

Overview of the Middle Grades Longitudinal Study of 2017–18 (MGLS:2017)

Technical Report

March 23, 2023

Shannon L. Russell

Yuan Zhang

American Institutes for Research

Carolyn Ann Grim

Study Director

National Center for Education Statistics

U.S. Department of Education

Miguel Cardona
Secretary

Institute of Education Sciences

Mark Schneider
Director

National Center for Education Statistics

Peggy G. Carr
Commissioner

The National Center for Education Statistics (NCES) is the primary federal entity for collecting, analyzing, and reporting data related to education in the United States and other nations. It fulfills a congressional mandate to collect, collate, analyze, and report full and complete statistics on the condition of education in the United States; conduct and publish reports and specialized analyses of the meaning and significance of such statistics; assist state and local education agencies in improving their statistical systems; and review and report on education activities in foreign countries.

NCES activities are designed to address high-priority education data needs; provide consistent, reliable, complete, and accurate indicators of education status and trends; and report timely, useful, and high-quality data to the U.S. Department of Education, the Congress, the states, other education policymakers, practitioners, data users, and the general public. Unless specifically noted all information contained herein is in the public domain.

We strive to make our products available in a variety of formats and in language that is appropriate to a variety of audiences. You, as our customer, are the best judge of our success in communicating information effectively. If you have any comments or suggestions about this or any other NCES product or report, we would like to hear from you. Please direct your comments to

NCES, IES, U.S. Department of Education
Potomac Center Plaza (PCP)
550 12th Street, SW
Washington, DC 20202

March 2023

The NCES Home Page address is <https://nces.ed.gov>.

The NCES Publications and Products address is <https://nces.ed.gov/pubsearch>.

This publication is only available online. To download, view, and print the report as a PDF file, go to the NCES Publications and Products address shown above.

This report was prepared in part under Contract No. ED-IES-12-D-0002/0007 with the American Institutes for Research. Mention of trade names, commercial products, or organizations does not imply endorsement by the U.S. Government.

Suggested Citation

Russell, S.L., and Zhang, Y. (2023). *Overview of the Middle Grades Longitudinal Study of 2017-18 (MGLS:2017)* (NCES 2023-055). U.S. Department of Education, Washington, DC: National Center for Education Statistics. Retrieved [date] from <https://nces.ed.gov/pubsearch>.

Content Contact

Carolyn Ann Grim
(202) 245-7046
Carolyn.Grim@ed.gov

Acknowledgments

The study team would like to recognize the more than 19,000 students and their parents who participated in this study. The study team also thanks the teachers and administrators from the more than 750 schools across the United States for allowing NCES and its representatives to work with their students and for providing information about their students and schools. Special recognition is extended to the chief state school officers, district superintendents and staff, and private school officials for their assistance in promoting the study in their schools.

Contents

Acknowledgments	3
Overview	6
<i>Sample Design</i>	<i>6</i>
<i>Data Collection</i>	<i>7</i>
<i>Direct Cognitive Assessments</i>	<i>9</i>
Reading and Mathematics	10
Executive Function	11
<i>Student Survey.....</i>	<i>11</i>
<i>Response Rates and Nonresponse Bias Analysis.....</i>	<i>12</i>
<i>Data Reliability</i>	<i>18</i>
Nonsampling Errors	19
Sampling Errors and Weighting	19
<i>Available Data Files and Documentation</i>	<i>22</i>
References.....	23
Appendix A: Variables of Interest	24
<i>Student and Family Characteristics.....</i>	<i>25</i>
<i>School Characteristics</i>	<i>27</i>

List of Tables

Table 1. School response rates, by school type: School years 2017–18 and 2019–20.....	13
Table 2. Student assessment and parent survey overall response rates, by school type: School years 2017–18 and 2019–20.....	13
Table 3. Overall longitudinal response rates for student assessment and parent survey, by school type: School years 2017–18 and 2019–20.....	14
Table 4. Math teacher survey overall response rates, by school type: School years 2017–18 and 2019–20.....	15
Table 5. Special education teacher survey overall response rates, public school only: School years 2017–18 and 2019–20.....	15
Table 6. Overall school administrator survey response rates, by school type: School years 2017–18 and 2019–20.....	16
Table 7. Student, family, and school characteristics: Unweighted frequency and percentage distribution of the MGLS:2017 sample, by student, family, and school characteristics: Spring 2018 and spring 2020.....	17

List of Exhibits

Exhibit 1. MGLS:2017 base-year data collection (MS1) and follow-up (MS2) student session components for in-school and out-of-school participants.....	8
Exhibit 2. MGLS:2017 school and student analysis weights.....	20

Overview

The Middle Grades Longitudinal Study of 2017-18 (MGLS:2017) was the first study conducted by the National Center for Education Statistics (NCES) to follow a national sample of students from sixth to eighth grade. MGLS:2017 collected data on this cohort of students' academic, social, and interpersonal status and growth as they transitioned from elementary school and then prepared to enter high school.

The Main Study Base Year (MS1) data collection occurred in the second half, or spring semester, of the 2017-18 school year. The Main Study follow-up (MS2) data collection took place 2 years later during the spring semester of the 2019-20 school year. Most of the students were in eighth grade during MS2.¹ Information was collected from students, their parents or guardians, their teachers, and their school administrators. The data collection included direct and indirect assessments of mathematics, reading, and executive function, as well as indirect assessments of socioemotional development. MGLS:2017 field staff provided additional information about the school environment through an observational checklist.

This technical report provides general information about the study and the data files and technical documentation that are available. More information about MGLS:2017 can be found at the MGLS website (<https://nces.ed.gov/surveys/mgls/>). Research and development (R&D) restricted-use data files that contain information collected and derived from the MS1 and MS2 data collections are available.

The data are only available as R&D restricted-use products to ensure that researchers understand their limitations and analyze the data accordingly. MGLS:2017 experienced low response rates, which affects the precision of the data for certain key subgroups and limits the kinds of analyses that the MGLS:2017 data were intended to support. For example, the data do not support the production of official statistics for the population of students who were in sixth grade in the U.S. in the 2017-18 school year.

Data files include a school-level file, a student-level file, and two assessment item-level files (one for each round of data collection). Additional detail about the study design, instrumentation, data collection, and data are provided in the *User's Manual for the MGLS:2017 Data File, Restricted-Use Version* (NCES 2023-013) distributed with the study data files.

Sample Design

A national sample was selected to enable inferences about brick-and-mortar schools in the United States educating sixth-graders and about students in such schools who were in sixth grade during the fall of the 2017-18 school year. The sample design was developed to meet a set of statistical precision requirements and to produce reliable, national-level estimates for key reporting characteristics, including students' race/ethnicity, school type (i.e., public or private), and census region where the school is located. In addition, the sixth-grade sampling included a focus on students who were served in one of three disability categories specified in the Individuals with Disabilities Education Act (IDEA): autism, emotional disturbance, and specific learning disability. Autism and emotional disturbance are low-incidence disabilities. To try to ensure that enough of these students were in the study to support representative national estimates, they were oversampled, resulting in them being sampled at higher rates than possible with fully random sample draws.

¹ Students who were not in eighth grade in MS2 because they had been held back or advanced in grade were still included.

The sample for MGLS:2017 employed a multi-stage sampling design with schools selected in the first stage and students within schools selected in the second stage. In the first stage of sampling, public and private schools that offered instruction in sixth grade were stratified by school type (public, Catholic, other private), nine Census regions (New England, Middle Atlantic, East North Central, West North Central, South Atlantic, East South Central, West South Central, Mountain, and Pacific), and estimated prevalence (high/low) of students with autism or emotional disturbance. Schools were then selected, within sampling stratum, with probability proportional to a composite measure of size. In the second stage of sampling, students enrolled in sixth grade and ungraded students ages 11-13 whose Individualized Education Program (IEP) indicated a primary classification of autism, emotional disturbance, or a specific learning disability were selected within each sampled school using simple random sampling within student sampling strata within schools. The domains were ordered such that if a student was counted in a domain, they were not included in later domains in the list. The domain order was students who 1) had autism, 2) had an emotional disturbance, 3) had a specific learning disability, 4) were Asian and not Hispanic, 5) were Hispanic, 6) were Black, and 7) were of any other race.

As discussed below in the section on response rates, the study experienced low participation among both selected schools and individuals asked to participate. As a result, the participating sample sizes for some groups defined by the characteristics used to develop the sampling plan were too small to meet the study precision requirements and support reliable, national-level estimates. While the MGLS:2017 data may be adequate for some types of analyses, they must be used with caution.

Data Collection

The MGLS:2017 base-year data collection (MS1) was conducted in January through July 2018 of the 2017–18 school year, when students were in the sixth grade. The follow-up data collection (MS2) was conducted 2 years later, in January through July 2020, when the majority of students were in eighth grade. The student data collection component, or student session, included direct assessments measuring students' mathematics achievement, reading achievement, and executive function and a student survey capturing students' attitudes and behaviors and family and school experiences.

In the base year, more than 99 percent of students who participated did so via in-school group sessions held during the school day. The assessments and student survey were web-based and self-administered on Chromebooks provided by the study in a proctored setting. Sessions were scheduled to take up to 90 minutes to complete. The small percentage of students who did not participate in the in-school group sessions completed the student session out-of-school on their own device. During the follow-up round, student data collection activities moved from a primarily in-school model to an out-of-school model midway through the collection period due to the impact of the COVID-19 pandemic. As a result, while student sessions were conducted in schools from January through mid-June in 2018 during the base-year round, in the follow-up round in-school student sessions were conducted only from January through mid-March of 2020. The remaining sessions were conducted online out-of-school from April through July. Therefore, during the follow-up round, about 30 percent of students who participated did so in the out-of-school student session.

There were differences in which components and which version of some assessments were administered depending on whether the student session was administered in school or out of school. Exhibit 1 provides a summary of the components included in the in-school and out-of-school

administrations of the student session. Note that an option to take a 45-minute or a 75-minute session was offered to students in MS2 beginning April 27, 2020.

Exhibit 1. MGLS:2017 base-year data collection (MS1) and follow-up (MS2) student session components for in-school and out-of-school participants

Component	In-school session	Out-of-school session	
		45-minute option	75-minute option ¹
Student survey	Yes	Yes	Yes
Reading assessment	Router and second-stage form	Router only	Router and second-stage form
Math assessment	Router and second-stage form	Router only	Router and second-stage form
Executive function tasks	Yes	No	No
Height and weight measurement	Yes	No	No

¹ A 75-minute out-of-school session option was available in MS2 only.

SOURCE: U.S. Department of Education, National Center for Education Statistics, Middle Grades Longitudinal Study of 2017-18 (MGLS:2017), spring 2018 and spring 2020.

Information about student and family characteristics, background, and experiences was collected through parent/guardian surveys conducted in both rounds of data collection. The parent survey included questions about family involvement in their child’s education and family characteristics that are associated with academic achievement and other student outcomes. Questions were designed to gather information to supplement the information collected from students and teachers about the students’ educational experiences and to learn about parents’ expectations for their children’s academic attainment in high school and beyond. Parent surveys were conducted from January 2018 through August 2018 for the base-year data collection and from January 2020 through July 2020 for the follow-up data collection. Parents could participate on their own via a web-based survey or over the phone with an interviewer via a computer-assisted telephone interview (CATI). Surveys were completed in English or Spanish, in accordance with the parents’ language preference. Across both rounds of the study, all materials asked that “the person most knowledgeable about the student’s school and home life” participate in the survey.

In addition, information about students’ schools and the school environment was collected through school administrator, math teacher, and special education teacher (when applicable) surveys in both rounds of data collection. The web-based school administrator survey gathered information about school factors that may influence student development, motivation, and learning, such as a school’s characteristics and staffing, and included items designed to capture a general description of the schools’ students and the services and programs available to them. In addition, school administrators answered a set of questions about their background and experience.

The math teacher surveys were designed to collect information on the class instructional environment, teacher characteristics and working environment, and teacher ratings of individual students' attitudes and confidence in math. The web-based surveys consisted of teacher-level, class-level, and student-level items, the latter referred to as teacher student reports or TSRs. The respondent for the math teacher instrument was the person who provided the student with mathematics instruction. The teacher-level items collected information about teachers' educational background, demographics, and teaching experience, as well as school-level services and characteristics such as special programs offered by the school, school climate, and instructional leadership. The class-level portion included questions measuring potential class-level correlates of students' mathematics achievement such as class size, heterogeneity of ability in a class, and students' opportunities to learn the content assessed on the direct mathematics assessment. The student-level information included class attendance, behavior, and performance for individual sampled students.

The purpose of the special education teacher survey was to gather information about the teacher or related service providers serving students with IEPs and services for students with disabilities during the middle grades. The respondent for these instruments was identified as the person who provided the sampled student with special education instruction or the person who knew the most about the student's IEP. The special education teacher or related service provider completed a teacher student report (TSR) for each sampled student linked to that teacher. Teacher-level information included information such as teacher training, experience, and instructional practices. The student-level information included category of disability, IEP goals and evaluation, and student skills in multiple academic domains. Like the math teacher survey, the special education teacher survey was a web-based, self-administered instrument.

Finally, during visits to the sampled schools, field staff completed a school environment checklist that recorded their observations regarding the condition of school facilities, grounds, and the surrounding neighborhood.

Direct Cognitive Assessments

The study measured the progress that students made in reading and mathematics and their developmental trajectories, including executive function, in the middle grade years and as they prepared for high school. The direct cognitive assessments of reading, mathematics, and executive function were administered as unique components within the computer-based student session during MS1 and MS2.

Most students participated in the cognitive assessments. However, some students with disabilities did not participate in some assessments because their IEP restricted the administration of standardized assessments in certain domains. More students with disabilities were excluded from the reading assessment than from the other assessments. Also, some students who were English learners were not administered some assessments because the assessments were in English and the students' proficiency in English was too limited to obtain a valid estimate of the student's ability.

Cognitive assessment scores for reading, mathematics, and executive function developed for MGLS:2017 were produced using item response theory (IRT) methodology. Ability (theta) scores for mathematics, reading, and executive function are provided in the data file for both MS1 and MS2. The

theta scores (ability estimates) range from negative to positive values, with higher scores indicating higher ability.

The MGLS:2017 scale scores for mathematics and reading were created to provide a more user-friendly version of these cognitive assessment scores. A linear transformation that maintains the score distribution was applied to the IRT ability estimates to create scale scores that are all positive values. These scores can be compared within content area (e.g., mathematics or reading) across MS1 and MS2 to study growth in student skills and development. It must be noted that scale scores for different content areas in the MGLS:2017 have not been linked and are not comparable to each other. For example, it is not appropriate to interpret scores that are higher in one content area to mean that the student is doing better in that content area than in the other.

Reading and Mathematics

The mathematics and reading assessments used a multistage testing (MST) design. Each mathematics and reading assessment consisted of a router and three second-stage forms. The student's router performance determined which second-stage form (low, middle, or high difficulty) was administered. Items appropriate for use in both rounds were strategically placed in both the MS1 and MS2 assessment forms to strengthen the longitudinal scale. Additional information about the MGLS:2017 direct assessments is presented in detail in the *User's Manual for the MGLS:2017 Data File, Restricted-Use Version* (NCES 2023-013).

Mathematics. The MGLS:2017 mathematics assessment framework specifies the content to be assessed, covering important mathematics skills and knowledge acquired during the middle grades that are indicative of algebra readiness. A goal of the framework was to provide a strong longitudinal measure of students' abilities in mathematics, specifically their progress in the elements leading to algebra proficiency. The concept of algebra readiness is underpinned by four mathematics domains: number system, ratios and proportional relationships, expressions and equations, and functions. The number system includes fractions (and the prerequisite understanding of multiplication and division), the rational number system, negative numbers, and the use of exponents. Proportional relationships and reasoning (ratios, rates, and proportions) build on multiplicative reasoning. These, together with an understanding of expressions and equations, provide a foundation for understanding linear equations, slope, and functions. Within the four mathematics domains, several learning standards were included to describe the types of learning objectives that fall within each domain. A test blueprint specified the number of items needed to assess each mathematics learning standard within each of the four domains and along three depth-of-knowledge categories, specifically knowing, applying, and reasoning. Using this representation of the longitudinal model of growth, the MGLS:2017 mathematics assessment provides a single measure of middle-grade students' high school algebra readiness.

Reading. The MGLS:2017 reading assessment framework addresses foundational reading skills and more advanced reading comprehension. There is an emphasis on reading informational text that is purpose-driven and assessing a wide range of comprehension skills. Foundational reading skills include word recognition and decoding, vocabulary, morphology, sentence comprehension, reading efficiency, and basic reading comprehension of passages. Higher-level reading comprehension integrates all these skills into the ability to think critically about a text. This can include integrating, synthesizing, evaluating, and applying information from multiple texts. The reading assessments drew on preexisting measures

developed by the Educational Testing Service (ETS) for the Institute of Education Sciences (IES) under a Reading for Understanding grant (award numbers R305G040065 and R305F100005).

The MGLS:2017 reading assessment provides the data for creating two scores for each round of data collection that represent an overlapping continuum of reading skills and abilities. The two sets of reading scores available on the data file are: a) reading foundation scores for MS1 and MS2 derived from the router portion of the assessment that are measures of reading foundational skills and b) overall reading scores for MS1 and MS2 derived from both the router and second-stage forms of the reading assessment that are measures of both reading foundational skills and advanced reading comprehension.

Executive Function

The MGLS:2017 student session included measurements of executive function as part of the battery of direct cognitive assessments. Executive function permits individuals to self-regulate, engage in purposeful and goal-directed behaviors, and conduct themselves in a socially appropriate manner. Self-regulation is needed for social success, academic and career success, and good health outcomes. Executive function includes capacities such as shifting (cognitive and attentional flexibility), inhibitory control, and working memory (Hoyle and Davisson 2011; Miyake et al. 2000).

MGLS:2017 included measures of cognitive flexibility, inhibitory control, and working memory. The Hearts and Flowers task assesses cognitive flexibility and inhibition of a natural prepotent response (Diamond et al. 2004). A prepotent response is a type of cognitive bias that occurs when an individual responds quickly to a test stimulus without conscious thought or deliberation. The Hearts and Flowers task was administered via touch-sensitive tablet computer (i.e., Chromebook) to students who completed the MGLS:2017 student session in-school.² For example, the congruent condition (Hearts) is presented first with instructions to press the key on the same side as a heart appears. Next, the incongruent condition (Flowers) is presented. Flowers requires inhibition of the prepotent response of pressing the key on the same side as the flower appears. This test was administered in both MS1 and MS2.

The Spatial 2-Back task (Hoang and Fellows n.d.; Romer 2009) is an n-back task that assesses working memory with nonverbal stimuli and measures inhibitory control. An n-back test is a type of cognitive neuroscience task used to measure working memory and the ability to actively keep information in the mind and manipulate it. In the typical n-back test, a participant is shown a sequence of stimulus items (such as letters or numbers) and is asked to identify when the current item matches a previously seen item some number of steps back. This "number of steps back" is referred to as the n. The Spatial 2-Back task in MGLS:2017 used line drawings of unfamiliar objects presented in a series with each object visible for 2000ms. This test was included only in MS2 as it was found to be too challenging for most sixth-grade students.

Student Survey

A key goal of MGLS:2017 was to understand the supports that students need for academic success, high school readiness, and positive life development, such as high school graduation, college and career readiness, and healthy lifestyles. In each round of data collection, a student survey was fielded as part of the student session. The student survey collected information on a variety of topics, some of which were

² The executive function tasks were not fielded in the student sessions administered out-of-school due to the need to capture timing information and due to the sensitivity of these measures to variations in the testing environment that were more controlled under the in-school administration.

single-item indicators such as behaviors, out-of-school time use, and family, school, and classroom environments. Topics that represented psychological or latent traits were covered by multi-item sets for which results were transformed into scale scores. These multi-items sets measured constructs in the areas of attitudes, opinions and socioemotional states related to social relationships, social supports, and academic engagement. Specific constructs covered include growth mindset, math self-efficacy, subjective task value related to math, school belongingness, peer social support, teacher support, conscientiousness, optimism, sensation-seeking, and academic engagement. The student survey data augment the information collected from the mathematics, reading, and executive function assessments to provide a deeper understanding of the social and contextual factors related to students' academic and non-academic circumstances and outcomes.

Response Rates and Nonresponse Bias Analysis

The sixth-grade round of the MGLS:2017 Main Study administration (MS1) obtained participation from about 570 schools and about 14,280 cases in which the student or parent participated. The follow-up (MS2) administration of MGLS:2017, conducted when most students were in the eighth grade, obtained participation from about 560 schools from the base-year sample. An augmentation sample of about another 190 schools supplemented and refreshed the base-year sample. It included some schools that were sampled for and did not participate in MS1 but then did participate in MS2. From the schools that were added in MS2 came about 4,840 students, all of whom were in eighth grade.

The study sample and sampling procedures were designed to meet certain precision requirements to allow for analysis of key subgroups of students. Unfortunately, due to low response rates at both the school and student levels, some of the minimum requirements were not met in the end, for example for students in private schools and students in the focal disability groups. This section describes the school-level response rates and overall response rates for the student assessments and parent interviews for the MS1 and MS2 data collections.

School response rates, shown in Table 1 below, reflect schools that allowed student sampling and a student session to take place. For MS1, the weighted response rate for all schools was 41 percent, with a 45 percent response rate for public schools and a 32 percent response rate for private schools. To accommodate the introduction of the school augmentation sample in MS2, a base-year sampled school is considered a base-year responding school as of MS2 if the school allowed student sampling and a student session to take place in MS1 or MS2. MS2 response rates are determined considering MS2 participation for schools defined as a base-year responding school as of MS2 as well as schools new for MS2. The weighted response rate for all schools in MS2 was 34 percent, with a 33 percent response rate for public schools and a 35 percent response rate for private schools.

Table 1. School response rates, by school type: School years 2017–18 and 2019–20

School type	MS1 weighted response rate ¹	MS2 weighted response rate ²
Total	41	34
School type		
Public	45	33
Private	32	35

¹ Estimates were weighted using the MS1 school sampling weight.

² Estimates were weighted using the MS2 school sampling weight.

SOURCE: U.S. Department of Education, National Center for Education Statistics, Middle Grades Longitudinal Study of 2017–18 (MGLS:2017), spring 2018 and spring 2020.

Overall student response rates are the product of the school-level response rates and student-level response rates for the relevant component, i.e., the student assessment, the parent survey, etc. A student was a respondent to the student assessment if sufficient data had been collected for at least one of the following: math assessment, reading assessment, Hearts and Flowers task, 2-Back task, or height and/or weight measurement. Response rates for the parent surveys represent the percentages of students whose parents completed the survey. Each of the response rate estimates at the student level was weighted using the school sampling weight times a student sampling adjustment.

Table 2 presents the overall response rates for the student assessment and parent survey for school years 2017–18 and 2019–20. In MS1, the overall student assessment weighted response rate was 33 percent, with a 36 percent response rate for students in public schools and a 28 percent response rate for students in private schools. The overall parent survey weighted response rate was 21 percent, with a 23 percent response rate for students in public schools and an 18 percent response rate for students in private schools. In MS2, the overall student assessment weighted response rate was 21 percent, with a 20 percent response rate for students in public schools and a 22 percent response rate for students in private schools. The overall parent survey weighted response rate was 17 percent, with a 17 percent response rate for students in public schools and a 19 percent response rate for students in private schools.

Table 2. Student assessment and parent survey overall response rates, by school type: School years 2017–18 and 2019-20

School type	Student assessment ¹		Parent survey ²	
	MS1 weighted response rate	MS2 weighted response rate	MS1 weighted response rate	MS2 weighted response rate
Total	33	21	21	17
School type				
Public	36	20	23	17
Private	28	22	18	19

¹ Student had sufficient data for at least one of the following: mathematics assessment, reading assessment, an executive function assessment, and height or weight measurement.

² Student's parent participated in the survey.

SOURCE: U.S. Department of Education, National Center for Education Statistics, Middle Grades Longitudinal Study of 2017–18 (MGLS:2017), spring 2018 and spring 2020.

Longitudinal response rates are intended to provide a measure of the degree to which respondents who participated in multiple study rounds represent the target population. The overall longitudinal response rate was calculated by multiplying two values: a school-level weighted response rate as of MS1 and a student-level weighted response rate where the student-level weighted response rate was calculated as the weighted proportion of students who responded in MS1 and MS2. Table 3 presents the student assessment and parent survey student-level overall longitudinal response rates based on the school and students sampled in MS1. The overall weighted longitudinal response rate for the student assessment was 21 percent, with a 23 percent response rate for students in public schools and a 19 percent response rate for students in private schools. The overall weighted longitudinal response rate for the parent survey was 16 percent, with an 18 percent response rate for students in public schools and a 15 percent response rate for students in private schools.

Table 3. Overall longitudinal response rates for student assessment and parent survey, by school type: School years 2017–18 and 2019–20

School type	Weighted response rate ¹	
	Student assessment ²	Parent survey ³
Total	21	16
School type		
Public	23	18
Private	19	15

¹ Estimates were weighted using the school sampling weight times a student sampling adjustment.

² Student had sufficient data in both 2017–18 and 2019–20 for at least one of the following: mathematics assessment, reading assessment, an executive function assessment, and height or weight measurement.

³ Student’s parent participated in both the 2017–18 and 2019–20 surveys.

SOURCE: U.S. Department of Education, National Center for Education Statistics, Middle Grades Longitudinal Study of 2017-18 (MGLS:2017), spring 2018 and spring 2020.

Response rates for the school staff surveys represent the weighted percentages of students with school staff data. Table 4 presents the overall response rates for the math teacher survey for school years 2017–18 and 2019–20. In MS1, the overall weighted response rate for the math teacher survey was 23 percent for the teacher-level survey (25 percent and 21 percent for students in public and private schools, respectively) and 20 percent for the student-level survey (21 percent and 19 percent for students in public and private schools, respectively). In MS2, the overall weighted response rate was 13 percent for the teacher-level survey (12 percent and 17 percent for students in public and private schools, respectively) and 11 percent for the student-level survey (11 percent and 15 percent for students in public and private schools, respectively).

Table 4. Math teacher survey overall response rates, by school type: School years 2017–18 and 2019–20

School type	Math teacher survey (teacher-level) ¹ response rates		Math teacher survey (student-level) ² response rates	
	MS1 weighted response rate	MS2 weighted response rate	MS1 weighted response rate	MS2 weighted response rate
Total	23	13	20	11
School type				
Public	25	12	21	11
Private	21	17	19	15

¹ Math teacher linked to a student completed the teacher survey.

² Math teacher linked to a student completed a report for the student.

SOURCE: U.S. Department of Education, National Center for Education Statistics, Middle Grades Longitudinal Study of 2017–18 (MGLS:2017), spring 2018 and spring 2020.

The special education teacher surveys were only administered in public schools, so the school response rates used in constructing overall response rates for special education teacher surveys specifically use school response rates for public schools. Table 5 presents the overall response rates for the special education teacher survey for school years 2017–18 and 2019–20. In MS1, the weighted overall response rate for the special education teacher surveys was 15 percent for the teacher-level survey and 15 percent for the student-level survey. In MS2, the weighted overall response rate was 9 percent for the teacher-level survey and 9 percent for the student-level survey.

Table 5. Special education teacher survey overall response rates, public school only: School years 2017–18 and 2019–20

Public school	Special education teacher survey (teacher-level) ¹ response rates		Special education teacher survey (student-level) ² response rates	
	MS1 weighted response rate	MS2 weighted response rate	MS1 weighted response rate	MS2 weighted response rate
Total	15	9	15	9

¹ Special education teacher linked to student completed teacher survey.

² Special education teacher linked to student completed a report for the student.

SOURCE: U.S. Department of Education, National Center for Education Statistics, Middle Grades Longitudinal Study of 2017–18 (MGLS:2017), spring 2018 and spring 2020.

Table 6 presents the overall response rates for the school administrator survey for school years 2017–18 and 2019–20. The overall weighted response rate was 27 percent (29 percent and 25 percent for students in public and private schools, respectively) in MS1. In MS2, the overall weighted response rate

for the school administrator survey was 13 percent (13 percent and 16 percent for students in public and private schools, respectively).

Table 6. Overall school administrator survey response rates, by school type: School years 2017–18 and 2019–20

School type	School administrator survey (student-level) ¹ response rates	
	MS1 weighted response rate	MS2 weighted response rate
Total	27	13
School type		
Public	29	13
Private	25	16

¹ Student’s school administrator participated in the survey.

SOURCE: U.S. Department of Education, National Center for Education Statistics, Middle Grades Longitudinal Study of 2017–18 (MGLS:2017), spring 2018 and spring 2020.

A nonresponse bias analysis was conducted to determine if substantial bias was introduced as a result of unit-level nonresponse at the school level and student level. Nonresponse bias was estimated for characteristics available in the sampling frames, which would have been known for most participating and nonparticipating schools and students, such as student race, Hispanicity, sex, and participation during the COVID-19 pandemic. Nonresponse bias was estimated for each variable or characteristic as the difference between the weighted means (proportions) of the respondents and the full sample, using the school base weight. Each estimated nonresponse bias was first tested using a *t* test to determine if it significantly differed from zero at the 5 percent significance level. Relative bias was computed as the ratio of the estimated bias to the weighted full-sample mean. Second, using the weights adjusted for nonresponse, the nonresponse bias was re-estimated and tested for significance. These tests were complemented by effect size calculations.

Finally, to better understand the effect of poststratification on efforts to reduce nonresponse bias, two additional sets of estimates were created. The first set of estimates, equal to the difference in respondent means before and after poststratification, corresponds to the effect of poststratification on nonresponse adjustments. The second set of estimates, equal to the difference between base-weighted full-sample means and the post-stratified respondent means, corresponds to the cumulative effects of all weighting and adjustment steps. These analyses found that the weight adjustments were generally successful in reducing, if not completely eliminating, bias for the characteristics measured. However effective the weight adjustments were for correcting bias in characteristics available in the frame information, there may still be bias on other measures in the MGSL:2017 data that do not correspond to frame information. More detail on the nonresponse bias analysis is available in data documentation. Note that it is still important to acknowledge limitations in understanding the extent of the nonresponse bias given the low response rates at the school level and for individual study components at the student level.

Further details about the students included in the participating sample for the MGLS:2017 are provided in Table 7 below. Readers should keep in mind these estimates are for the sample that participated in the MGLS and do not present a national population of sixth-graders.

Table 7. Student, family, and school characteristics: Unweighted frequency and percentage distribution of the MGLS:2017 sample, by student, family, and school characteristics: Spring 2018 and spring 2020

Student, family, and school characteristics	n	Percent
Total	19,120	100.0
Student's sex		
Male	10,040	52.5
Female	9,080	47.5
Student's race/ethnicity ¹		
American Indian/Alaska Native, non-Hispanic	260	1.4
Asian, non-Hispanic	1,060	5.6
Black, non-Hispanic	2,330	12.3
Hispanic	5,140	27.0
Native Hawaiian/Pacific Islander, non-Hispanic	80	0.4
White, non-Hispanic	8,410	44.2
More than one race, non-Hispanic	1,750	9.2
Primary language in the student's home		
English	16,180	85.3
Language other than English	2,500	13.2
Two or more languages spoken equally	290	1.5
Poverty status ²		
Income below 100 percent of the federal poverty level	2,770	20.6
Income at or above 100 percent of the federal poverty level	10,690	79.4
Parents' highest level of education ³		
Less than high school	1,380	10.2
High school diploma or equivalent	4,160	30.9
Vocational/technical diploma after high school or associate's degree	2,310	17.2
Bachelor's degree	2,990	22.2
Master's degree (M.A., M.S.) or doctorate degree (Ph.D., Ed.D.)	2,630	19.5
School type		
Public	16,520	86.4
Catholic or other private	2,600	13.6

See notes at end of table.

Table 7. Student, family, and school characteristics: Unweighted frequency and percentage distribution of the MGLS:2017 sample, by student, family, and school characteristics: Spring 2018 and spring 2020—Continued

Student, family, and school characteristics	n	Percent
Percentage of students eligible for free or reduced-price lunch at school ⁴		
Less than 25 percent	2,130	13.8
25 to 50 percent	4,660	30.2
51 to 75 percent	4,300	27.9
More than 75 percent	4,340	28.1
School locale		
City	5,790	30.3
Suburban	7,000	36.6
Town	2,650	13.9
Rural	3,680	19.2
Grade configuration of school		
Prekindergarten/kindergarten/first grade through sixth grade	1,750	9.3
Prekindergarten/kindergarten/first grade through eighth grade	3,340	17.6
Fifth grade/sixth grade through eighth grade	11,090	58.6
Other ⁵	2,740	14.5

¹ Black, non-Hispanic includes African American. Hispanic includes Latino and Latina.

² Poverty status for Main Study Base Year (MS1) is based on U.S. Census thresholds for 2017, and poverty status for Main Study Follow-up (MS2) is based on U.S. Census thresholds for 2019. The Census thresholds identify incomes determined to meet household needs, given the household size. For example, in 2017 a family of two with a householder under age 65 was below the poverty threshold if its income was lower than \$16,414; in 2019 a family of two with a householder under age 65 was below the poverty threshold if its income was lower than \$17,120.

³ Parents' highest level of education is the highest level of education achieved by either of the parents or guardians in a two-parent household or by the only parent or guardian in a single-parent household.

⁴ Percentage of students eligible for free or reduced-price lunch only applies to public schools.

⁵ "Other" includes "Prekindergarten/Kindergarten/First grade through 12th grade" and any other grade configurations not listed in the table.

NOTE: This table describes students enrolled in sixth grade in school year 2017-18 but incorporates data collected either in the spring of 2018 or the spring of 2020. All student, family, and school characteristics are drawn from the Follow-up (MS2) data when the Main Study Base Year (MS1) data are missing (i.e., the student/school joined in MS2 or the student attended the same school across both rounds of data collection but MS1 school information is missing). Poverty status and parents' highest level of education include some imputed values. Sample sizes are rounded to the nearest 10. Detail may not sum to total due to rounding and/or missing data.

SOURCE: U.S. Department of Education, National Center for Education Statistics, Middle Grades Longitudinal Study of 2017-18 (MGLS:2017), Base Year and Follow-Up Restricted-Use Data File.

Data Reliability

Estimates produced using data from MGLS:2017 are subject to two types of error: nonsampling errors and sampling errors. Nonsampling errors are errors made in the collection and processing of data.

Sampling errors occur because the data are collected from a sample rather than a census of the population.

Nonsampling Errors

Nonsampling error is the term used to describe variations in the estimates that may be caused by population coverage limitations, as well as data collection, processing, and reporting procedures. The sources of nonsampling errors are typically nonresponse, differences in respondents' interpretations of the meaning of the questions, response differences related to the specific time the survey was conducted, and mistakes in data preparation.

In general, it is difficult to identify and estimate either the amount of nonsampling error or the bias caused by this error. In MGLS:2017, efforts were made to prevent such errors from occurring and to compensate for them where possible (e.g., field testing items and assessments, using survey questions that had been tested and used in previous surveys, multiday field staff training, field staff certification sessions, and monitoring field staff performance and field data quality throughout the collection period).

One aspect of nonsampling error that can be measured is nonresponse. Nonresponse bias due to a respondent not participating at all was discussed earlier in the response rate section. However, item-level nonresponse bias is also possible in situations where a respondent skips some items they should have answered. Analysis of potential bias due to item nonresponse is typically conducted for those items with a response rate below 85 percent. Across all the MGLS:2017 surveys, there were 823 school-level variables and 2,802 student-level variables. Of these, 318 school-level and 1,130 student-level variables had response rates below 85 percent. In MS1, 84 percent of variables associated with the full parent survey had response rates below 85 percent. In MS2, 81 percent of variables associated with the full parent survey had response rates below 85 percent. In both MS1 and MS2, all items from the student survey were above the 85 percent threshold. An item nonresponse bias analysis was not conducted because of the low level of response at the unit level.

Sampling Errors and Weighting

Analyses using MGLS:2017 data should be weighted to account for school and student sampling, school and student nonresponse, and sampling frame coverage limitations. The analysis weights provided with the MGLS:2017 data incorporate adjustments to account for nonresponse and for school and student population coverage issues that arose from use of a school sampling frame that contained information on schools as of the 2013-14 school year. Adjustments were made in the student-level weights to make the data representative of students, or subsets of those students, enrolled in grade 6, in the United States, including the District of Columbia, in the 2017-18 school year.

The MGLS:2017 provides multiple analytic weights to support data analysis at the school and student levels. Teacher and parent data may be analyzed using student analysis weights, but inference is to the MGLS:2017 student population, not the populations of teachers and parents associated with the MGLS:2017 student population. While the MGLS:2017 study selected random samples of schools and students, it did not select random samples of parents or teachers that would be nationally representative. The design supports statements such as "6th-grade students with teachers who had more than 10 years of experience had an average math score of X." It does not support a statement such as "X percent of teachers had 10 years or more of experience." The school- and student-level data files created for MGLS:2017 include analysis weights that account for population coverage, as noted above, and aspects of the sample design and data collection, including unequal probabilities of selection and nonresponse. Fewer weights were developed for use with data collected in MS2 compared to MS1

because of the lower-than-expected response rates across multiple study components. Exhibit 2 describes the analytic weights available on the data files.

Exhibit 2. MGLS:2017 school and student analysis weights

School Year	Weight	Description
School weights		
2017-18	W1SCHL	School sampling weight adjusted for school nonresponse in MS1 ¹ and coverage
	W1ADMIN	School sampling weight adjusted for school nonresponse in MS1 and administrator nonresponse in MS1 ² and coverage
2019-20	W12SCHL	School sampling weight adjusted for school nonresponse in MS1 and MS2 ³ and coverage
Student Weights		
2017-18	W1S0	Student base weight adjusted for student nonresponse in MS1 and coverage
	W1SP0	Student base weight adjusted for student and parent nonresponse in MS1 and coverage
	W1PY0	Student base weight adjusted for parent and math teacher-student level nonresponse in MS1 and coverage
	W1SPM0	Student base weight adjusted for student, parent, and math teacher level nonresponse in MS1 and coverage
	W1SM0	Student base weight adjusted for student and math teacher level nonresponse in MS1 and coverage
	W1SY0	Student base weight adjusted for student and math teacher-student level nonresponse in MS1 and coverage

See notes at end of exhibit.

Exhibit 2. MGLS:2017 school and student analysis weights —Continued

School Year	Weight	Description
	W1SA0	Student base weight adjusted for student and school administrator nonresponse in MS1 and coverage
	W1SMA0	Student base weight adjusted for student, math teacher level, and school administrator nonresponse in MS1 and coverage
	W1N0	Student base weight adjusted for unknown eligibility, ⁴ special education teacher nonresponse in MS1, and coverage
	W1PN0	Student base weight adjusted for unknown eligibility, ⁴ parent and special education teacher nonresponse in MS1, and coverage
	W1P0	Student base weight adjusted for parent nonresponse in MS1 and coverage
	W1PA0	Student base weight adjusted for parent and school administrator nonresponse in MS1 and coverage
	W1SPMA0	Student base weight adjusted for student, parent, math teacher level, and school administrator nonresponse in MS1 and coverage
2019-20	W12S	Student base weight adjusted for student nonresponse in MS1, student nonresponse in MS2, and coverage.
	W12SP2	Student base weight adjusted for student nonresponse in MS1, student nonresponse in MS2, parent nonresponse in MS2, and coverage
	W12SP1P2	Student base weight adjusted for student nonresponse in MS1, student nonresponse in MS2, parent nonresponse in MS1, parent nonresponse in MS2, and coverage

¹ Schools sampled in MS1 were defined as respondents if at least one sampled student, or their parent, responded in MS1.

² Schools sampled in MS1 were defined as respondents if at least one sampled student, or their parent, responded in MS1 and the school administrator responded in MS1.

³ Schools sampled in MS1 or MS2 were defined as respondents if at least one sampled student, or their parent, responded in MS1 or MS2.

⁴ This weight represents students enrolled in a special education program in grade 6 in public schools in the 2017-18 school year. Special education teacher surveys were only requested for students in special education. However, the special education status was not known for all students, so an unknown eligibility adjustment was constructed and applied in addition to nonresponse and coverage adjustments.

SOURCE: U.S. Department of Education, National Center for Education Statistics, Middle Grades Longitudinal Study of 2017-18 (MGLS:2017), Base Year and Follow-Up Restricted-Use Data File.

Analysis of MGLS:2017 data requires the use of statistical software designed to compute standard errors that account for the complex sampling design used to select the schools and students that participated in the study. Software packages that are not designed for complex surveys calculate estimates under the assumption of a simple random sample design. This assumption is incorrect when analyzing MGLS:2017 data because it does not account for the clustering of students within schools. This can lead to estimated variances and confidence intervals that are too small and to incorrect results for statistical tests of differences.

The MGLS:2017 data file contains variables that allow for two methods of estimating variances of survey statistics produced from a study with a complex sample design: Taylor series linearization and balanced repeated replication (BRR) (Wolter 2007).

Available Data Files and Documentation

There are research and development (R&D) restricted-use data files available for MGLS:2017 that contain information collected and derived from the MS1 and MS2 data collections. Data files include a school-level file, a student-level file, and two assessment item-level files (one for each round of data collection).

For more information about the available data files please see the MGLS:2017 study website at <https://nces.ed.gov/surveys/mgls>, and the *User's Manual for the MGLS:2017 Data File, Restricted-Use Version* (NCES 2023-013).

References

- Diamond, A., Briand, L., Fossella, J., & Gehlbach, L. (2004). Genetic and neurochemical modulation of prefrontal cognitive functions in children. *American Journal of Psychiatry*, *161*(1), 125-132.
- Hoyle, R. H., & Davisson, E. K. (2011, January). Assessment of self-regulation and related constructs: Prospects and challenges. In *NRC Workshop on Assessment of 21st Century Skills*.
- Lennon, J, Herget, D., Kline, T., Wilson, D., Rogers, J., Ingold-Smith, M., Honeycutt, L., Parish, E., Brown, D., Kinney, S., Black, S., McLean, M., (2023). *Middle Grades Longitudinal Study of 2017–18 (MGLS:2017), User’s Manual for the MGLS:2017 Data Restricted Use Version (NCES 2023-013)*. U.S. Department of Education. Washington, DC: National Center for Education Statistics. Retrieved [date] from <https://nces.ed.gov/pubsearch/>.
- Miyake, A., Friedman, N. P., Emerson, M. J., Witzki, A. H., Howerter, A., & Wager, T. D. (2000). The unity and diversity of executive functions and their contributions to complex “frontal lobe” tasks: A latent variable analysis. *Cognitive psychology*, *41*(1), 49-100.
- Romer, D., Betancourt, L., Giannetta, J.M., Brodsky, N.L., Farah, M., & Hurt, H. (2009). Executive cognitive functions and impulsivity as correlates of risk taking and problem behaviour in preadolescents. *Neuropsychologia*, *47*, 2916-2926.
- Shah, V. B. B. B., Barnwell, B. G., & Bieler, G. S. (2005). SUDAAN user’s manual, release 9. *Research Triangle Park, NC: Research Triangle Institute*.
- Wolter, K. M. (2007). *Introduction to variance estimation* (Vol. 53). New York: Springer.

Appendix A: Variables of Interest

Variables of Interest

This section provides additional details about the variables included in table 7. Variable names from the data file are written in capital letters and appear in brackets or parentheses in the variable descriptions.

Student and Family Characteristics

Several variables presented in table 7 were derived by combining information from one or more data sources or multiple items in the MGLS:2017 data collection. More information on the derivation of key variables is provided in chapter 7 of the *User's Manual for the MGLS:2017 Data File, Restricted-Use Version* (NCES 2023-013).

Student's sex [X_SEX]. Information on student's sex was obtained from schools at the time of sampling. It was also collected from students and their parents in the MS1 and MS2 student and parent surveys. The composite variable indicating student's sex was derived using data from these three sources, with an order of preference for which source should be used. Specifically, these sources of data were prioritized by using information from the student survey, if it was available; then, from the parent survey, if it was available. Information collected from schools was used if neither student nor parent information was available.

Student's race/ethnicity [X_STUDRAC]. Information on the student's race and ethnicity was collected in the student survey, the parent survey, and from schools. When there were multiple sources of race/ethnicity data, the student survey data were given priority. When student data were missing, parent-reported data were used. If both the student and parent survey data were missing, information from the school was used.

The composite variable, X_STUDRAC, includes the following categories:

- American Indian/Alaska Native, non-Hispanic
- Asian, non-Hispanic
- Black, non-Hispanic
- Hispanic, no race specified
- Hispanic, race specified
- More than one race, non-Hispanic
- Native Hawaiian/Pacific Islander, non-Hispanic
- White, non-Hispanic

In table 7, the two Hispanic categories were collapsed into one category, thereby classifying a student as Hispanic if the student indicated their ethnicity was Hispanic, regardless of whether a race was identified and what that race was.

Primary language in the student's home [X1LANGST, X2LANGST]. These composite variables indicate whether English was the primary language in the student's home, if a different (i.e., non-English) language was the primary language in the student's home, or if two or more languages were spoken equally. The parent survey (P1ENGLANG and P1PRMLNG for MS1; P2ENGLANG and P2PRMLNG for MS2) was the primary source of information about the primary language spoken in the student's home. If this information was missing in the parent survey, the student survey question about English as a first

language (S1LANG1ENG for MS1; S2LANG1ENG for MS2) was used as the source. If both sources were missing, values were set to -9 (not ascertained) for X1LANGST and X2LANGST.

Poverty status [X1POVTY100_I, X2POVTY100_I]. The federal poverty-level status composite variables (X1POVTY100_I for MS1, X2POVTY100_I for MS2) were derived from household income (X1INCOME_I and X2INCOME_I for MS1 and MS2, respectively) and the total number of household members (X1HTOTAL_I and X2HTOTAL_I for MS1 and MS2, respectively). The household income composite variable for each round has the same category definitions. The income composite variables also include imputed values for cases for which parent data were missing. The midpoint value of the household income composite variables was used to develop the federal poverty-level status composite variables. For example, if the MS1 household income composite variable, X1INCOME_I, is equal to 9 (\$40,001-\$45,000), the value 42,500 was used as part of the input to the federal poverty-level status composite variable, X1POVTY100_I. The same applies to MS2.

Like the household income variable, the total number of household members was collected in the parent survey in both MS1 and MS2. The household count composite variable for each round (X1HTOTAL_I, X2HTOTAL_I) provides the count of the total number of household members at the time of the parent survey and was imputed when the response from the parent survey was missing.

The MS1 federal poverty-level status composite variable, X1POVTY100_I, was created based on the MS1 household count composite variable, X1HTOTAL_I, the midpoint value of X1INCOME_I, and on US Census thresholds for 2017. For example, if a household contained three members (two adults and one child under the age of 18) and the household income was lower than \$19,730, the household was classified as being below 100 percent of the federal poverty level. If a household with three members had an income of \$19,730 or more, the household was classified as being at 100 percent of the federal poverty level or above.

Similarly, the MS2 federal poverty-level status composite variable, X2POVTY100_I, was created based on the MS2 household count composite variable, X2HTOTAL_I, the midpoint value of X2INCOME_I, and on US Census thresholds for 2019. For example, if a household contained three members (two adults and one child under the age of 18) and the household income was lower than \$20,578, the household was classified as being below 100 percent of the federal poverty level. If a household with three members had an income of \$20,578 or more, the household was classified as being at 100 percent of the federal poverty level or above.

Parents' highest level of education [X1PAR1EDU_I, X1PAR2EDU_I, X2PAR1EDU_I, X2PAR2EDU_I]. Parents' highest level of education was the highest level of education achieved by either of the parents or guardians in a two-parent household or by the only parent or guardian in a single-parent household. Parents' highest level of education was derived from parent survey information about parents' educational attainment. The composite variables X1PAR1EDU_I, X1PAR2EDU_I, X2PAR1EDU_I, and X2PAR2EDU_I include the following categories: less than high school; high school diploma or the equivalent; vocational/technical diploma after high school; associate's degree; bachelor's degree; master's degree (M.A., M.S.); and doctoral degree (Ph.D., Ed.D.). In addition, values were imputed for cases in which parent respondent data were missing. More information on imputation is provided in chapter 5 of the *User's Manual for the MGLS:2017 Data File, Restricted-Use Version* (NCES 2023-013).

In table 7, the parent education composite variable values available in the data file were collapsed into five categories: less than high school, high school diploma or the equivalent, vocational/technical diploma after high school or associate's degree, bachelor's degree, and master's degree (M.A., M.S.) or doctorate degree (Ph.D., Ed.D.).

School Characteristics

The school characteristics composite variables used in table 7 were derived from information collected through the Common Core of Data (CCD), a universe data collection of all public schools in the United States that is conducted by NCES, and the Private School Universe Survey (PSS), a universe data collection of all private schools in the United States that is conducted by NCES. More information on the derivation of key variables is provided in chapter 7 of the *User's Manual for the MGLS:2017 Data File, Restricted-Use Version* (NCES 2023-013).

School type [X1CONTROL, X2CONTROL]. These composite variables indicate whether the student's school was a public school or a private school in MS1 and MS2, respectively.

Percentage of students eligible for free or reduced-price lunch [X1FRPMEAL, X2FRPMEAL]. These composite variables indicate the percentage of students in the school who were eligible for free or reduced-price school lunch. The MS1 school lunch composite variable, X1FRPMEAL, was derived from the sum of the free lunch and reduced-price lunch variables in the CCD for the 2017-18 school year. When data were not available from the 2017-18 school year, the percentages of students eligible for free and reduced-price lunch in the previous two school years (i.e., 2015-16 and 2016-17) were averaged.

Similarly, the MS2 school lunch composite variable, X2FRPMEAL, was derived from the sum of the free lunch and reduced-price lunch variables in the CCD for the 2019-20 school year. When data were not available from the 2019-20 school year, the percentages of students eligible for free and reduced-price lunch in the previous two school years were averaged.

Some schools receive Title I funding due to the percentage of their students living in households with incomes below the federal poverty threshold; these schools do not require students to apply for free or reduced-price lunch individually. Instead, the schools certify the entire school population as eligible. In these cases, if free and reduced-price lunch estimates were not available in the CCD, the composite value was set at 100 percent.

In table 7, the categories for the free or reduced-price lunch composite variables were collapsed into four categories: Less than 25 percent, 25 to 50 percent, 51 to 75 percent, and more than 75 percent.

School locale [X1LOCALE, X2LOCALE]. These composite variables indicate the locality or urbanicity type of the student's school. The MS1 school locale composite variable, X1LOCALE, was derived from the CCD for the 2017-18 school year and the PSS data for the 2017–18 school year. The MS2 school locale composite variable, X2LOCALE, was derived from the CCD for the 2019-20 school year and the PSS data for 2019–20. When data were missing from the 2017-18 CCD or PSS or from the 2019-20 CCD or PSS, the prior year's data were used. The locale categories reported in the CCD or PSS are the following:

- City, Large: Territory inside an urbanized area and inside a principal city with a population of 250,000 or more;

- City, Midsize: Territory inside an urbanized area and inside a principal city with a population less than 250,000 and greater than or equal to 100,000;
- City, Small: Territory inside an urbanized area and inside a principal city with a population less than 100,000;
- Suburb, Large: Territory outside a principal city and inside an urbanized area with a population of 250,000 or more;
- Suburb, Midsize: Territory outside a principal city and inside an urbanized area with population less than 250,000 and greater than or equal to 100,000;
- Suburb, Small: Territory outside a principal city and inside an urbanized area with a population less than 100,000;
- Town, Fringe: Territory inside an urban cluster that is less than or equal to 10 miles from an urbanized area;
- Town, Distant: Territory inside an urban cluster that is more than 10 miles and less than or equal to 35 miles from an urbanized area;
- Town, Remote: Territory inside an urban cluster that is more than 35 miles from an urbanized area;
- Rural, Fringe: Census-defined rural territory that is less than or equal to 5 miles from an urbanized area, as well as rural territory that is less than or equal to 2.5 miles from an urban cluster;
- Rural, Distant: Census-defined rural territory that is more than 5 miles but less than or equal to 25 miles from an urbanized area, as well as rural territory that is more than 2.5 miles but less than or equal to 10 miles from an urban cluster; and
- Rural, Remote: Census-defined rural territory that is more than 25 miles from an urbanized area and is also more than 10 miles from an urban cluster.

For Table 7, the school locale composite variables were collapsed into four categories: city, suburban, town, and rural.

Grade configuration [X1LOWGRD/X1HIGGRD, X2LOWGRD/X2HIGGRD]. These composite variables indicate the lowest and highest grades offered at the student’s school in MS1 and MS2. For example, the category “Prekindergarten/kindergarten/first grade through sixth grade” includes students who attended schools with a lowest grade of prekindergarten, kindergarten, or first grade and a highest grade of sixth grade. The MS1 grade configuration composite variables, X1LOWGRD and X1HIGGRD, were drawn primarily from the CCD and PSS for the 2017-18 school year and supplemented with information collected by field staff during data collection. The MS2 grade configuration composite variables, X2LOWGRD and X2HIGGRD, were drawn primarily from the CCD and PSS for the 2019-20 school year (the most recent year available) and supplemented with information collected by field staff during data collection.

For table 7, information about the lowest and highest grades offered at the school was used to create the following four categories of grade configuration:

- Prekindergarten/kindergarten/first grade through sixth grade (i.e., PK/K/01-06);
- Prekindergarten/kindergarten/first grade through eighth grade (i.e., PK/K/01-08);
- Fifth grade/sixth grade through eighth grade (i.e., 05/06-08); and
- Other, which includes prekindergarten/kindergarten/first grade through 12th grade (i.e., PK/K/01-12) and any other grade configurations not listed.