Children's knowledge and skills at kindergarten entry in Illinois: Results from the first statewide administration of the Kindergarten Individual Development Survey

Appendix A. About the Kindergarten Individual Development Survey

Appendix B. Methods

Appendix C. Supporting analyses

See https://go.usa.gov/xVtfA for the full report.

Appendix A. About the Kindergarten Individual Development Survey

This appendix provides information on the Kindergarten Individual Development Survey, including details about the survey items and the process used to develop and pilot the survey.

Survey items

The 14 required items on the Kindergarten Individual Development Survey measure children's skills at kindergarten entry. Teachers rate each child on each item in one of six categories: building–earlier, building–middle, building–later, integrating–earlier, integrating–middle, and integrating–later. When rating children, teachers are provided with a definition for each item (table A1) and a definition and example of children's behaviors that are associated with each rating category for an item (figure A1).

Teachers can also report that they are unable to rate a skill. But according to guidance from the state, teachers should be able to rate all skills for all children, including English learner children and children who were eligible for an individualized education program. Teachers could observe English learner children using both English and children's home language. In general education settings, teachers are instructed to rate English learner children based on their reciprocal communication and conversation, their comprehension of age-appropriate text, their phonological awareness, and their letter and word knowledge. In addition, general education teachers should also rate English learner children using two additional alternate items and the child's communication skills in English.[1] Teachers in Spanish-English bilingual classrooms can fill out additional ratings that assess learning to communicate in Spanish.

---

[1] About 3 percent of teachers who submitted data used the alternate items, and these teachers were clustered in 41 districts. By contrast, 65 percent of all teachers taught English learner children, and 394 districts included English learner children.

Table A1. Description of the 14 required items in the Kindergarten Individual Development Survey provided to teachers when rating children, fall 2017 administration in Illinois

Item	Definition
Curiosity and initiative in learning	Child explores the environment in increasingly focused ways to learn about people, things, materials, and events.
Self-control of feelings and behavior	Child increasingly develops strategies for regulating feelings and behavior, becoming less reliant on adult guidance over time.
Engagement and persistence	Child increasingly persists in understanding or mastering activities, even if they are challenging or difficult.
Relationships and social interactions with familiar adults	Child develops close relationships with one or more familiar adults (including family members) and interacts in an increasingly competent and cooperative manner with familiar adults.
Relationships and social interactions with peers	Child becomes increasingly competent and cooperative in interactions with peers and develops friendships with several peers.
Classification of objects into groups based on their attributes	Child shows an increasing ability to compare, match, and sort objects into groups according to their attributes.
Number sense of quantity	Child shows developing understanding of number and quantity.
Ability to add and subtract small quantities	Child shows an increasing ability to add and subtract small quantities of objects.
Knowledge of shapes	Child shows an increasing knowledge of shapes and their characteristics.
Communication and use of language (expressive)	Child's communication develops from nonverbal communication to using language with increasingly complex words and sentences.
Reciprocal communication and conversation	Child engages in back-and-forth communication that develops into increasingly extended conversation.
Comprehension of age-appropriate text	Child develops capacity to understand details and ideas from age-appropriate text presented by adults.
Phonological awareness	Child shows increasing awareness of the sounds (elements) that make up language, including the ability to manipulate them in language.
Letter and word knowledge	Child shows increasing awareness of letters in the environment and their relationship to sound, and increasing understanding that letters make up words.

Source: Illinois State Board of Education, 2015.

Figure A1. Example of the item definition, rating categories, rating category definitions, and examples provided to teachers when rating children for the Kindergarten Individual Development Survey, fall 2017 administration in Illinois

Developmental Domain: ATL—REG — Approaches to Learning—Self-Regulation

ATL-REG 1: Curiosity and Initiative in Learning
Child explores the environment in increasingly focused ways to learn about people, things, materials, and events

Mark the latest developmental level the child has mastered:

Building			Integrating		
Earlier ○	Middle ○	Later ○	Earlier ○	Middle ○	Later ○
Explores through simple observations, manipulations, or asking simple questions	Explores by engaging in specific observations, manipulations, or by asking specific questions	Carries out simple investigations using familiar strategies, tools, or sources of information	Carries out multi-step investigations, using a variety of strategies, tools, or sources of information	Carries out experiments with things or materials, by systematically modifying actions and reacting to the results	Finds out about things, people, or events by comparing multiple sources of information, including experiments, books and pictures, and asking questions
Examples					
▸ Watches the fish in the fish tank intently after a conversation about how fish breathe underwater. ▸ Drops a marble in a maze and uses hands to follow its path as it rolls to the bottom. ▸ Asks, "What's that doing?" when seeing the compact disc player in the listening center.	▸ Compares leaves gathered on a nature walk by color or shape. ▸ Asks, "How do I make the story play?" while in the learning center. ▸ Manipulates pattern blocks to make different shapes. ▸ Squeezes a sponge to see how it works.	▸ Uses a magnetic wand to figure out which objects on a table it will lift up. ▸ Uses a magnifying glass to observe a caterpillar closely, and describes its pattern of colors and number of legs. ▸ Changes the compact disc to listen to a new story. ▸ Uses a communication device to learn about the new pet guinea pig.	▸ Examines images from informational books or a computer to learn about the habitats of different animals. ▸ Looks through a prism held up to the light, directing its motion until a rainbow of colors appears on the wall. ▸ Sets up a project, with an adult, that involves investigating the growth of lima bean plants with different amounts of water, and documents their growth.	▸ Makes a wooden block ramp steeper and steeper and runs a small metal car down it each time to find out what happens. ▸ Adds blue paint to a saucer of yellow paint a few drops at a time, stirring after each addition, to see how the green color changes. ▸ Kicks a ball into a play soccer goal repeatedly, placing the ball farther away (and at different angles) before each kick. ▸ Watches a cup of snow to see how long it takes to melt.	▸ Communicates, "But that's different from what my daddy told me," and asks, "Why?," after hearing an adult's response to a question about why plants are green. ▸ Sets up a ramp to experiment with whether it is true that objects roll down steeper ramps more quickly, after a peer shows that objects roll down steeper ramps more quickly than shallow ramps. ▸ Gathers information from books and the internet to create an environment for the classroom butterflies. ▸ Creates a model of a bridge, consulting pictures of bridges, talking with an adult, and experimenting with creating a bridge across a divide.

○ Child is emerging to the next developmental level
○ If you are unable to rate this measure, explain here:

Source: Illinois State Board of Education, 2015.

Survey development

Creating and validating the survey took several years. The process included developing the full 55 item instrument, piloting the survey, creating and validating a 14 item version of the survey to minimize the burden of data collection for teachers, and preparing for statewide implementation of the survey.

The Kindergarten Individual Development Survey was adapted from an existing kindergarten entry assessment in California at the request of the Illinois State Board of Education with the help of an advisory committee and the WestEd Center for Child and Family Studies. Beginning in 2010 a planning committee (consisting of staff from the Illinois State Board of Education, experts in child development, elementary school teachers and principals, advocates, and other educators) partnered with the WestEd Center for Child and Family Studies to develop the survey. The planning committee adapted the Desired Results Developmental Profile–Kindergarten, the kindergarten entry assessment used in California, to align with the Illinois Early Learning and Development Standards, the Illinois Early Learning Standards—Kindergarten, and the Illinois Social Emotional Learning Standards.

More than 400 districts in the state participated in piloting the Kindergarten Individual Development Survey over five years. From 2012 to 2016 Illinois piloted the survey with 26,495 Illinois kindergarteners in 2,913 kindergarten classrooms across 422 of the state's 745 districts using a version that contained 55 items. In 2015, based on feedback from the pilot, the Illinois State Board of Education requested that the WestEd Center for Child and Family Studies produce a shorter version that could still reliably and validly measure children's social and emotional, language and literacy, and math skills at kindergarten entry.

Using content experts and based on empirical analyses, the Illinois State Board of Education and its partners selected 14 items hypothesized to measure three domains. The Illinois State Board of Education—along with the advisory committee of early childhood advocates and educators, elementary teachers and administrators, child development specialists, and researchers from the University of California–Berkeley Evaluation and Research Center and the WestEd Center for Child and Family Studies—selected 14 items from the full survey based on the substantive importance of each item reported in the research literature and psychometric analyses confirming that the items loaded on factors representing key areas of development (K. Kriener-Althen, personal communication, June 8, 2017). Teachers were required to enter ratings for the 14 selected items for each child; teachers could also choose to enter ratings for the remaining 41 items.

A three-dimensional item response theory model was fit to data from the 2014/15 pilot to produce measures of three domains: language and literacy development, cognition: mathematics, and approaches to learning/social-emotional development (University of California–Berkeley Evaluation and Assessment Research Center & WestEd Center for Child and Family Studies, 2015). The values for these subscales ranged from –7 to 6 on a logit scale; higher logit values are associated with a higher probability of a higher score on an item or scale and signify higher levels of the corresponding measured trait. The internal scale reliability was above .90 for all three subscales (K. Kriener-Althen, personal communication, June 8, 2017). The correlation between the approaches to learning/social-emotional development subscale and the language and literacy development subscale was .97, the correlation between the approaches to learning/social-emotional development subscale and the cognition: mathematics subscale was .93, and the correlation between the cognition: mathematics subscale and the language and literacy development subscale was .98. In addition, a criterion zone boundary was created to indicate whether a child's skills in each domain passes a threshold associated with kindergarten readiness. The Illinois State Board of Education used this indicator of whether a child was ready in its statewide reporting of the 2017 data.

The state offered training and coaching prior to the 2017 survey administration. The state required universal use of the Kindergarten Individual Development Survey for the first time in fall 2017. To gain access to the KIDStech website, teachers were required to attend a one-day training session organized by the Illinois State Board of Education that covered how to use the survey and how to enter the ratings for each child online. Regional coaches facilitated the trainings in spring, summer, and fall 2017 in nine regions of the state. The state offered two additional types of teacher training during the 2017/18 school year: "Observing and Documenting for the Kindergarten Individual Development Survey" and the "Kindergarten Individual Development Survey Refresher." The Illinois State Board of Education also provided optional half-day workshops that school principals, assistant principals, and district administrators were encouraged to attend. In addition, the Regional Educational Laboratory Midwest offered training on data use for kindergarten teachers during the 2017/18 school year.

References

Illinois State Board of Education. 2015. *Kindergarten Individual Developmental Survey: User's guide & instrument*. Springfield, IL: Author.

University of California–Berkeley Evaluation and Assessment Research Center & WestEd Center for Child and Family Studies. (2015). *Preliminary KIDS analysis: Three subscales using a subset of measures across domains.* Berkeley, CA, and San Francisco, CA: Authors.

Appendix B. Methods

This appendix contains information about the study sample and methods used in the psychometric, descriptive, and qualitative analyses.

Study sample

The data from the fall 2017 administration of the Kindergarten Individual Development Survey included observations on 125,800 children in 6,275 classrooms in 2,077 public schools in 745 districts. According to data from the Illinois State Board of Education, 97 percent of schools and districts serving kindergartners reported data. The analytic sample excluded 9,683 children for whom teachers had entered ratings for fewer than 12 of the survey's 14 items and 2,401 children who had alternate ratings because of their English learner status.[2] The analytic sample included 113,716 children in 5,872 classrooms in 2,039 public schools in 738 districts. See table B1 for the demographic characteristics of kindergartners in the analytic sample and the entire 2017/18 Illinois kindergarten cohort.

Table B1. Demographic characteristics of children in the analytic sample and the entire 2017/18 Illinois kindergarten cohort for the fall 2017 administration of the Kindergarten Individual Development Survey (percent)

Subgroup	Analytic sample (n = 113,716)	2017/18 Illinois kindergarten cohort (n = 132,378)
Boys	51.0	51.3
English learner children[a]	26.5	19.1
Children eligible for the national school lunch program	47.4	48.0
Children eligible for an individualized education program	10.2	11.7
American Indian children	0.3	0.3
Asian children	5.4	5.4
Black children	16.3	16.6
Hispanic children	25.4	25.4
Pacific Islander children	0.1	0.1
White children	48.3	47.9
Multiracial children	4.3	4.4
Total sample size	113,716	132,378

a. The data on English learner children in the analytic sample and in the 2017/18 Illinois kindergarten cohort are from different points in time. The difference in the percentage of English learner children in each group may be due to the fact that the number of children classified as English learner children changes over the course of the school year.

Source: Authors' analyses of fall 2017 Kindergarten Individual Development Survey data and demographic data for the 2017/18 Illinois kindergarten cohort provided by the Illinois State Board of Education.

Exploratory and confirmatory factor analysis (research question 1)

To identify the latent constructs measured by the Kindergarten Individual Development Survey, the study team conducted both an exploratory factor analysis and a confirmatory factor analysis and used multiple approaches to determine the optimal factor structure.

---

[2] These alternate items were not included in the psychometric analyses because there was no way to equate the alternate items with the rest of the items.

Because all the items are on a numerical scale, the study used a polychoric correlation matrix and full information maximum likelihood estimation with Huber-White covariance adjustment (specifically, the robust maximum likelihood estimation option in Mplus) to account for missing data and non-normality. The analytic sample of 113,716 children was divided into two random subsamples of equal size. Exploratory factor analysis was conducted with the first random subsample to determine the optimal factor structure, and then confirmatory factor analysis was conducted with the second random subsample to test the model fit. The confirmatory factor analysis results served as validation for the empirical latent constructs identified by the exploratory factor analysis (Thompson, 2004).

For the exploratory factor analysis, squared multiple correlations were used as the initial communality estimates. The scree test of eigenvalues was used to provide information on the number of potential underlying factors. The eigenvalues in the scree plot represent the amount of variance that can be explained by each additional factor. Usually, factors with eigenvalues less than 1 will be discarded.

Geomin and promax rotations were tried to obtain the optimal factor structure. The final factor structure was selected on the basis of four criteria: the number of items with all loadings below 0.4 (wherein a smaller number suggests better fit; Stevens, 2009; Yates, 1988), the number of double loaders (wherein a smaller number suggests better fit), the closeness to a simple structure (the simpler the structure, the better; Fabrigar, Wegener, MacCallum, & Strahan, 1999), and the meaningfulness of each factor (Gadermann, Guhn, & Zumbo, 2012). In addition, the study team examined the correlation between factors to ensure sufficient discrimination validity of the factor structure. Given the clustering of children within classrooms, the study team also examined results from multilevel factor analyses using a saturation approach and applying both oblimin and geomin rotations to determine whether the factor structure was consistent when classroom-level variance was controlled for.

For the confirmatory factor analysis, goodness-of-fit statistics, such as comparative fit index and root mean square error of approximation (RMSEA), were used to determine model fit. According to Kline (2005), a comparative fit index of more than .9 and an RMSEA of less than .08 (with an upper limit of less than .1) signal acceptable fit. A comparative fit index of more than .95 and an RMSEA of less than .05 suggest close fit (Hu & Bentler, 1999). The sample size and the complexity of the factor structure influence the RMSEA (Iacobucci, 2010). As such, this goodness of fit statistic was not a main criterion.

After identifying the optimal factor structure, both a Rasch score and a sum-score index were calculated for each domain. The Rasch score has two advantages over the sum-score index, or the simple average of items. First, the Rasch scores are on an interval scale, whereas the sum-score indices are on a numerical scale. Second, the Rasch scores are less skewed (and therefore closer to a normal distribution) than the sum-score indices. The correlation between Rasch scores and sum-score indices was found to be more than .95. To simplify the presentation of the results, the report presents sum-score indices.

Finally, the study team examined reliability for domains in the retained factor structure using two approaches: an internal consistency approach in which Cronbach's alpha was calculated and a Rasch modeling approach in which a person reliability index was calculated. Reliability numbers higher than .9 indicate that a factor has high reliability, and numbers higher than .7 suggest that a factor has acceptable reliability (Kline, 2013). Factors with more items are more likely to have higher reliability.

Multiple-group confirmatory factor analyses and differential item functioning analysis (research question 2)

Multiple-group confirmatory factor analyses were conducted to determine whether factor invariance was achieved across subgroups. Factor invariance across subgroups was established step by step: first configural

invariance, then metric invariance (weak factor invariance), and finally scalar invariance (strong factor invariance). At each step a stronger case of invariance was established when there was no sizable difference in the model fit statistics (Muthén & Asparouhov, 2002), indicated by either the nonsignificance of the chi-square difference test (Marsh & Grayson, 1994) or the minimal shift in comparative fit index and RMSEA (change in comparative fit index of less than .01 and change in RMSEA of less than .015; Cheung & Rensvold, 2002).

Differential item functioning analysis was used to test whether any of the 14 items on the Kindergarten Individual Development Survey functioned differently for one subgroup of children relative to another. Log odds ratios were used to screen the differential item functioning, which did not require an adequate model fit assumption according to item response theory methods (Cohen, Kim, & Baker, 1993; Liu & Agresti, 1996; Penfield & Algina, 2003). The threshold used for these analyses was a log odds ratio of 0.64 (Zwick, Thayer, & Lewis, 1999).[3] Log odds ratios below 0.43 indicate negligible differential item functioning, log odds ratios of 0.43–0.64 indicate small to moderate differential item functioning, and log odds ratios above 0.64 indicate moderate to large differential item functioning. An absolute value above 0.64 for the log odds ratio was the criterion for potential bias in an item to avoid a high rate of type I error (Penfield, Alvarez, & Lee, 2009; Zieky, 1993). Log odds ratios larger than 0.64 were flagged for potential bias, with negative numbers indicating potential bias favoring the focal group, and positive numbers indicating potential bias favoring the reference group.

Multilevel modeling (research questions 4 and 5)

Differences in children's knowledge and skills at kindergarten entry across key subgroups. Multilevel modeling was used to estimate the differences between child subgroups on measures of learning and social skills and measures of academic knowledge and skills. Children (level 1) were nested in classrooms (level 2) and districts (level 3) to account for clustering of standard errors. The multilevel model took the following form:

$$Y_{ijk} = \pi_{0jk} + \pi_{1jk}(focal\ subgroup) + e_{ijk} \qquad \text{(B1)}$$

$$\pi_{0jk} = \beta_{00k} + r_{0jk} \qquad \text{(B2)}$$

$$\pi_{1jk} = \beta_{10k} + r_{1jk} \qquad \text{(B3)}$$

$$\beta_{00k} = \gamma_{000} + \mu_{00k} \qquad \text{(B4)}$$

$$\beta_{10k} = \gamma_{100} + \mu_{10k} \qquad \text{(B5)}$$

In equation B1, Y_{ijk} is one of the two continuous outcome domain measures (the mean of individual item scores of learning and social skills or academic knowledge and skills) for child *i* in classroom *j* and district *k*; π_{0jk} is the intercept, which here can be interpreted as the classroom mean for classroom *j* in district *k* for children in the reference group (for example, if the focal subgroup is boys, the reference group would be girls); π_{1jk} is the difference in the average scores of the subgroup and the reference group in classroom *j* and district *k*; and e_{ijk} is the child-level error term for child *i* in classroom *j* and district *k*. Notably, equation B1 includes the indicator (if represented by a single binary variable) or indicators (if represented by a set of dummy variables) for one subgroup at a time rather than including all the demographic characteristics as control variables.

This specification will produce estimates of gaps between subgroups. Notably, these estimates do not account for economic disparities or other observable differences between subgroups. The study team made this analytic choice in order to meet the Midwest Early Childhood Education Research Alliance's request for information on

---

[3] Log odds ratios range from negative infinity to positive infinity. A value of 0.64 means that the odds for one group to have a favorable answer on an item is about twice as much as the odds for the other group. For example, if the probability of having a favorable rating for one group is 50 percent, the probability for the other group would be 67 percent, representing a large discrepancy between the two groups.

overall disparities between subgroups rather than on gaps net of other characteristics. Both the intercept and the slope on *focal subgroup* are allowed to vary across schools and districts, as shown in equations B2–B5. In equation B2, β_{00k} is the classroom-level intercept, which can be interpreted as the average of classroom average scores, and r_{0jk} represents the classroom-level error term. Equation B3 contains β_{10k}, the classroom-level average achievement gap, and r_{1jk}, which is the classroom-level error term in this achievement gap across teachers. In equation B4, the classroom-level average for the reference group is a function of the district-level average score, γ_{000}, and μ_{00k}, the district-level error term. Finally, in equation B5, the classroom-level average difference between the reference group and the subgroup is modeled as a function of the district-level average achievement gap and μ_{10k}, the district-level error in this average achievement gap.

Multilevel models were used to provide estimates of skill gaps. To ease interpretation, skill gaps were translated into instruction days. This translation was performed by dividing the gap by the mean effect size for annual achievement gain on reading during the kindergarten year across seven nationally normed tests, 1.52 standard deviations, and multiplying the result by 180 days.

Association between school-level poverty and knowledge and skills at kindergarten entry. To answer research question 5, the study team included school-level poverty—the percentage of children in a school who were eligible for the national school lunch program—in stepwise multilevel models. To estimate the association between school-level poverty and children's scores on the domains of the Kindergarten Individual Development Survey, the study team used a multilevel regression with children nested in schools nested in districts (step 1):

$$Y_{ijk} = \pi_{0jk} + e_{ijk} \tag{B6}$$

$$\pi_{0jk} = \beta_{00k} + \beta_{01k}(PercentNSLP) + r_{0jk} \tag{B7}$$

$$\beta_{00k} = \gamma_{000} + \mu_{00k} \tag{B8}$$

$$\beta_{10k} = \gamma_{100} + \mu_{10k} \tag{B9}$$

In equation B6, Y_{ijk} is one of the two continuous domain scores (learning and social skills or academic knowledge and skills) for child i in school j in district k; π_{0jk} is the intercept, which can be interpreted as the school average score within school j and district k; e_{ijk} is the child-level error term; β_{00k} is the school-level intercept; β_{01k} is the association between percentage of children in a school who were eligible for the national school lunch program and the Kindergarten Individual Development Survey subscale score; and r_{0jk} is the school-level error term. Both the school-level intercept and the school-level coefficient on the percentage of children eligible for the national school lunch program were allowed to vary across districts, as shown in equations B8 and B9.

Next, the study team included all child-level characteristics except for the child-level indicator of national school lunch program eligibility in the multilevel model (step 2):

$$Y_{ijk} = \pi_{0jk} + \boldsymbol{\pi_{1jk}}(Covariates\ Except\ NSLP) + e_{ijk} \tag{B10}$$

$$\pi_{0jk} = \beta_{00k} + \beta_{01k}(PercentNSLP) + r_{0jk} \tag{B11}$$

$$\beta_{00k} = \gamma_{000} + \mu_{00k} \tag{B12}$$

$$\beta_{10k} = \gamma_{100} + \mu_{10k} \tag{B13}$$

The model is the same as the one listed for step 1 except for the inclusion of a vector of child-level covariates indicated by $\boldsymbol{\pi_{1jk}}$ that represents the coefficients on covariates for age as of September 1, 2017, gender, race/ethnicity, English learner status, and individualized education program status. The slopes on the child-level covariates were not allowed to vary across districts. Because the percentage of children in a school who were

eligible for the national school lunch program is an aggregate of individual children's national school lunch program eligibility within the school, including the child-level indicator of national school lunch program eligibility may explain most of the variances from the model in step 1. Therefore, in step 2 the study team included in the model child-level characteristics described in the prior section, excluding national school lunch program eligibility. Finally, the study team estimated the following multilevel model that included all child-level characteristics, including child-level national school lunch program eligibility (step 3):

$$Y_{ijk} = \pi_{0jk} + \pi_{1jk}(Covariates\ Including\ NSLP) + e_{ijk} \qquad \text{(B14)}$$

$$\pi_{0jk} = \beta_{00k} + \beta_{01k}(PercentNSLP) + r_{0jk} \qquad \text{(B15)}$$

$$\beta_{00k} = \gamma_{000} + \mu_{00k} \qquad \text{(B16)}$$

$$\beta_{10k} = \gamma_{100} + \mu_{10k} \qquad \text{(B17)}$$

The model is the same as the one listed for step 2 except that the vector of child-level covariates in equation B14 now includes a child-level indicator for national school lunch program eligibility in addition to the covariates for age as of September 1, 2017, gender, race/ethnicity, English learner status, and individualized education program status. The slopes on the child-level covariates were not allowed to vary across districts.

Teacher and principal interviews (research questions 5 and 6)

Teachers and principal interview sample. In fall 2017 the Illinois State Board of Education used regional coaches to train teachers in nine regions of the state on how to administer the Kindergarten Individual Development Survey. The study team asked the regional coaches to nominate one or more schools in their region to participate in the interviews. The regional coaches recommended teachers or principals based on familiarity with them gained through conducting trainings and fielding questions from teachers and principals. Five regional coaches nominated one or more schools in their region to participate in the interviews. The study team interviewed one kindergarten teacher and one principal in each of nine schools individually. Regional coaches may have nominated educators who they considered to have ample experience in administering the survey, as shown by the fact that five of the nine teacher participants had administered the survey for at least one year prior to the statewide implementation. As such, the findings here are not representative of all educators' experiences in the state. See table B2 for a description of the sample.

Table B2. Description of the interview sample of kindergarten teachers and school principals related to the fall 2017 administration of the Kindergarten Individual Development Survey in Illinois

Teacher–principal pair[a]	First year that teacher administered the survey[b]	Number of measures collected[c]	Teacher assessed children one-on-one	Teacher had elementary teaching certificate	Teacher had an early childhood teaching certificate	Teacher had more than three years of experience	Principal taught at the elementary school level previously
1	2013	21		✓	✓	✓	✓
2	2014	31		✓		✓	✓
3	2014	14			✓	✓	✓
4	2014	14		✓		✓	✓
5	2016	14		✓		✓	✓
6	2017	14				✓	✓
7	2017	14		✓			✓
8	2017	14	✓	✓	✓	✓	✓
9	2017	14	✓	✓	✓	✓	✓

a. A pair of educators, consisting of one kindergarten teacher and one principal, from each school participated in the interviews.
b. Teachers who administered the survey prior to 2017 participated in the pilot.
c. The pilot instrument contained 55 measures; in 2015, based on feedback from the pilot, the Illinois State Board of Education requested that the WestEd Center for Child and Family Studies produce a shorter version of the instrument that could still reliably and validly measure children's social and emotional, language and literacy, and math skills at kindergarten entry. That version contained 14 items. The Illinois State Board of Education required teachers to rate each child on those 14 items; teachers could also choose to rate children on the remaining 41 items.
Source: Authors' analysis of data from interviews with nine pairs of Illinois teachers and principals in 2017.

Teachers in the sample had varied teaching background and differences in prior experience with the Kindergarten Individual Development Survey. Teachers had between 0 and 25 years of experience teaching kindergarten. Seven teachers had an elementary teaching certificate, one teacher was in the process of obtaining an elementary teaching certificate, and four teachers had an early childhood teaching certificate. Of the nine teachers interviewed, four had completed the Kindergarten Individual Development Survey for the first time in fall 2017, and five had begun administering the survey at various stages during the pilot and had between one and four years of experience with it. All teachers had collected the 14 required measures; two of the teachers who had participated during the pilot had collected data on and rated more than the 14 required measures.

Principals in the sample had a range of prior administrative and teaching experience. The total time that the principals had served in administration roles prior to 2017 ranged from 0 to 13 years. Prior classroom teaching experience ranged from 3 to 18 years, spent at the elementary or middle school levels. Two principals had taught special education. None of the principals had taught in kindergarten classrooms, although one principal had taught grades 1 and 2, and another had been a reading interventionist for K–5 children.

Teacher and principal interview protocols. The interviews included questions about educators' teaching or administration background, impressions of the Kindergarten Individual Development Survey, experiences administering the survey with all children (including English learner children and children with disabilities), data entry and data use, and suggestions for improvement. See boxes B1 and B2 for the interview protocols.

Box B1. Interview protocol for teachers

Hello, I am _____ with the Regional Educational Laboratory Midwest.

Thank you for deciding to participate in this interview. The purpose of this interview is to provide the state with feedback from teachers about barriers they encountered in administering the KIDS, in addition to suggestions on how the KIDS might be improved in future years. The information collected from the interviews will be reported back to the state, but no identifying information about your name, school, or district will be attached to the reports. The interview will last approximately 40 minutes.

Your participation in this interview is completely voluntary, and any information you provide will be confidential. As a reminder, this conversation is being recorded, and after the transcription of our conversation, the digital file will be destroyed. All identifying information will be kept confidential. Do I have permission to record you? [Note: If the respondent wishes not to be recorded, take notes but do not proceed with recording. If the respondent consents to being recorded, please record the interview.] Do you have any questions before we begin?

1. Background

    a. How many years have you been teaching? How long have you taught kindergarten?

    b. Do you have an early childhood or elementary teaching license?

2. Impressions of the Kindergarten Individual Development Survey (KIDS)

    a. Based on the training you received, what is the purpose of KIDS? How can these data be used?

    b. How did you feel about being asked to assess all of your children using KIDS? Did you have any concerns prior to administration?

    c. Tell me about your experience administering KIDS in your classroom.

        i. Did it take as about the length of time that you expected?

        ii. Did you have enough information to enter ratings for all children?

        iii. Were you able to collect observational data during normal instruction?

    d. Did you encounter any difficulties while collecting data for KIDS? (If the respondent does not mention the following topics in his or her answer, probe about time, burden, and observing children with special needs or dual language learners.)

    e. Did you encounter any difficulties in entering the data into the system? If so, please explain.

    f. What was it like to explain the KIDS results to parents? What challenges did you face in doing so?

    g. How does the administration of KIDS differ from the practices you typically use to assess children at the beginning of the school year? [if none, skip]

        i. Did it require more or less time entering data into the system?

        ii. In your opinion, did it result in more or less useful data?

        iii. Will you continue using the other practices you mentioned in addition to using KIDS?

3. Suggestions for improving KIDS

    a. Is there anything that the state (or your district or school) could have provided to make collecting data for KIDS easier? (If the respondent does not mention the following topics, probe about training, guidance on how to enter the data, and additional support from paraprofessionals).

    b. Is there anything that you learned from collecting data for KIDS that you think might be helpful to other teachers who will collect data for KIDS in the future?

4. Is there any additional information that you would like to share about your experience with KIDS this fall?

This question concludes the interview. Thank you for your time. If you have any questions about this study, please feel free to contact the Principal Investigator for this study, Jill Bowdon, Ph.D.

The interview protocol of the principals focused on the principal's role in supporting the administration of the Kindergarten Individual Development Survey. The full protocol is shown below in box B2.

Box B2. Interview protocol for principals

Hello, I am _____ with the Regional Educational Laboratory Midwest.

Thank you for deciding to participate in this interview. The purpose of this interview is to provide the state with feedback from school principals about barriers they encountered in overseeing the administration of the KIDS, in addition to suggestions on how KIDS might be improved in future years. The information collected from the interviews will be reported back to the state but no identifying information about your name, school, or district will be attached to the reports. The interview will last approximately 40 minutes.

Your participation in this interview is completely voluntary, and any information you provide will be confidential. As a reminder, this conversation is being recorded, and after the transcription of our conversation, the digital file will be destroyed. All identifying information will be kept confidential. Do I have permission to record you? [Note: If the respondent wishes not to be recorded, take notes, but do not proceed with recording. If the respondent consents to being recorded, please record the interview.] Do you have any questions before we begin?

1. Background

    a. How long have you been a school principal at this school? How many years have you been a school principal total, including years served at other schools?

    b. Did you teach prior to becoming a principal? If so, what grades and for how long?

2. The principal's role in the administration and use of the Kindergarten Individual Development Survey (KIDS)

    a. Administration

        i. What do you see as your role in the administration of KIDS?

        ii. Did any of your teachers express concerns to you prior to administering KIDS? If so, what were they and how did you address those concerns?

        iii. To the best of your knowledge, did the teachers at your school share any resources or guidance with each other about how to administer KIDS?

        iv. What type of feedback did you hear from teachers about their administration of KIDS this fall? Did they perceive any challenges in administering KIDS?

    b. Use

        i. How would you, as an administrator, like to use the KIDS data?

        ii. Did you hear any feedback from teachers about their use of KIDS this fall?

3. Feedback on support for administering and using the KIDS data

    a. Is there anything that the state (or your district or school) could have provided to make collecting data for KIDS easier? (If the respondent does not mention the following topics, probe about training, guidance on how to enter the data, and additional support from paraprofessionals).

    b. What suggestions do you have on how to make the administration of KIDS smoother in years to come?

4. Is there any additional information that you would like to share about KIDS prior to the conclusion of the interview?

This question concludes the interview. Thank you for your time. If you have any questions about this study, please feel free to contact the Principal Investigator for this study, Jill Bowdon, Ph.D.

Interview data analysis. The study team conducted a 30-minute semistructured interview with each of the 18 participants. The study team qualitatively analyzed the interviews by first identifying themes related to challenges that the literature suggests teachers may face in administering the survey. These themes included finding time to assess each child, using technology-based systems to enter data, and using the survey with English learner children or children with an individualized education program. Additional codes were created to include emergent themes related to facilitators and barriers to implementation. Suggestions for improvement also were coded based on emergent themes. The study team met to reconcile findings and determine final themes across individuals and teacher-principal pairs.

References

Cheung, G. W., & Rensvold, R. B. (2002). Evaluating goodness-of-fit indexes for testing measurement invariance. *Structural Equation Modeling*, *9*(2), 233–255. https://eric.ed.gov/?id=EJ651463.

Cohen, A. S., Kim, S.-H., & Baker, F. B. (1993). Detection of differential item functioning in the graded response model. *Applied Psychological Measurement*, *17*(4), 335–350. https://eric.ed.gov/?id=EJ485766.

Fabrigar, L. R., Wegener, D. T., MacCallum, R. C., & Strahan, E. J. (1999). Evaluating the use of exploratory factor analysis in psychological research. *Psychological Methods*, *4*(3), 272–299.

Gadermann, A. M., Guhn, M., & Zumbo, B. D. (2012). Estimating ordinal reliability for Likert-type and ordinal item response data: A conceptual, empirical, and practical guide. *Practical Assessment, Research & Evaluation*, *17*(3), 1–13. https://eric.ed.gov/?id=EJ977577.

Hu, L., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling, 6*(1), 1–55. http://eric.ed.gov/?id=EJ576570.

Iacobucci, D. (2010). Structural equations modeling: Fit indices, sample size, and advanced topics. *Journal of Consumer Psychology*, *20*(1), 90–98.

Kline, P. (2013). *Handbook of psychological testing*. London, UK: Routledge.

Kline, R. B. (2005). *Principles and practice of structural equation modeling* (2nd ed.). New York, NY: Guilford Press.

Liu, I.-M., & Agresti, A. (1996). Mantel-Haenszel-type inference for cumulative odds ratios with a stratified ordinal response. *Biometrics*, *52*(4), 1223–1234.

Marsh, H. W., & Grayson, D. (1994). Longitudinal confirmatory factor analysis: Common, time-specific, item-specific, and residual-error components of variance. *Structural Equation Modeling, 1*(2), 116–145.

Muthén, B., & Asparouhov, T. (2002). *Latent variable analysis with categorical outcomes: Multiple-group and growth modeling in Mplus*. *Mplus web notes, 4*(5): 1–22. Retrieved October 21, 2018, from http://www.statmodel.com/download/webnotes/CatMGLong.pdf.

Penfield, R. D., & Algina, J. (2003). Applying the Liu-Agresti estimator of the cumulative common odds ratio to DIF detection in polytomous items. *Journal of Educational Measurement, 40*(4), 353–370.

Penfield, R. D., Alvarez, K., & Lee, O. (2009). Using a taxonomy of differential step functioning form to improve the interpretation of DIF in polytomous items: An illustration. *Applied Measurement in Education, 22*(1), 61–78*. https://eric.ed.gov/?id=EJ822892.

Stevens, J. P. (2009). *Applied multivariate statistics for the social sciences* (5th ed.). New York, NY: Routledge.

Thompson, B. (2004). *Exploratory and confirmatory factor analysis: Understanding concepts and applications.* Washington, DC: American Psychological Association.

Yates, A. (1988). *Multivariate exploratory data analysis: A perspective on exploratory factor analysis.* Albany, NY: State University of New York Press.

Zieky, M. (1993). Practical questions in the use of DIF statistics in test development. In P. W. Holland & H. Wainer (Eds.), *Differential item functioning* (pp. 337–347). Hillsdale, NJ: Lawrence Erlbaum.

Zwick, R., Thayer, D. T., & Lewis, C. (1999). An empirical Bayes approach to Mantel-Haenszel DIF analysis. *Journal of Educational Measurement, 36*(1), 1–28. https://eric.ed.gov/?id=EJ590611.

Appendix C. Supporting analyses

This appendix includes key supporting analyses for the psychometric and descriptive findings of the report. It includes the results of the exploratory and confirmatory factor analyses and the examination of the reliability of the two identified factors as well as findings related to the multigroup factor analyses and the differential item functioning analysis. It also contains histograms and results from multilevel models associated with the descriptive questions.

Two-factor structure

To answer research question 1, the study team conducted both exploratory and confirmatory factor analyses. Methodological details about these approaches are in appendix B. The eigenvalues from the exploratory factor analysis are pictured in a scree plot in figure C1. A scree plot can help determine the optimal number of factors to be extracted from the data. The graph indicates that no more than two factors should be retained, since factors with eigenvalues less than 1 are typically discarded.

Figure C1. Scree plot from exploratory factor analysis for the fall 2017 administration of the Kindergarten Individual Development Survey in Illinois

Eigenvalue
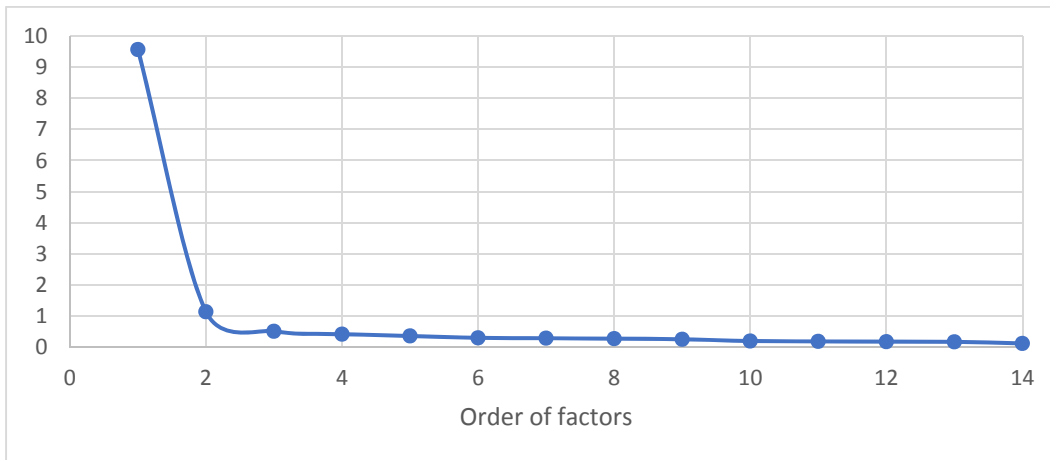
Note: The sample consisted of 56,844 children in 2,021 schools.
Source: Authors' analyses of fall 2017 Kindergarten Individual Development Survey data provided by the Illinois State Board of Education.

For the confirmatory factor analysis, goodness-of-fit statistics, such as comparative fit index and root mean square error of approximation, were used to determine model fit (table C1). Based on these analyses, the study team opted to use the two-factor solution it identified in exploratory factor analyses rather than the three-factor solution identified during the pilot phase and reported by the state.

Table C1. Detailed findings on the goodness of fit for different factor structures based on confirmatory factor analyses of the fall 2017 administration of the Kindergarten Individual Development Survey in Illinois

Number of factors	Comparative fix index	Root mean square error of approximation	Standardized root mean residual	Tucker Lewis index	Correlation between factors	Substantive description of identified factors
Factor structures identified by the study team						
1	.94	.17	.07	.93	na	• Overall readiness
2	.97	.13	.03	.96	.64	• Learning and social skills • Academic knowledge and skills
3	.99	.09	.02	.98	.42–.79	• Learning and social skills • Academic knowledge and skills • Not meaningful
4	.99	.08	.01	.99	.22–.80	• Learning and social skills • Literacy and language development • Cognition: mathematics • Not meaningful
Factor structure identified during the pilot phase and reported on by Illinois State Board of Education						
3	.96	.13	.03	.96	.58–.70	• Approaches to learning/social-emotional development • Language and literacy development • Cognition: mathematics

na is not applicable.

Note: The sample consisted of 56,844 children in 2,021 schools. The fit statistics for the three-factor model identified during the pilot phase and reported on by the Illinois State Board of Education are no better than the fit statistics for two-factor model from exploratory factor analyses, despite having more parameters. Moreover, the correlation between the approaches to learning/social-emotional development factor and the language and literacy development factor and the correlation between the language and literacy development factor and the cognition: mathematics factor are both .70, meaning that there is less discriminant validity.

Source: Authors' analyses of fall 2017 Kindergarten Individual Development Survey data provided by the Illinois State Board of Education.

The study team made this decision for three main reasons. First, the two-factor solution was supported by exploratory factor analysis with a majority of the whole state population. Second, the discriminant validity for the two-factor solution was better than that for the three-factor solution identified during the pilot phase and reported by the state. The .64 correlation is lower than the .70 correlation between the language and literacy development factor and the other two factors in the state-reported three-factor solution. Third, the fit statistics for the three-factor model identified during the pilot phase and reported by the Illinois State Board of Education are no better than the fit statistics for the two-factor model from exploratory factor analyses, despite having more parameters.

Reliability of the two factors

The study team examined the reliability of the two factors. Reliability of 0.9 or higher is considered excellent (Ferketich, 1991; Zinbarg et al., 2005). The learning and social skills factor had an estimated internal consistency of .94, a Rasch person reliability of .90, and a McDonald's omega of 0.95. The academic knowledge and skills factor had an estimated internal consistency of .92, a Rasch person reliability of .90, and a McDonald's Omega of 0.94. See table 1 in the main text for the factor loadings for the two domains.

Invariance of factors across child subgroups

The study team used multigroup factor analyses to test the invariance of factors across subgroups. Although the chi-square test is always significant (a finding that may be expected in an analysis with large samples, as is the case with this study), the shift in both fit statistics is minimal, thus establishing strong factor invariance (table C2).

Table C2. Summary of factor invariance results for the fall 2017 administration of the Kindergarten Individual Development Survey in Illinois

Grouping variable	Configural invariance established	Chi-square test nonsignificant	Shift in comparative fit index ≤ .01	Shift in root mean square error of approximation ≤ .015	Strong factor invariance achieved
Gender	Yes	No	Yes	Yes	Yes
English learner status	Yes	No	Yes	Yes	Yes
Eligibility for the national school lunch program	Yes	No	Yes	Yes	Yes
Eligibility for an individualized education program	Yes	No	Yes	Yes	Yes
Race/ethnicity	Yes	No	Yes	Yes	Yes

Note: The sample consisted of 113,716 children in 2,027 schools.

Source: Authors' analyses of fall 2017 Kindergarten Individual Development Survey data provided by the Illinois State Board of Education.

Testing for differential item functioning across subgroups

The study team used differential item functioning analysis to understand whether any of the 14 required items on the Kindergarten Individual Development Survey functioned differently for one subgroup of children relative to another. Differential item functioning signals the possibility that an item does not measure child ability accurately for a subgroup. However, in this case, differential item functioning analyses cannot determine whether the item is functioning differently because of measure or rater bias, real differences in skills for a child subgroup, or random errors. Additional analyses, such as a review by a panel of child development and culture experts, would be needed to determine whether the differences indicate bias or reflect true differences in abilities between the subgroups for these items or just random errors.

Three of the 140 contrasts in the analysis indicated differential item functioning (tables C3–C5). Differential item functioning occurs when two subgroups of children have identical scores for a domain—for example, learning and social skills or academic knowledge and skills—but differ in the likelihood of having a high rating for a particular item that is part of that score in that domain. In table C3 each of the child subgroups listed in the third column had identical scores for the domain listed in the second column. However, the favored group (listed in the fourth column) was more likely to receive a higher rating for the item listed in the first column. For example, Asian children generally received higher ratings on letter and word knowledge compared with Hispanic children who had equal scores for the academic knowledge and skills domain.

Table C3. Items from the fall 2017 administration of the Kindergarten Individual Development Survey in Illinois that functioned differently for one subgroup relative to another

Item	Domain	Subgroup comparison	Favored subgroup
Communication and use of language (expressive)	Learning and social skills	English learner children and non–English learner children	Non–English learner children
Engagement and persistence	Learning and social skills	Asian children and Black children	Asian children
Letter and word knowledge	Academic knowledge and skills	Asian children and Hispanic children	Asian children

Note: The sample for the first comparison consisted of 32,823 children in 1,651 schools. The sample for the second comparison consisted of 23,800 children in 1,601 schools. The sample for the third comparison consisted of 113,716 children in 2,027 schools.

Source: Authors' analyses of fall 2017 Kindergarten Individual Development Survey data provided by the Illinois State Board of Education.

More detailed differential item functioning estimates in log ratio units for each item and each comparison are presented in tables C4 and C5, by domain. Each row represents an item within the domain, and each column stands for a subgroup comparison.

Table C4. Differential item functioning summary for the learning and social skills domain of the fall 2017 administration of the Kindergarten Individual Development Survey in Illinois

Item[a]	Focal subgroup (contrast in log odds estimates)									
	Female[b] (n = 113,716)	English learner children (n = 113,716)	Children eligible for the national school lunch program (n = 113,716)	Children eligible for an individualized education program (n = 113,716)	Hispanic children (White children as reference group) (n = 84,567)	Asian children (White children as reference group) (n = 64,094)	Black children (White children as reference group) (n = 75,444)	Hispanic children (Black children as reference group) (n = 44,173)	Asian children (Black children as reference group) (n = 23,800)	Hispanic children (Asian children as reference group) (n = 32,823)
Curiosity and initiative in learning (ALT_REG 1)	0.391	−0.047	0.143	−0.274	0.068	−0.222	0.072	−0.008	−0.287	0.281
Self-control of feelings and behavior (ALT_REG 2)	−0.268	−0.302	−0.098	−0.053	−0.243	−0.24	0.014	−0.25	−0.256	0.023
Engagement and persistence (ALT_REG 3)	−0.078	−0.478	−0.027	−0.041	−0.268	−0.595	0.068	−0.332	−0.67	0.361
Relationships and social interactions with familiar adults (SED 3)	−0.05	−0.074	−0.229	−0.376	−0.118	0.218	−0.192	0.077	0.413	−0.349
Relationships and social interactions with peers (SED 4)	−0.234	−0.226	−0.17	0.041	−0.263	0.012	−0.073	−0.192	0.078	−0.253
Communication and use of language (expressive) (LLD 3)	0.12	0.649	0.271	0.38	0.517	0.46	0.139	0.369	0.335	0.009
Reciprocal communication and conversation (LLD 4)	0.162	0.473	0.112	0.305	0.306	0.347	−0.035	0.351	0.38	−0.071

Note: Items with a log odds ratios larger than 0.64 were flagged for differential item functioning. Each column represents a set of differential item functioning analyses. The column header represents the focal group, and the rest of the children in the sample form the reference group. Negative numbers indicate potential bias favoring the focal group, and positive numbers indicate potential bias favoring the reference group.

a. Text in parentheses refers to Kindergarten Individual Development Survey item numbers.

b. Male children is the reference group.

Source: Authors' analyses of fall 2017 Kindergarten Individual Development Survey data provided by the Illinois State Board of Education.

Table C5. Differential item functioning summary for the academic knowledge and skills domain of the fall 2017 administration of the Kindergarten Individual Development Survey in Illinois

Item[a]	Focal subgroup (contrast in log odds estimates)									
	Female children (n = 113,716)	English learner children (n = 113,716)	Children eligible for the national school lunch program (n = 113,716)	Children eligible for an individualized education program (n = 113,716)	Hispanic children (White children as reference group) (n = 84,567)	Asian children (White children as reference group) (n = 64,094)	Black children (White children as reference group) (n = 75,444)	Hispanic children (Black children as reference group) (n = 44,173)	Asian children (Black children as reference group) (n = 23,800)	Hispanic children (Asian children as reference group) (n = 32,823)
Comprehension of age-appropriate text (LLD 6)	−0.178	0.167	−0.155	0.383	−0.025	0.562	−0.059	0.039	0.607	−0.562
Phonological awareness (LLD 8)	−0.16	0.004	−0.061	0.215	−0.049	−0.072	−0.205	0.161	0.129	0.002
Letter and word knowledge (LLD 9)	−0.037	0.029	0.132	−0.294	0.137	−0.57	−0.259	0.377	−0.301	0.689
Classification (COG MATH 1)	−0.054	−0.049	−0.093	−0.023	−0.123	0.423	0.25	−0.372	0.18	−0.533
Number sense of quantity (COG MATH 2)	0.181	−0.141	0.039	−0.052	−0.023	−0.276	0.018	−0.039	−0.279	0.241
Number sense of math operations (COG MATH 3)	0.185	−0.247	0.152	0.005	−0.081	−0.321	0.217	−0.296	−0.531	0.259
Shapes (COG MATH 6)	0.115	0.206	−0.02	−0.158	0.131	0.229	0.107	0.017	0.131	−0.126

Note: Items with a log odds ratios larger than 0.64 were flagged for differential item functioning. Each column represents a set of differential item functioning analyses. The column header represents the focal group, and the rest of the children in the sample form the reference group. Negative numbers indicate potential bias favoring the focal group, and positive numbers indicate potential bias favoring the reference group.

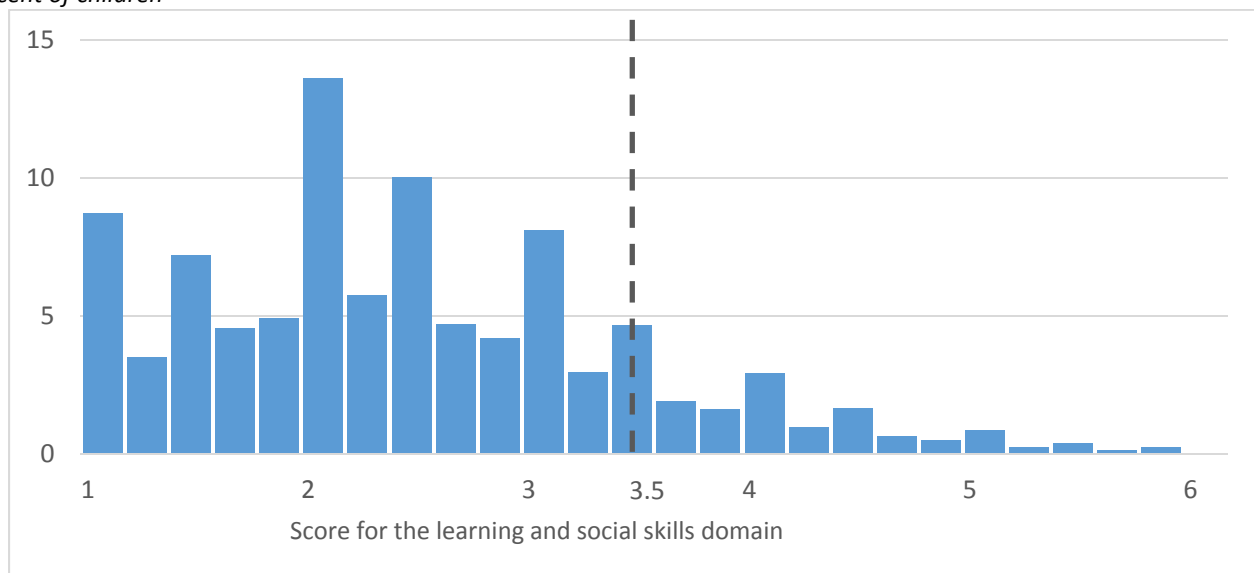a. Text in parentheses refers to Kindergarten Individual Development Survey item numbers.

Source: Authors' analyses of fall 2017 Kindergarten Individual Development Survey data provided by the Illinois State Board of Education.

Distributions of scores on the two identified scales in the 2017 Kindergarten Individual Survey data

The study team used histograms to understand the distributions of scales found in the 2017 Kindergarten Individual Development Survey data. A score represents a combination of ratings from several individual items on the Kindergarten Individual Development Survey. Teachers rate a child's skills on an individual item using the qualitative categories described in appendix A (building–earlier, building–middle, building–later, integrating–earlier, integrating–middle, and integrating–later). These rating categories were translated into a numerical scale and then combined to form scores for the learning and social skills domain and the academic knowledge and skills domain. Most children in the sample had a score on the lower end of the six-point scales for both the learning and social skills domain and the academic knowledge and skills domain (figures C2 and C3). Relative to the scale's midpoint (3.5), about 88 percent of children had a lower average score for the learning and social skills domain, and about 85 percent had a lower average score for the academic knowledge and skills domain.

Figure C2. Distribution of scores for the learning and social skills domain of the fall 2017 administration of the Kindergarten Individual Development Survey in Illinois
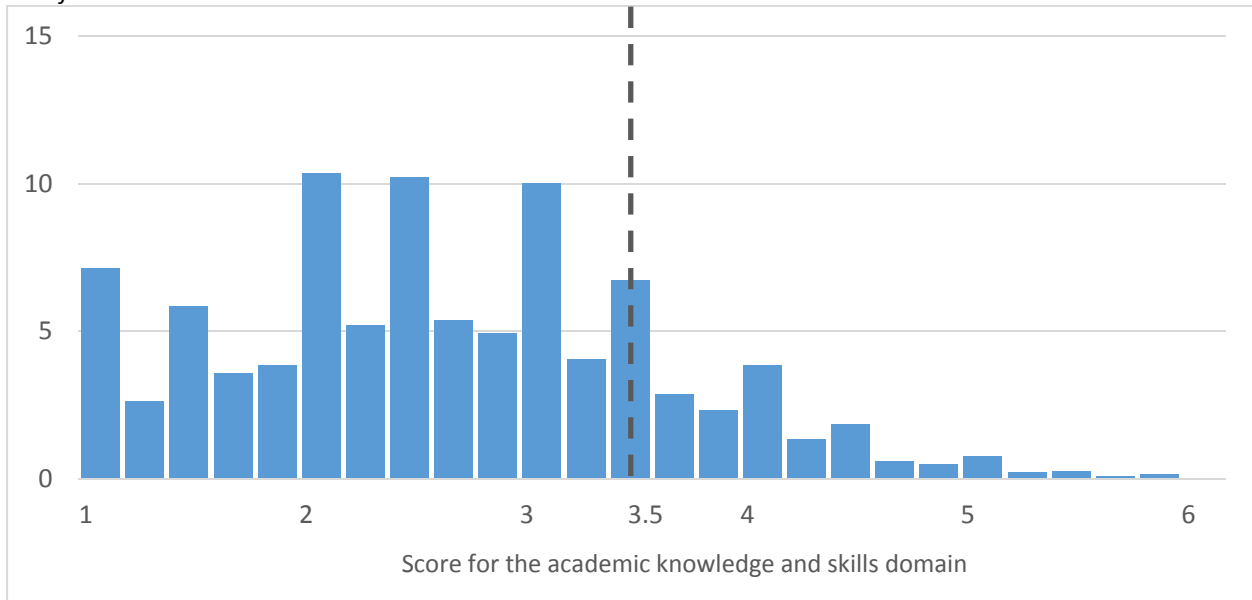
Percent of children

Note: The sample consisted of 113,716 children in 2,039 schools.
Source: Authors' analyses of fall 2017 Kindergarten Individual Development Survey data provided by the Illinois State Board of Education.

Figure C3. Distribution of scores for the academic knowledge and skills domain of the fall 2017 administration of the Kindergarten Individual Development Survey in Illinois

Percent of children

Note: The sample consisted of 113,716 children in 2,039 schools.
Source: Authors' analyses of fall 2017 Kindergarten Individual Development Survey data provided by the Illinois State Board of Education.

Associations between school-level poverty and scores on the two identified factors in the 2017 Kindergarten Individual Development Survey data

This section contains detailed findings related to research question 5. More details about the methods used to generate these findings can be found in appendix B. The association between school-level poverty (the percentage of children in a school who were eligible for the national school lunch program) and child-level scores for the learning and social skills domain was significant in models that did not account for child-level eligibility for the national school lunch program (table C6 and figure C4). Model 1 includes only a school-level indicator of the concentration of poverty, as operationalized by the percentage of children in the school who were eligible for the national school lunch program. The coefficient on the fixed effect for school-level poverty is –0.5 and is statistically significant ($p < .001$). Model 2 is the same as model 1 but adds child-level control variables for age as of September 1, 2017, gender, race/ethnicity, English learner status, and eligibility for an individualized education program. In model 2 the coefficient on school-level poverty is statistically significant ($p < .001$). Finally, model 3 is the same as model 2 but adds a child-level control variable for child eligibility for the national school lunch program. In model 3 the school-level poverty indicator is no longer associated with child scores for the learning and social skills domain.

The association between school-level poverty and child-level scores for the academic knowledge and skills domain was significant in models that did not account for child-level eligibility for the national school lunch program (table C7 and figure C5). School-level poverty was still significantly associated with child-level scores after child-level control variables were added for child age as of September 1, 2017, gender, English learner status, eligibility for an individualized education program, and race/ethnicity and after eligibility for the national school lunch program was added to those other control variables.

Table C6. Results of multilevel model testing of the association between school-level poverty and child-level scores for the learning and social skills domain of the fall 2017 administration of the Kindergarten Individual Development Survey in Illinois, by model

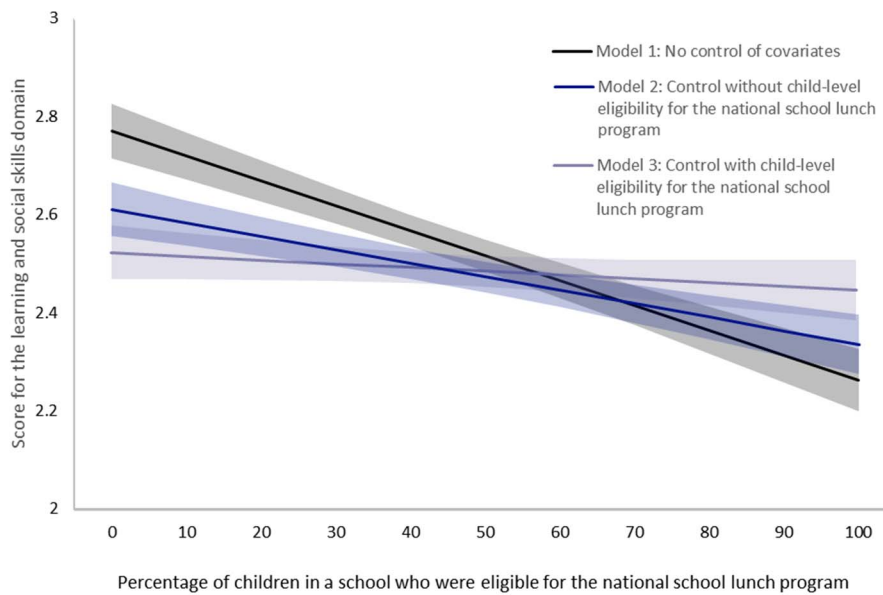| Fixed effect	Model 1			Model 2			Model 3		
	Estimate	Standard error	*p*-value	Estimate	Standard error	*p*-value	Estimate	Standard error	*p*-value
Intercept	2.77	0.03	<.001	0.43	0.05	<.001	0.41	0.05	<.001
School-level poverty[a]	−0.50	0.05	<.001	−0.28	0.05	<.001	−0.07	0.05	.14
Child-level poverty[b]	na	na	na	na	na	na	−0.28	0.01	<.001

na is not applicable.

Note: Estimates were generated using multilevel models with 113,716 children nested in 2,000 schools in 736 districts. Model 1 does not include any child-level covariates or school-level variables other than school-level poverty. Model 2 includes five child-level covariates (age as of September 1, 2017, gender, race/ethnicity, English learner status, and eligibility for an individualized education program) and school-level poverty but no other school-level variables. Model 3 includes all the covariates in model 2 as well as child-level poverty.

a. The percentage of children in a school who were eligible for the national school lunch program. It is a continuous variable ranging from 0 percent to 100 percent.

b. A binary indicator of whether a child was eligible for the national school lunch program.

Source: Authors' analyses of fall 2017 Kindergarten Individual Development Survey data provided by the Illinois State Board of Education and data from U.S. Department of Education (2019).

Figure C4. Visualization of the association between school-level poverty and child-level scores for the learning and social skills domain of the fall 2017 administration of the Kindergarten Individual Development Survey in Illinois, by model

Note: The black line represents results from a multilevel model that contained no child-level controls. The dark blue line represents results from a multilevel model that controlled for child-level characteristics, excluding child-level national school lunch program eligibility. The light blue line represents results from a multilevel model that controlled for child-level characteristics, including child-level national school lunch program eligibility. The sample size was 113,716 children nested in 2,039 schools in 738 districts.

Source: Authors' analysis of fall 2017 Kindergarten Individual Development Survey data provided by the Illinois State Board of Education and data from U.S. Department of Education (2019).

Table C7. Results of multilevel model testing of the association between school-level poverty and child-level scores for the academic knowledge and skills domain of the fall 2017 administration of the Kindergarten Individual Development Survey in Illinois, by model

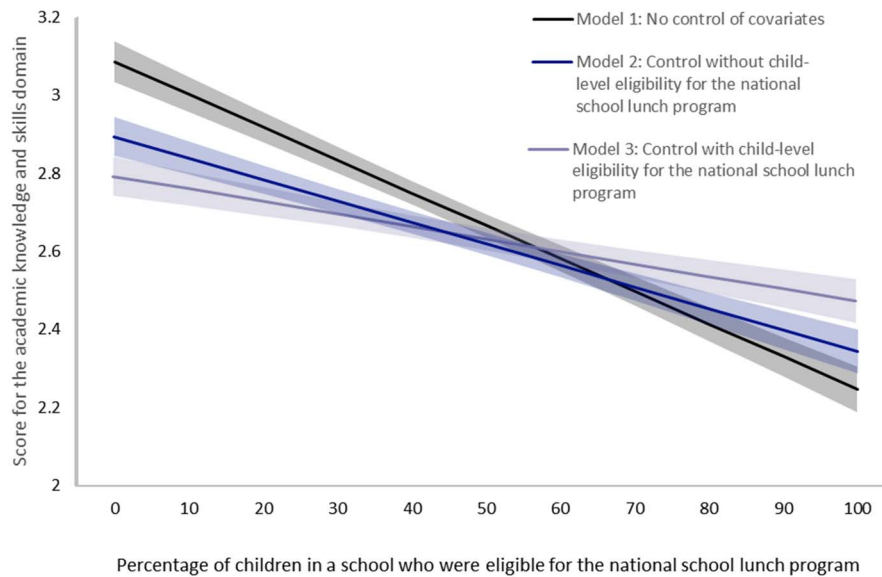|	Model 1			Model 2			Model 3		
Fixed effect	Estimate	Standard error	*p*-value	Estimate	Standard error	*p*-value	Estimate	Standard error	*p*-value
Intercept	3.09	0.03	<.001	0.53	0.05	<.001	0.51	0.05	<.001
School-level poverty[a]	−0.83	0.05	<.001	−0.55	0.05	<.001	−0.31	0.05	<.001
Child-level poverty[b]	na	na	na	na	na	na	−0.32	0.01	<.001

na is not applicable.

Note. Estimates were generated using multilevel models with 113,716 children nested in 2,000 schools in 736 districts. Model 1 does not include any child-level covariates or school-level variables other than school-level poverty. Model 2 includes five child-level covariates (age as of September 1, 2017, gender, race/ethnicity, English learner status, and eligibility for an individualized education program) and school-level poverty but no other school-level variables. Model 3 includes all the covariates in model 2 as well as child-level poverty.

a. The percentage of children in a school who were eligible for the national school lunch program. It is a continuous variable ranging from 0 percent to 100 percent.

b. A binary indicator of whether a child was eligible for the national school lunch program.

Source: Author's analyses of fall 2017 Kindergarten Individual Development Survey data provided by the Illinois State Board of Education and data from U.S. Department of Education (2019).

Figure C5. Visualization of the association between school-level poverty and child-level scores for the academic knowledge and skills domain of the fall 2017 administration of the Kindergarten Individual Development Survey in Illinois

Note: The black line represents results from a multilevel model that contained no child-level controls. The dark blue line represents results from a multilevel model that controlled for child-level characteristics, excluding child-level national school lunch program eligibility. The light blue line represents results from a multilevel model that controlled for child-level characteristics, including child-level national school lunch program eligibility. The sample size was 113,716 children nested in 2,039 schools in 738 districts.

Source: Authors' analyses of fall 2017 Kindergarten Individual Development Survey data provided by the Illinois State Board of Education and data from U.S. Department of Education (2019).

References

Ferketich, S. (1991). Focus on psychometrics: Aspects of item analysis. *Research in Nursing & Health, 14*(2), 165–168.

Zinbarg, R.E., Revelle, W., Yovel, I., & Li, W. (2005). Cronbach's alpha, Revelle's beta, McDonald's omega: Their relations with each other and two alternative conceptualizations of reliability. *Psychometrika, 70,* 123–133.