

Slide 1 of 26

Title Slide: Data Collected Through the ECLS-K

Slide 2 of 26

This module provides detailed information about the direct and indirect assessment used to gather information on children's cognitive knowledge and skills, socioemotional development, and physical development.

Additionally, this module highlights some design issues that should be considered when analyzing ECLS-K data.

Throughout this module, underlined blue screen text indicates a link to additional resources.

Slide 3 of 26

The ECLS-K assessments were designed to measure knowledge, skills, and development at particular points in time as well as growth across time. The purpose of the cognitive assessments was to measure a broad range of important knowledge and skills representing what is typically covered by schools' curricula for the particular grade of interest. The design of the study was to be as inclusive as possible, capturing direct assessment information from a diverse, national population of children. Thus, study procedures were developed to accommodate children with varying needs and abilities, such as children with disabilities and those from homes in which English was not the primary language spoken. For example, accommodations were made for children with disabilities such as allowing a health care aide to be present during the assessment or allowing the children to use adaptive devices that they typically used at school. A very small subset of children were excluded from the direct assessments, including those whose Individualized Education Programs, or IEPs, indicated that they should not participate in standardized assessments, those who required an assessment in Braille or a sign language interpreter, and those whose IEPs required them to be assessed using large print materials. Children excluded from the direct assessments were still included in other aspects of the study, including the indirect assessments of their cognitive knowledge and skills. Lastly, the design of the assessments was sensitive to the time needed for administration because the ECLS-K assessed children at the school during their school day. Thus, assessments were developed that captured as much information as possible in as little time as possible to still yield psychometrically sound measures.

Slide 4 of 26

The ECLS-K assessment included both direct and indirect measures of children's knowledge, skills, and development in multiple domains. Information was collected on children's cognitive knowledge and skills, socioemotional development, and physical development.

Data Collected Through the ECLS-K

There are two ways to access the information within this module. You can click on one of the links above, which will take you directly to detailed information about the specific domain of interest. Or you can click the 'next' button to advance to the next screen of the module and view the contents in the order presented on this slide. At the end of each section for a particular domain, you will be provided with options to either return to this page and select a different domain or click the 'next' button to move directly into the section for the next domain.

Slide 5 of 26

In the cognitive domain, children's reading and mathematics knowledge and skills were measured in every round of data collection. A general knowledge assessment that included questions related to both science and social studies was included in the kindergarten and first-grade collections. The general knowledge assessment was replaced with an assessment focusing only on science knowledge and skills in the third-, fifth-, and eighth-grade collections.

Slide 6 of 26

The ECLS-K required direct measures that provided information on the development of children's reading, mathematics, general knowledge, and science knowledge skills from kindergarten entry through spring of their eighth-grade year. A thorough review of existing measures and assessments showed that there was no one existing assessment that met the study goals. For this reason, assessments were developed specifically for use in the ECLS-K.

To develop psychometrically sound assessments, experts in test development conducted a review of grade-level curricula and standards and consulted child development and content area experts to determine what specific knowledge and skills should be measured within each content area. Information collected during this process of review and consultation was used to develop an assessment framework that guided the construction of the assessments for each grade in the study, including the selection of test items. When possible, well-tested items from existing measures were used. When no items tapping desired content that worked within the ECLS-K assessment design were available, new items were developed. The assessment items underwent rigorous pilot tests and large scale field tests. Items that performed the best were selected for the final assessments used in the national data collections.

More detailed information about the development of the direct cognitive assessments is provided in the study psychometric reports available under the Technical/Methodological heading on the ECLS-K website's study publications and products page.

Slide 7 of 26

The ECLS-K assessments are two-stage adaptive assessments consisting of a first-stage routing test and second-stage tests of different difficulty levels. For each domain, all children first received the same set of items in the routing test. Based on their

Data Collected Through the ECLS-K

performance on these items, children were routed to a second-stage test that was tailored to their demonstrated level of ability. As a result of this assessment design, no child was administered all of the items in the assessment.

Administration was un-timed in that children were given as much time as they needed to complete the assessments. However, the assessments were designed to take an average of about an hour per child. From kindergarten through fifth grade, the assessments were administered one-on-one by a trained assessor using computer-assisted personal interview, or CAPI, technology. The computer prompted the assessor to administer the items using a visual stimulus shown to the children in a spiral-bound book called an easel. For each assessment item, the computer also provided the assessor with a standardized administration protocol, including the question to be read verbatim to the child and any instructions that should be provided to the child. Assessors entered the children's responses into the computer.

In the eighth-grade collection, groups of ECLS-K students who attended the same school were assessed in a single, proctored, paper-and-pencil group administration.

Slide 8 of 26

As noted earlier, efforts were made to include children who spoke a language other than English in the ECLS-K assessment.

Field staff checked the school records to determine the children's home language or, if records were not available, requested this information directly from children's teachers.

In the kindergarten and first-grade data collections, a brief language screener, the Oral Language Development Scale, or OLDS, was given to those children whose home language was not English. The screener determined if children understood English well enough to receive the direct child assessment in English. Children who achieved at least a minimum score on the language screener received the full ECLS-K direct assessment battery in English. Children who did not achieve the established cut score on the language screener received a reduced version of the ECLS-K assessments, as shown in this diagram.

Spanish-speaking children who did not achieve the cut score on the English OLDS were administered a Spanish version of the OLDS, as well as the mathematics and psychomotor assessments translated into Spanish. They also had their height and weight measured. Children whose primary home language was not Spanish who did not achieve at least a minimum score on the language proficiency screener only had their height and weight measured.

The OLDS was not administered beyond first grade because most children had passed the OLDS by the spring first-grade data collection.

Slide 9 of 26

Several types of scores were produced for each of the direct cognitive assessments.

Since the direct cognitive assessments were adaptive and included two stages, no child was administered all the items in the assessment. Item Response Theory, or IRT, procedures were used to calculate total scores regardless of which second-stage test a child received.

Four types of IRT-based scores were produced for each direct cognitive assessment for each round of data collection. IRT procedures yield an ability estimate, referred to as theta. The theta score is an estimate of a child's ability based on his or her performance on the items in the assessment that the child was administered.

Theta can be used to estimate the total number of items a child would have answered correctly if he or she had been administered all of the questions in the domain specific assessment. The scale score is the estimated total correct.

Both theta and overall scale scores are indicators of children's knowledge and skills at any given point in time and can be used to look at knowledge and skills over time. Therefore, both of these scores are appropriate for analysis of status in one round and of growth across rounds.

A standardized t-score is also provided for each domain in each round of data collection. The t-scores, which have a mean of 50 and a standard deviation of 10, are based on the samples of children assessed in a given round. These scores are recommended for analyses that examine assessment data at one point in time. T-scores are not well-suited for longitudinal analyses intended to examine growth over time, because they are normed to have a mean of 50 based on the sample assessed at each point in time.

A series of proficiency probabilities that reflect children's performance on a cluster of four items at about the same level of difficulty is provided for both reading and mathematics. Proficiency probabilities, which present information about performance in levels, can be used to better understand where growth is occurring by looking at what kinds of skills children are demonstrating, and, in particular, the difficulty of the those skills. The "highest proficiency level mastered" score indicates the highest level for which a student was estimated to be able to answer at least three of the four items in the cluster correctly. These scores are discussed in more detail in the next two slides.

Some items in the reading assessment measured a common construct but were not included in the proficiency probability scores because they did not fit well within any of the levels. These items were used to develop two different scores, a "print familiarity" and "decoding text" score. The print familiarity score provides information about children's performance on items assessing the children's knowledge of the conventions of print, for example that reading occurs from left to right and that after a line of text ends, reading continues at the beginning of the next line. The decoding text score

Data Collected Through the ECLS-K

indicates the child's ability to read words that were unlikely to be part of most children's everyday vocabulary but could be sounded out phonetically.

Lastly, raw number right scores indicating the actual number of routing items a child answered correctly are provided for each content area. The information provided by these scores is somewhat limited, because the scores are based on a subset of items tapping many different constructs and provide no information about the difficulty of the items a child answered correctly.

Slide 10 of 26

This illustration shows the reading proficiency probability scores for the cognitive assessment in reading. The scores are identified as levels. The levels are hierarchical, with each level having items of higher difficulty than the items in the level below it. For example, the items in level 6 are more difficult than the items in level 4, which are more difficult than the items in level 2.

The names of the levels, shown here in parentheses, identify specific skills in order to provide a description of the type of skill that is either dominant in the cluster or the most difficult skill in the cluster. In reading, the easiest level is associated with letter recognition, while the most difficult level is associated with evaluating complex syntax.

Scores on the proficiency probabilities range from 0 to 1 and indicate the probability that a child has acquired the skills represented in the cluster of items. The higher a child's score is on a given proficiency probability, the more likely it is that the child had acquired the skills measured in the item set. These proficiency probabilities are not meant to be interpreted as subscale scores.

Slide 11 of 26

This illustration shows the mathematics proficiency probability scores for the cognitive assessment in mathematics. The easiest level is associated with recognition of numbers and shapes, while the most difficult level is associated with items related to the calculation of area and volume. As with the proficiency probabilities for reading, the levels are hierarchical, the level names indicate the type of skill that is either dominant in the cluster or the most difficult skill in the cluster, scores range from 0 to 1, and the scores indicate the probability that a child has acquired the skills represented in the cluster of items. The scores are not meant to be interpreted as subscale scores.

Slide 12 of 26

The ECLS-K also included indirect assessments of children’s cognitive knowledge and skills. These assessments, referred to as the Academic Rating Scale, or ARS, were designed to overlap and augment the direct cognitive measures. Classroom teachers were presented with a list of grade-level knowledge and skills and asked to provide information for each ECLS-K child in his or her classroom as to whether the child demonstrated proficiency on the focal skill. The items on the ARS provide information about both skills and process in the areas of language and literacy, mathematical thinking, general knowledge, social studies, and science. In kindergarten and first grade, the social studies and science domains were measured within an ARS pertaining to a broader General Knowledge domain. In third grade, there was a separate ARS for social studies and one for science. Social studies knowledge and skills were not measured indirectly after the third-grade data collection.

Slide 13 of 26

The data file includes both item-level data for each question included in the ARS and scale scores that represent the students’ teacher-reported performance on a hierarchy of skills, knowledge, and behavior related to a particular domain. The ARS scale was designed to provide information on children’s abilities at a given point in time, not necessarily over time. In addition, although the items are similar across grades, the actual items reflect performance criteria that increase in difficulty from one time point to the next. Moreover, the ARS scores are placed on different metrics relative to the item difficulty in a given grade. Therefore, change scores should not be calculated between time points. However, covariance models may be used to compare teachers’ ratings of performance in different grades.

This concludes the information about the direct and indirect assessments of cognitive knowledge and skills in the ECLS-K. To return to the full list of assessment domains to make another selection, click the “Return to List” button.

Slide 14 of 26

Both direct and indirect assessments were used to measure children’s socioemotional development. Direct assessment of children’s socioemotional development began in the third-grade data collection with items included in the student survey. Socioemotional development was directly assessed again in fifth and eighth grade. The indirect assessments consisted of questions about children’s socioemotional functioning that were asked of parents in the kindergarten and first-grade collections and of teachers from kindergarten through fifth grade.

Slide 15 of 26

Students completed the Self-Description Questionnaire, or SDQ, which was included in the student survey fielded in third, fifth, and eighth grade. The SDQ provides information about socioemotional development in many areas: students' perceived interest and competence in reading, mathematics, and all subjects; students' relationships with peers; externalizing problem behaviors, such as fighting and arguing with other children; and internalizing problem behaviors, such as feeling sad or lonely. The eighth-grade student survey included only a subset of the constructs measured in the earlier grades. The data file includes scale scores for each of these constructs, with values ranging from "1 = not at all true" to "4 = very true." Additionally, a supplementary dataset containing all the item-level SDQ data is available online.

Slide 16 of 26

As part of the eighth-grade student survey, children also provided information on their self-concept and locus of control. The self-concept items asked children about their level of agreement with statements about certain perceptions about themselves, such as "I feel good about myself" and "I am able to do things as well as most other people."

The locus of control items asked children about their level of agreement with statements concerning the amount of control of their own lives they thought they had, such as "I don't have enough control over the direction my life is taking" and "When I make plans, I am almost certain I can make them work."

The data file includes scores for each of these constructs that are standardized to have a mean of zero and a standard deviation of 1. The self-concept score is based on seven items, and a higher score indicates a more positive self-concept. The locus-of-control score is based on eight items, and a higher score indicates a greater perception of control over one's life.

Slide 17 of 26

Parents and teachers provided information on how often the child exhibited certain social skills and behaviors, with response options ranging from "Never" to "Very Often." The questions asked of parents and teachers measured similar constructs, but the questions were tailored to each respondent and the context in which they were observing the behaviors, that is, in the home or at school.

The items assessing socioemotional development were asked together in a set referred to as the Social Rating Scale, or SRS. The SRS includes some items taken verbatim from the Social Skills Rating System, or SSRS, some items that are modifications of original SSRS items, and some items that measure the same kinds of skills and behaviors captured in the SSRS but use wording developed specifically for the ECLS-K.

Slide 18 of 26

Several social-skill scales were developed based on teachers' responses to the SRS questionnaire items. The score on each scale is the mean rating on the items included in the scale. The teacher scales describe children's socioemotional development in five areas: self-control, interpersonal skills, externalizing problem behaviors, internalizing problem behaviors, and approaches to learning. For the third- and fifth-grade data collections, an additional scale describing peer relations was created. This scale combines responses on both the interpersonal items and self-control items that relate to peers.

These scales are included in the data file for each round in which the questions used to derive them were asked. Due to copyright restrictions, item-level data cannot be obtained for the items comprising any of the scales without publisher permission and a restricted-use license, with one exception. The approaches to learning items were developed specifically for the ECLS-K, rather than being taken from an existing, copyrighted source.

Note that the SRS items were not included in the eighth-grade data collection. Due to the organization of instruction in the majority of middle and high schools, where students have a different teacher for each subject, it would have been operationally difficult to determine which teacher could best answer these questions on socioemotional development.

Slide 19 of 26

Several social-skill scales also were developed based on parents' responses to the SRS questionnaire items, which were included in the parent interview only in the kindergarten and first-grade rounds of data collection. The score on each scale is the mean rating on the items included in the scale. The parent scales describe children's socioemotional development in five areas: self-control, social interaction, impulsive or overactive behaviors, sadness or loneliness, and approaches to learning. As with the teacher scales, the parent scales are included in the data file for each round in which the questions used to derive them were asked. The only item-level data that are publicly available are for the approaches to learning items. Other item-level data can be obtained with publisher permission and a restricted-use license.

This concludes the information about the direct and indirect assessments of socioemotional development in the ECLS-K. To return to the full list of assessment domains to make another selection, click the "Return to List" button.

Slide 20 of 26

Student's physical development was directly assessed in the ECLS-K.

At every data collection point in the study, students were weighed and had their height measured. Body Mass Index, also known as BMI, was calculated using this height and

Data Collected Through the ECLS-K

weight information. Variables indicating children's height, weight, and BMI are provided in the data file for every round of data collection.

Slide 21 of 26

In the fall of kindergarten only, both fine and gross motor skills were assessed. In order to assess fine motor skills, kindergartners were asked to copy basic figures, draw a person, and build a gate with wooden blocks. Gross motor skills were assessed by having the students balance on each foot, hop on one foot, skip, and walk backward in a straight line. The data file includes a fine motor skills composite variable that is the sum of the points for the seven fine motor skill tasks administered: build a gate, draw a person, and copy five simple figures. Children could receive up to two points for each of the first two tasks and one point for each of the figures, so the score ranges from 0 to 9. The data file also includes a gross motor skills composite variable that is the sum of the points for the four gross motor skill tasks administered. Children could receive up to two points for each of the tasks, so the score ranges from 0 to 8. A third composite, which is the sum of the fine and gross motor skills composite variables, is also available. Item-level motor skill data are not available due to copyright restrictions.

This concludes the information about the assessments of physical development in the ECLS-K. To return to the full list of assessment domains to make another selection, click the "Return to List" button.

Slide 22 of 26

When analyzing the assessment data, researchers should consider certain study design features that could have an effect on the analysis results, one of which is the timing of the assessment. Although efforts were made to assess students within a limited timeframe, each round of data collection spanned a few months. Most of the fall assessments were conducted between early October and mid-November, but some assessments fell outside this timeframe. Similarly, most of the spring assessments were conducted between mid-April and late May, but some assessments fell outside this timeframe. Children assessed earlier in a data collection round tended to have lower achievement scores than children assessed on a later date in that data collection round. Therefore, it is recommended that researchers account for the timing of the assessment in their analyses.

Additionally, because most children within a school were assessed within a few days of one another, the age at which children were assessed varies. Therefore, it is recommended that researchers account for the child's age at the time of assessment in their analyses of assessment data, because age may be related to various outcomes, for example physical development.

Slide 23 of 26

A consideration specific to longitudinal analyses is the elapsed time between assessments, which was not the same for all children. Children with more elapsed time between assessments tended to show greater achievement gains than those with

Data Collected Through the ECLS-K

shorter elapsed times because they had more time between assessments to acquire knowledge and skills. Elapsed time between assessments may also be an important consideration when examining health or physical development data.

Slide 24 of 26

Note that an incorrect version of the theta scores was included in the original release of the Kindergarten Through Eighth Grade Full Sample Public-Use Data and Documentation (2009-005) data file.

Researchers interested in using the theta scores in their analyses should review the Errata Sheet: ECLS-K K–8 Full Sample Public-Use Data File (NCES 2010-052) March 2010. The errata data file provides the correct version of the theta scores for rounds 1, 2, 3, 4, 5 and 6. The Round 7 thetas included in the ECLS-K K–8 Full Sample Public-Use Data File (NCES 2009-005) are correct; however, for ease of use, they are included in the errata data file.

To return to the full list of assessment domains to make another selection, click the “Return to List” button.

Slide 25 of 26

This module has provided you with detailed information about the direct and indirect assessments used to collect data on children’s cognitive knowledge and skills, socioemotional development, and physical development. Additionally, design issues that should be considered when analyzing ECLS-K data were also presented.

Slide 26 of 26

Important resources that have been provided throughout the module are summarized in this slide.

You may now proceed to the next module in the series, or click the exit button to return to the landing page.