2006 and kindergarten 2007 data collections. Section 1.1 discusses the target population for the ECLS-B. Section 2.1 summarizes analytic issues related to the timing of the assessments, the periodicity of the study, and the samples of children included in the two rounds of kindergarten data collection. It also includes examples of research questions that can be appropriately addressed with data from the kindergarten collections. Section 3.1 discusses analysis of the ECLS-B children as they enter kindergarten for the first time. Lastly, section 4.1 describes the ECLS-B sample of kindergarten repeaters and explains the possibilities for and limitations of analysis of this subsample.

### 1.1 Target Population

The ECLS-B is a birth cohort sample representative of children born in the United States in 2001. During the first two rounds of data collection (9 months and 2 years), information was collected to coincide with children's birth dates (i.e., the study employed a rolling, year-round data collection). Starting with the third round (preschool) and continuing through the final rounds (kindergarten 2006 and 2007), information was collected to coincide with the school academic calendar in order to minimize differences in the quantity of exposure children had to formal educational settings. For example, during the preschool round, data generally were collected during the fall and winter of 2005.<sup>1</sup> Likewise, during the kindergarten 2006 round, data were collected during the fall and winter of 2006.<sup>2</sup> Despite this change in the period of data collection, the cohort remains an age-based cohort. It is not a grade-based cohort.

Researchers should keep the target population in mind in order to develop questions appropriate for the study design. Specifically, researchers should frame their thinking in terms of children who were born in the United States in 2001, the children's development across the early years, and their outcomes and experiences as they entered kindergarten. The sample was not freshened in any round to be representative of a cross-section of children of a particular age or grade at the time the data were collected. For example, while the third round of collection is referred to as the preschool

---

<sup>1</sup> Every effort was made to collect data from children during this time period. However, some children were not assessed until later in the academic calendar year.

<sup>2</sup> As was true for the preschool collection, some children were not assessed until later in the academic calendar year.

ECLS-B Considerations for Analysis of ECLS-B Data > Analysis of Children at Kindergarten Entry  
> Slide 36 of 44

collection, the ECLS-B is not a study just of children in preschool, nor is it representative of all children in preschool, because the study does not include children who came to the United States after birth. Rather, the ECLS-B's study design supports research on children born in the U.S. during 2001, their development across the early years, their outcomes and experiences when they were preschool age, and how their experiences at preschool age relate to their experiences when they were kindergartners.

Similarly, data collected during the kindergarten 2006 and kindergarten 2007 rounds of the ECLS-B are *not* representative of all children enrolled in kindergarten in those data collection years. For example, neither collection includes kindergartners who were born outside the U.S., and the kindergarten 2006 collection excludes children who were repeating kindergarten in 2006 after having attended kindergarten in 2005. It would be inappropriate to make statements about all kindergartners in the United States using ECLS-B data. As mentioned above, the data can be used to describe the cohort of children born in the U.S. in 2001 at the time they entered kindergarten.

## 2.1 Timing of the Assessment and Data Collection and Samples of Children in the Kindergarten Collections

### 2.1.1 Timing of the assessment and data collection

In both longitudinal and cross-sectional analyses it is critical to consider assessment timing: children's age at assessment, assessment dates, and the time interval between assessments. Assessment dates ranged from August 2005 to June 2006 for the preschool round, from September 2006 to March 2007 for the kindergarten 2006 round, and from November 2007 to March 2008 for the kindergarten 2007 round.<sup>3</sup> Children assessed in the later part of one of these rounds of data collection may have an advantage over children assessed earlier in the data collection period because they had more time to learn the skills and knowledge being assessed. Substantial differences in the interval between assessments also may affect analysis of gain scores, or measures of how much children have learned over a particular time period. The ECLS-B includes information on the data file about the child's age at assessment (X\*ASAGE) and month and year of assessment (X\*ASMTMM; X\*AMMTYY) so that researchers can account for child age, the timing of the assessment, and the time between assessments when applicable for their research question.<sup>4</sup>

---

<sup>3</sup> The 9-month data were collected on a rolling basis from the fall of 2001 to the fall of 2002. The 2-year data were collected on a rolling basis from the fall of 2003 to the fall of 2004.

<sup>4</sup> Analysts looking for round-specific variable names should replace the asterisk (\*) with the number 1 for the 9-month wave, the number 2 for the 2-year round, the number 3 for the preschool round, the number 4 for the kindergarten 2006 round, and the number 5 for the kindergarten 2007 round.

## ECLS-B Considerations for Analysis of ECLS-B Data > Analysis of Children at Kindergarten Entry > Slide 36 of 44

When using ECLS-B kindergarten round data, it is also important that users keep in mind that the data collection period was the fall of the academic school year for most children. Therefore, sample children did not have a full year of kindergarten experience at the time of assessment so conclusions cannot be drawn about how school and teacher characteristics affect child outcomes. Rather, these data provide information about the abilities and experiences of children born in 2001 when they entered school for the first time or when they started their second year of kindergarten for those who repeated the grade in the 2007-08 school year.

### 2.1.2 Samples of children included in each of the kindergarten data collections

The transition from an age-based data collection to a grade-based/academic year collection resulted in kindergarten data not being collected for all children in the same round of data collection. Specifically, information about the children's experiences at kindergarten entry had to be collected in two school years, 2006-07 and 2007-08, because children born later in 2001 were not age-eligible for kindergarten in 2006-07 and some children experienced a delayed entry (i.e., they did not enter kindergarten when they were age-eligible to do so). The data collection in the 2006-07 school year, i.e., the kindergarten 2006 collection, included all children in the ECLS-B, irrespective of their enrollment in kindergarten; approximately 75 percent of children were in kindergarten or higher for this round. For the kindergarten 2007 data collection, which occurred in the 2007-08 school year, only a subsample of the ECLS-B children were included. This subsample comprised children who were not in kindergarten or higher in the 2006 collection, children who were in kindergarten in the 2006 collection and were identified by parents as repeating kindergarten in 2007-08, and twins of children in each of these groups.

Analysis of data collected during the kindergarten 2006 and kindergarten 2007 rounds requires careful consideration of the group of children to which users want to generalize and whether the outcomes of interest may be related to school exposure and/or the child's age at the time of assessment. Due to the differences in the samples included in each kindergarten collection, data from each round, or from both rounds combined, are better suited to answer different types of research questions. For example, if users are interested in children's physical development (e.g., average height and weight) at about age 5, analysts could use the physical measurements collected in kindergarten 2006 when most ECLS-B children were about this age. However, analysts who are interested in academic outcomes such as children's early reading knowledge and skills should be aware that estimates of overall averages for all children using data collected in kindergarten 2006 reflect children with different levels of exposure to formal

**ECLS-B Considerations for Analysis of ECLS-B Data > Analysis of Children at Kindergarten Entry  
> Slide 36 of 44**

schooling, since approximately 75 percent had at least started kindergarten or a higher grade at the time of assessment and about 25 percent were not yet in kindergarten.

Data from the kindergarten 2006 round are best suited to answer questions about children born in 2001 when they were approximately 5 years old. Examples of such questions are:

- What percentage of children born in the U.S. in 2001 were overweight or at risk for being overweight when they were about 5 years old? What child and family characteristics are related to the likelihood that a child born in 2001 was overweight when he or she was about 5 years old?
- What percentage of children born in the U.S. in 2001 had been diagnosed with certain illnesses, such as asthma or repeated ear infections, by the time they were about 5 years old? What are the correlates of these diagnoses?

Combining kindergarten 2006 data with kindergarten 2007 data allows analysts to answer questions about children born in 2001 as they entered kindergarten.

Examples of such questions are:

- What child and family characteristics relate to children's early reading and math ability as they enter kindergarten?
- Is participation in early care and education prior to kindergarten entry related to children's cognitive and/or socioemotional development as they enter kindergarten?
- What percentage of children began kindergarten on time (i.e., when they were age-eligible to do so), what percentage of children had a delayed entry, and what differences exist between these groups of children?

Finally, some research questions may be best answered by analyzing data collected only in the kindergarten 2007 round of data collection. Primarily, users interested in studying children who repeat kindergarten will rely on data collected for part of the subsample followed in 2007. Examples of possible research questions are:

- What percentage of children born in 2001 attended kindergarten in 2006 and repeated kindergarten in 2007?
- Within the cohort, did children who were repeating kindergarten in the 2007-08 school year have significantly different assessment scores than children entering kindergarten for the first time in 2007-08?
- How is the likelihood of repeating kindergarten related to child, family, and school characteristics? Who is most likely to repeat kindergarten?

## ECLS-B Considerations for Analysis of ECLS-B Data > Analysis of Children at Kindergarten Entry > Slide 36 of 44

Note that there are limitations to the analyses that can be done to answer the latter question. These limitations are discussed more fully in section 4.1.

### 3.1 Analysis of children as they enter kindergarten for the first time

In order to accurately estimate any characteristic for the ECLS-B children when they entered kindergarten, analysts should utilize the data from the round in which the child was first in kindergarten. Due to the variety of types of kindergarten programs that the children attended, it is not always completely clear from the data obtained from parents or teachers, or both, if and when a child was first in kindergarten. For example, some children attended a 1-year kindergarten program in a public or private school, some attended a 2-year kindergarten program, some attended kindergarten in an early learning center where they also attended preschool, and some were homeschooled. This section identifies the variables available on the data file that provide information on children's kindergarten enrollment status and makes recommendations about how to restructure the data and what weights to use in order to assure that the information analyzed is reflective of the children's first kindergarten year.

#### 3.1.1 Identifying children's point of kindergarten entry

Several composite variables exist on the data file to help users identify the appropriate sample of children for their analysis. The variable XKWHENK<sup>5</sup> provides information on when children were first in kindergarten (see table 1). The variable takes into account data from both kindergarten collections to identify when children were first enrolled in school or homeschooled and, for non-homeschooled children, their grade level (a very small number of children were identified as being in first grade in the first round in which they were enrolled in school). The majority of children are identified as enrolled in kindergarten (and not homeschooled) for the first time in 2006 or 2007. XKWHENK also identifies children who were enrolled in school for the first time in 2006 or 2007 who were identified as being in first grade, whose grade level was unknown, or who were in ungraded or multigrade classrooms, and children who were not enrolled in school at the time of either the 2006 or 2007 parent interview. Users must decide how they want to treat homeschooled children and children not identified as being in kindergarten when conducting their analysis of children as they entered kindergarten for the first time.

---

<sup>5</sup> The prefix for this composite variable is "XK" as opposed to X4 or X5 because this variable was created using information from both 2006 and 2007 data collections. It is defined for all cases with valid data in kindergarten 2006. All variables and weights with the prefix "XK" are intended for use in analysis of data related to children's first experience in kindergarten.

ECLS-B Considerations for Analysis of ECLS-B Data > Analysis of Children at Kindergarten Entry > Slide 36 of 44

**Table 1. ECLS-B sample’s kindergarten enrollment status: 2006–07 and 2007–08**

XKWHENK value	XKWHENK value label	n	Percent
1	Not enrolled in school in 2006 or 2007	#	0.2
2	First-time kindergartner in 2006	5,050	72.0
3	First-time kindergartner in 2007	1,550	22.2
4	Straight to 1 <sup>st</sup> grade in 2006	#	0.2
5	Straight to 1 <sup>st</sup> grade in 2007	#	0.3
6	First enrolled in school in 2006 (grade unknown or ungraded)	100	1.3
7	First enrolled in school in 2007 (grade unknown or ungraded)	#	0.3
8	Homeschooled (beginning in 2006)	100	1.6
9	Homeschooled (beginning in 2007)	50	0.4
99	Unknown <sup>1</sup>	100	1.5

# Rounds to zero.

<sup>1</sup>Cases classified as “unknown” are cases that were not yet enrolled in kindergarten in 2006 and that did not respond to the survey in 2007.

NOTE: Table percentages are unweighted and, therefore, are only representative of the ECLS-B sample as opposed to all children born in the U.S. in 2001. Unweighted sample sizes (*n*) have been rounded to the nearest 50. Detail may not sum to total due to rounding.

SOURCE: U.S. Department of Education, National Center for Education Statistics, Early Childhood Longitudinal Study, Birth Cohort (ECLS-B), kindergarten 2006 (2006–07) and kindergarten 2007 (2007–08) data collections.

### 3.1.2 Creating “kindergarten round” variables

In order to conduct analyses of the ECLS-B children’s experiences at kindergarten entry, users will need to create new “kindergarten round” variable from existing variables so that the data analyzed for a given case pertain to the round in which the child was first in kindergarten. For example, if users wish to determine the mean early reading scale score for children when they enter kindergarten, users would need to construct a new early reading scale score variable using data from the 2006 early reading scale score (X4RSCR2) for those children who were first-time kindergartners in 2006 and the 2007 early reading scale score (X5RSCR2) for those children who were first-time kindergartners in 2007. Users will need to create new variables in this way for the majority of the variables they use in any analysis of the ECLS-B children at kindergarten entry.

The variable XKKDATA can be used to determine the round of data to be used for each child. Variables from the kindergarten 2006 round should be used when XKKDATA=1, and variables from the kindergarten 2007 round should be used when XKKDATA=2 for this kind of analysis.

The tables below demonstrate the differences in estimates that are obtained using data from the two kindergarten collections separately compared to those that are obtained when the data are combined to represent the time of kindergarten entry. Tables 2 and 3 give the frequency distributions for several commonly-used



**ECLS-B Considerations for Analysis of ECLS-B Data > Analysis of Children at Kindergarten Entry > Slide 36 of 44**

demographic variables separately for the kindergarten 2006 collection (which included all children, regardless of school enrollment or grade level) and the kindergarten 2007 collection (which included only children who were not in kindergarten or higher for the first time in 2006, kindergarten repeaters, and twins of these children). Table 4 gives the frequency distributions for these same variables using data from both rounds to provide estimates for the ECLS-B children when they were in kindergarten for the first time. Tables 5 and 6 give the mean early reading and math scale scores separately for kindergarten 2006 and kindergarten 2007. Table 7 gives the mean early reading and math scale scores using data from both rounds to provide estimates for the ECLS-B children when they were in kindergarten for the first time.

**Table 2.** Frequency and percentage distribution of children born in the U.S. in 2001, by selected child and family characteristics: 2006–07

Child and family characteristics	Unweighted n	Weighted N (to the nearest 100)	Weighted percentage distribution
Total	6,950	3,929,600	100.0
Child's sex			
Male	3,550	2,012,600	51.2
Female	3,450	1,917,000	48.8
Child's race/ethnicity			
White, non-Hispanic	2,850	2,106,700	53.8
Black, non-Hispanic	1,100	544,500	13.9
Hispanic	1,400	986,100	25.2
Asian, non-Hispanic	800	102,200	2.6
American Indian/ Alaska Native, non-Hispanic	250	19,700	0.5
Other, non-Hispanic	600	160,000	4.1
Family poverty status			
At or above poverty threshold	5,300	2,974,600	75.7
Below poverty threshold	1,650	955,000	24.3

NOTE: Data are weighted by W4R0. Only those cases with a valid weight are included in the estimates. Unweighted sample sizes (*n*) have been rounded to the nearest 50. Detail may not sum to total due to rounding.

SOURCE: U.S. Department of Education, National Center for Education Statistics, Early Childhood Longitudinal Study, Birth Cohort (ECLS-B), kindergarten 2006 data collection (2006–07).

**ECLS-B Considerations for Analysis of ECLS-B Data > Analysis of Children at Kindergarten Entry > Slide 36 of 44**

**Table 3.** Frequency and percentage distribution of children born in the U.S. in 2001 and in kindergarten in 2007-08, by selected child and family characteristics: 2007–08

Child and family characteristics	Unweighted n	Weighted N (to the nearest 100)	Weighted percentage distribution
Total	1,900	1,247,000	100.0
Child's sex			
Male	1,050	682,400	54.7
Female	800	564,600	45.3
Child's race/ethnicity			
White, non-Hispanic	900	696,800	55.9
Black, non-Hispanic	300	184,600	14.8
Hispanic	350	288,700	23.1
Asian, non-Hispanic	150	21,000	1.7
American Indian/ Alaska Native, non-Hispanic	50	7,200	0.6
Other, non-Hispanic	150	48,800	3.9
Family poverty status			
At or above poverty threshold	1,450	925,900	74.2
Below poverty threshold	450	321,100	25.8

NOTE: Data are weighted by W5R0. Only those cases with a valid weight are included in the estimates. Unweighted sample sizes (*n*) have been rounded to the nearest 50. Detail may not sum to total due to rounding.

SOURCE: U.S. Department of Education, National Center for Education Statistics, Early Childhood Longitudinal Study, Birth Cohort (ECLS-B), kindergarten 2007 data collection (2007–08).



**ECLS-B Considerations for Analysis of ECLS-B Data > Analysis of Children at Kindergarten Entry > Slide 36 of 44**

**Table 4.** Frequency and percentage distribution of children born in the U.S. in 2001, by selected child and family characteristics when they entered kindergarten for the first time: 2006–07 and 2007–08

Child and family characteristics	Unweighted n	Weighted N (to the nearest 100)	Weighted percentage distribution
Total	6,850	3,913,600	100.0
Child's sex			
Male	3,500	2,003,900	51.2
Female	3,400	1,909,700	48.8
Child's race/ethnicity			
White, non-Hispanic	2,800	2,097,200	53.7
Black, non-Hispanic	1,050	542,600	13.9
Hispanic	1,400	981,600	25.1
Asian, non-Hispanic	750	101,300	2.6
American Indian/ Alaska Native, non-Hispanic	250	19,400	0.5
Other, non-Hispanic	600	161,200	4.1
Family poverty status			
At or above poverty threshold	5,250	2,978,700	76.1
Below poverty threshold	1,600	934,900	23.9

NOTE: Data are weighted by WKR0. Only those cases with a valid weight are included in the estimates. Unweighted sample sizes (*n*) have been rounded to the nearest 50. Detail may not sum to total due to rounding.

SOURCE: U.S. Department of Education, National Center for Education Statistics, Early Childhood Longitudinal Study, Birth Cohort (ECLS-B), kindergarten 2006 (2006–07) and kindergarten 2007 (2007–08) data collections.

**ECLS-B Considerations for Analysis of ECLS-B Data > Analysis of Children at Kindergarten Entry  
> Slide 36 of 44**

**Table 5.** Average early reading and mathematics scale scores for children born in the U.S. in 2001, by selected child and family characteristics: 2006–07

Child and family characteristics	Average reading scale score	Average math scale score
Total	38.60	40.40
Child's sex		
Male	37.40	39.96
Female	39.85	40.86
Child's race/ethnicity		
White, non-Hispanic	40.67	42.58
Black, non-Hispanic	35.51	36.80
Hispanic	34.84	37.16
Asian, non-Hispanic	48.73	46.53
American Indian/ Alaska Native, non-Hispanic	31.10	33.72
Other, non-Hispanic	38.34	40.02
Family poverty status		
At or above poverty threshold	40.75	42.12
Below poverty threshold	31.73	34.92

NOTE: Data are weighted by W4R0.

SOURCE: U.S. Department of Education, National Center for Education Statistics, Early Childhood Longitudinal Study, Birth Cohort (ECLS-B), kindergarten 2006 data collection (2006–07).

**ECLS-B Considerations for Analysis of ECLS-B Data > Analysis of Children at Kindergarten Entry > Slide 36 of 44**

**Table 6.** Average early reading and mathematics scale scores of children born in the U.S. in 2001 and in kindergarten in 2007-08, by selected child and family characteristics: 2007–08

Child and family characteristics	Average reading scale score	Average math scale score
Total	48.95	47.72
Child's sex		
Male	47.56	47.19
Female	50.61	48.35
Child's race/ethnicity		
White, non-Hispanic	51.05	50.31
Black, non-Hispanic	47.65	44.64
Hispanic	44.42	43.48
Asian, non-Hispanic	56.31	53.17
American Indian/ Alaska Native, non-Hispanic	37.86	35.36
Other, non-Hispanic	49.01	46.50
Family poverty status		
At or above poverty threshold	50.79	49.51
Below poverty threshold	43.62	42.51

NOTE: Data are weighted by W5R0.

SOURCE: U.S. Department of Education, National Center for Education Statistics, Early Childhood Longitudinal Study, Birth Cohort (ECLS-B), kindergarten 2007 data collection (2007–08).

**ECLS-B Considerations for Analysis of ECLS-B Data > Analysis of Children at Kindergarten Entry > Slide 36 of 44**

**Table 7.** Average early reading and mathematics scale scores for children born in the U.S. in 2001 when they entered kindergarten for the first time, by selected child and family characteristics: 2006–07 and 2007–08

Child and family characteristics	Average reading scale score	Average math scale score
Total	44.00	44.07
Child's sex		
Male	43.07	43.89
Female	44.96	44.25
Child's race/ethnicity		
White, non-Hispanic	46.46	46.56
Black, non-Hispanic	41.26	40.47
Hispanic	39.44	40.33
Asian, non-Hispanic	52.15	48.87
American Indian/ Alaska Native, non-Hispanic	37.07	37.29
Other, non-Hispanic	44.22	43.82
Family poverty status		
At or above poverty threshold	46.07	45.68
Below poverty threshold	37.32	38.83

NOTE: Data are weighted by WKR0.

SOURCE: U.S. Department of Education, National Center for Education Statistics, Early Childhood Longitudinal Study, Birth Cohort (ECLS-B), kindergarten 2006 (2006–07) and kindergarten 2007 (2007–08) data collections.

**3.1.3 Analytic weights for use in analysis of children when they entered kindergarten for the first time**

There is a series of weight variables designed specifically for use in analyses that include data collected during the round in which the child first entered school or a homeschooling arrangement. Analysis weights for which the variable name prefix is “WK” have been developed for use in analyzing parent data, as well as parent data in conjunction with data from other ECLS-B components (child assessment, resident and nonresident father questionnaires, care provider interview, observation of the child care setting, teacher questionnaire, and school data), for children entering their first year of formal schooling, whether that was in the kindergarten 2006 or the kindergarten 2007 collection. For a more detailed discussion of these weights, including information about how to choose which weight to use in different analyses, see sections 5.1.5, 5.1.6, and 5.3.5.2 of the *Early Childhood Longitudinal Study, Birth Cohort (ECLS-B), Kindergarten 2006 and 2007 Data File User's Manual* (Snow et al. 2009).

## 4.1 Analysis of children who repeated kindergarten in 2007-08

As mentioned above, the kindergarten 2007 data collection included a subset of the ECLS-B sample children who were in kindergarten during the kindergarten 2006 data collection and were identified as repeating kindergarten in the 2007-08 school year.

To identify children who were repeating, the ECLS-B fielded a screener prior to the start of the 2007-08 school year for parents of children who were in kindergarten during the 2006 data collection. Children who were reported by the parent in the screener to be in kindergarten again during the 2007-08 school year were included in the kindergarten 2007 data collection along with children who were starting school for the first time in the 2007-08 school year.

### 4.1.1 Identifying children who were repeating kindergarten

To help users identify children repeating kindergarten, the data file includes the variable X5RPTR. This variable contains three categories: (1) Yes, child repeating kindergarten; (2) No, child not repeating kindergarten; and (3) child in 2-year, transitional, or ungraded program (table 8). The variable X5RPTR is set to “not applicable” for those children who were not yet enrolled in school in 2006 or were homeschooled in either 2006 or 2007. Additionally, X5RPTR is not consistent with XKWHENK for every case. Twins of those cases eligible for inclusion in the kindergarten 2007 round also were included regardless of their grade level in 2006 or their repeater status in 2007. Those children who entered kindergarten for the first time in 2006 and were not repeating in 2007 but were included in the kindergarten 2007 collection because their twin was either a first-time kindergartner or a kindergarten repeater in 2007 have X5RPTR = 2 (not repeating). Additionally, children who were in an ungraded or multigrade classroom in 2006 and were enrolled in first or second grade in 2007 were considered “not repeating.”

Users interested in limiting their analysis to children who repeated kindergarten during the 2007-08 school year do not need to create new variables as was explained above for analyses of children at kindergarten entry. The round 5 (kindergarten 2007) variables present information collected the year the children repeated. Depending on the nature of their research question, users may want to limit their analytic sample to those children who were repeating their kindergarten year in 2007-2008 by selecting children for whom X5RPTR = 1. Several variables can be used to identify children who were in kindergarten for the first time in 2006 who were not kindergarten repeaters in 2007. If the child was part of the 2007 data collection (i.e., if he or she was enrolled in a 2-year or transitional kindergarten class or was the twin of a case eligible for the 2007

ECLS-B Considerations for Analysis of ECLS-B Data > Analysis of Children at Kindergarten Entry > Slide 36 of 44

data collection), users can identify the child using X5RPTR = 2 or 3. Users should note that if a child was reported to be homeschooled for kindergarten in either 2006 or 2007, X5RPTR = -1 (not applicable) because X5RPTR is defined only for children who were enrolled in a school. Kindergartners in 2006 who were not part of the 2007 sample can be identified as non-repeaters using the kindergarten screener status variable (XKSTATUS). If XKSTATUS = 5,<sup>6</sup> then the child’s parent responded to the screener and indicated that the child would not be repeating kindergarten in 2007. Children who went directly into first grade in 2006 were not screened and were not eligible for the 2007 data collection.

It is important to note that data were collected for children repeating kindergarten only for those who were in kindergarten for the first time in 2006. Children who entered kindergarten for the first time in 2007 (either because they were not age-eligible to enter kindergarten in 2006-07 or because their parents chose to delay kindergarten entry) could have repeated kindergarten in the 2008-09 school year. Information about the repeater status and second year in kindergarten for children who were first in kindergarten in 2007-08 was not collected. For this reason, analysts should keep in mind that they cannot make statements about retention for the entire ECLS-B cohort.

**Table 8. ECLS-B sample’s kindergarten repeater status: 2007–08**

X5RPTR value	X5RPTR value label	n	Percent
1	Yes, child is repeating kindergarten	200	9.6
2	No, child is not repeating kindergarten	#	0.6
3	Child is in a 2-year, transitional, or ungraded program	50	3.6
-1	Not applicable <sup>1</sup>	1,650	86.2

# Rounds to zero.

<sup>1</sup>Cases are coded “not applicable” for repeater status if they were not yet enrolled in school in the kindergarten 2006 collection or if they were homeschooled in either round.

NOTE: Table percentages are unweighted and, therefore, are only representative of the ECLS-B sample as opposed to all children born in the U.S. in 2001. Unweighted sample sizes (*n*) have been rounded to the nearest 50. Detail may not sum to total due to rounding.

SOURCE: U.S. Department of Education, National Center for Education Statistics, Early Childhood Longitudinal Study, Birth Cohort (ECLS-B), kindergarten 2007 (2007–08) data collections.

<sup>6</sup> For more information on XKSTATUS, see the *Early Childhood Longitudinal Study, Birth Cohort (ECLS-B), Kindergarten 2006 and 2007 Data File User’s Manual* (Snow et al. 2009), where it is described in more detail.

ECLS-B Considerations for Analysis of ECLS-B Data > Analysis of Children at Kindergarten Entry  
> Slide 36 of 44

4.1.2 Analytic weights for analysis of children born in the U.S. in 2001 who were repeating kindergarten during the 2007-08 school year

There is a series of weight variables designed specifically for use in analysis of kindergarten 2007 data. Analysis weights for which the variable name prefix is “W5” have been developed for use in analyzing the 2007 parent data, as well as parent data in conjunction with data from other ECLS-B components (resident and nonresident father questionnaires, care provider interview, teacher questionnaire, and school data). Although the majority of W5 weights are defined for all respondents to the kindergarten 2007 data collection, data for children who were identified as repeating kindergarten at the 2007 round (identified by the variable X5RPTR) can be analyzed using the W5 weights by specifying the repeaters as an analysis domain. Additionally, the weight W54R0 was created for analyses examining the distribution of child’s grade and school enrollment status at the end of the study for the entire ECLS-B cohort. Analyses of kindergarten repeaters that include both repeaters and non-repeaters in the analytic sample may find this to be the most appropriate weight for their analysis because this weight accounts for nonresponse to both the 2006-2007 screener questionnaire as well as the 2007 data collection round. Additionally, children who were in kindergarten in 2006 and did not participate in the 2007 data collection do not have a valid weight for the W5 weights; therefore, W54R0 needs to be used to include these children in an analysis with kindergarten repeaters. For a more detailed discussion of these weights, including information about how to choose which weight to use in different analyses, see sections 5.1.5, 5.1.6, and 5.3.5.1 of the *Early Childhood Longitudinal Study, Birth Cohort (ECLS-B), Kindergarten 2006 and 2007 Data File User’s Manual* (Snow et al. 2009).