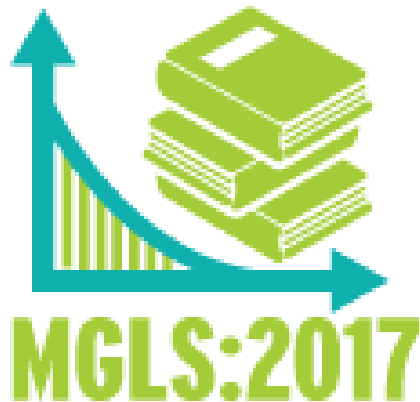


**National Center for Education  
Statistics  
(NCES)**

**Middle Grades Longitudinal Study of  
2016-2017 (MGLS:2017) Field Test  
2015 Recruitment**

**OMB #1850-NEW v. 1**

**Supporting Statement Part B**



March 6, 2014  
Revised June 3, 2014

## Table of Contents

### **PART B. COLLECTION OF INFORMATION EMPLOYING STATISTICAL METHODS.**

B.1. Respondent Universe and Sampling Methods.....	2
B.1.1. Selecting a Sample of Sites (Areas).....	2
B.1.2. Selecting Schools.....	5
B.1.3. Samples of Students and Their Parents and Teachers.....	6
B.1.4. Evaluating Mathematics Items, Reading Assessment, and Socioemotional and Executive Functions Measures.....	8
B.2. Procedures for the Collection of Information.....	9
B.3. Methods to Maximize Response Rates.....	10
B.4. Test of Procedures or Methods.....	11
B.5. Individuals Responsible for Study Design and Performance.....	11

### **Tables**

Table 1. Sample of Sites and Schools for MGLS Field Test.....	5
Table 2. Target Observations and Sample Sizes for Various Instruments.....	7
Table 3. Distribution of Sample of Students with Disabilities <sup>a</sup> .....	7

### **Appendices**

Appendix A: Recruitment Letter for School Districts/Diocese
Appendix B: Recruitment Letter for School Administrators
Appendix C: Frequently Asked Questions
Appendix D: Field Test Summary
Appendix E: Guide for Recruitment Calls
Appendix F: Project Staff Confidentiality Agreement

## **PART B. COLLECTION OF INFORMATION EMPLOYING STATISTICAL METHODS**

The Middle Grades Longitudinal Study of 2016–2017 (MGLS:2017) aims to obtain an understanding of students’ development and learning that occurs during the middle grade years, and the school and non-school influences associated with better mathematics and reading success, socioemotional health, and positive life development during middle school and beyond. The study will follow students as they enter and move through the middle grades (grades 6–8), starting with a nationally representative sample of 6<sup>th</sup> grade students in spring 2017, and annual follow-ups in spring 2018 and spring 2019, when most of the students in the sample will be in grades 7 and 8, respectively.

### ***B.1. Respondent Universe and Sampling Methods***

The respondent universe for the MGLS:2017 field test includes all students in grades 5–8 attending public or private schools in the United States (with sampling restricted to the contiguous United States). The national MGLS:2017 will sample grade 6 students at baseline and follow them through grades 7 and 8 (and possibly beyond). However, the field test must assess the instruments, measures, and items to be used on students in each grade (6–8) simultaneously. In addition, because of the anticipated range of mathematics knowledge and skills of sixth graders, we will use a sample of grade 5 students to help guard against floor effects. The field test also includes an oversample of students in grades 6–8 with certain types of disabilities: (1) emotional disturbance and (2) autism, with a target of 150 mathematics and reading assessments from students with these disabilities, including those that would occur in the course of sampling all 6-8 grade students. We also will target students with specific learning disabilities, but will have enough observations in this category so that oversampling them is not required. We will select additional sample of the students with emotional disturbance and autism to meet the target of 250 for the three groups.<sup>1</sup> We estimate that the additional sample will include 125 assessments beyond those that would occur without oversampling.

In addition to students, the respondent universe for the field test, and ultimately the main study, includes parents and teachers of students, and school administrators. These other groups of respondents will be linked to students selected for the MGLS:2017 field test.

We will select the sample for the field test in three stages (discussed in detail below). The stages will include purposively selected “sites” (geographic areas, usually comprising one or two counties or county equivalents), schools sampled with known probability within the sites, and students sampled randomly within the selected schools. We will collect data from teachers and parents linked to the sampled students. As discussed below, we expect to conduct the field test in 50 schools in five sites and collect data from 4,075 students ( 3,725 from those in grades 6–8, including an oversample of students with specific disabilities to allow us to meet the target of 250 for that group,<sup>2</sup> and 350 in grade 5), 600 teachers, 819 parents, and 50 school administrators. The initial samples of sites, schools, parents, and teachers selected to participate will be larger than the number of observations needed, to allow for replacement of sites that do not meet criteria (specified below) and nonresponding or ineligible units at other levels.

#### **B.1.1. Selecting a Sample of Sites (Areas)**

We will conduct the MGLS:2017 field test in 5 sites within the contiguous United States, a number that will allow for diversity in the sample but ensure the cost-effective collection of data. Given that we will not use the field test to make statistical estimates about a larger population, we can limit the number of sites and eliminate the need to adopt probability selection methods. Thus, we will select sites purposively rather than with probability sampling methods. We will select an initial group of 15 sites, of which 5 will be used for the field test. As discussed below, 10 of the sites will serve as potential replacement sites. Although we will use purposive selection, the definition of the sites will resemble that of the primary sampling units (PSUs)

<sup>1</sup> The target number of observations is at least 100 for all groups except for emotional disturbance; the target for this group is 50.

<sup>2</sup> The number 3,725 includes 3,600 from the basic sample, plus 125 additional students needed to meet targets for two of the groups of students with disabilities (all except those in the specific learning disability group).

that could be used for the national study.<sup>3</sup> The 15 sites will be selected so that if any site does not participate, it can be replaced by one with similar characteristics.

We next describe the characteristics required of the five sites that will be used in the study. Each site will comprise one or more counties within the contiguous United States and the District of Columbia. Each site must include at least one regular public school district and a private school serving grades 6–8, and must have 10 or more schools serving grades 6–8.<sup>4</sup> To be counted toward site requirements, districts and schools must not be operated by the state or federal government or serve correctional facilities or other institutionalized populations, as these schools will not be included in the sample. Schools where at least 80% of the students enrolled in the school have special education individual education programs (IEPs)<sup>5</sup> also are excluded (the oversampling of students with disabilities will be implemented in regular schools). In forming sites, we will rely on the CCD classification of whether schools are special education schools<sup>6</sup> in determining whether the site has enough schools to meet minimum criteria.

In addition to the minimal requirements for each site, other requirements affect the sampled sites in the aggregate, including their geographic distribution, their characteristics related to the oversampling of students with disabilities, and other characteristics of the sites, and the schools and students located within them. The final set of sites will include one each from the west, midwest, and northeast census regions; one from the south region, excluding the Washington, DC metropolitan area; and one from the Washington, DC metropolitan area (the Washington-Arlington-Alexandria, DC-VA-MD-WV Metropolitan Division). Each of the sites with the exception of the Washington DC area site will be contained within a single state.

To facilitate the oversampling of students with disabilities, we will require that the sites as a whole contain enough eligible schools with a higher than average prevalence of students with the specific disabilities.<sup>7</sup> The reason for this restriction is that to meet the oversample targets, without increasing the number of schools in the field test sample, the schools in the sample must contain more students with these disabilities than would average size schools with average prevalence rates for these groups. Our determination of whether this criterion is met will be based on both the number of students in grades 6-8 and prevalence rates for the specific disabilities. In our discussion of the student sample, below, we assume prevalence rates are the same as national estimated averages and an average school size of 600 students in grades 6-8, but different combinations of school size and prevalence rates could also allow us to meet the oversample targets.

Other aggregate requirements concern characteristics (not related to the disability oversample) of the sites, their schools, and their students. Specifically, at least three of the sites will include one or more science, technology, engineering, and mathematics (STEM) (particularly mathematics magnet) public school or private school in order to have high-achieving students to test for ceiling problems with the mathematics and reading assessment items. Additionally, two of the sites will include a public district serving a rural area to

---

<sup>3</sup> If the MGLS:2017 uses a multistage design, with sampled schools clustered within geographic areas, the field test experience can inform the main study. Even if MGLS:2017 does not cluster schools geographically, the field test methods of defining sites will not add to costs or detract from data quality. The sample will include schools with different grade configurations, but we will not require the sites to include specific configurations.

<sup>4</sup> We expect that two or more of the five final sites will include charter schools. Given that we do not plan to oversample charter schools, we will not require potential sites to include them. We will obtain counts of public districts and schools from the Common Core of Data (CCD) and use the Private School Survey (PSS) for private schools. For each file, we will use the version that is the most up to date at the time of sample selection.

<sup>5</sup> This definition is derived from the CCD definition of a special education school.

<sup>6</sup> See Keaton, P. (2012). *Documentation to the NCES Common Core of Data Public Elementary/Secondary School Universe Survey: School Year 2010–11* (NCES 2012-338rev). U.S. Department of Education. Washington, DC: National Center for Education Statistics. Retrieved [date] from <http://nces.ed.gov/pubsearch/pubsinfo.asp?pubid=2012338rev>.

<sup>7</sup> The LEA-level (CCD) file has a variable SPECED that gives the district-level count of students with an IEP. The school-level CCD file has a variable TYPEXX (“XX” is the year); TYPEXX=2 designates a special education school. We will use these variables to estimate the numbers of students with disabilities not in special education schools. When we contact districts, we will ask them for school-level information on students with disabilities.

be sure that different locales are represented in the field test.<sup>8</sup> The five sites will also be required to contain at least 20 schools that serve grade 5 (we anticipate having a larger number, but this is the minimum), along with one or more of grades 6–8.<sup>9</sup> We will also select sites that in the aggregate include low-income areas and areas with concentrations of African American students, Hispanic students, and English-language learners (ELLs).<sup>10</sup>

To define sites, we will start with the CCD Census of Local Education Authorities (LEA) or school districts. We will supplement the LEA data as needed with information extracted from the school-level CCD (e.g., number of schools serving some combination of grades 5–8, broken down by specific grades served). We then will use the augmented LEA data file to construct a county file.<sup>11</sup> Next, we will supplement the county file with information on private schools from the Private School Survey (PSS), resulting in a count of the number of private schools with some combination of grades 6–8 in a county. We will rely on school district websites as well as other websites to identify STEM schools.<sup>12</sup>

Sites generally will include one or two counties. We will form sites by first determining which counties would meet the minimum criteria on their own (one public LEA, one private school serving grades 6–8, and a total of 10 schools serving grades 6–8). If possible, counties that do not meet the minimum criteria will be linked to a contiguous county, so the two counties together meet the minimum criteria. To contain costs, as a general rule, no site will contain more than two counties; however, if needed, we will make an exception to ensure rural representation. We will exclude from the sampling frame counties that neither meet minimum criteria nor lend themselves to linkage to a contiguous county to form a site that meets the site criteria.

To account for possible refusals, especially by school districts within the sites, we will select 15 sites, 3 each from the west, midwest, and northeast Census regions; 3 from the south region, excluding the Washington, DC metropolitan area; and three from this metropolitan area, so that any group of 5 then will meet the aggregate criteria discussed above (STEM; private school; schools serving 5<sup>th</sup> grade; urbanicity; ethnic composition; and presence of low-income households, students with disabilities, and ELL students). These sets of sites each will comprise a “triplet” of eligible sites. We will designate one site within each triplet as the primary selection and the other two as potential replacement sites if the level of initial cooperation with the site is so low that (after excluding noncooperative districts or private schools) the site cannot meet minimum site requirements specified above (i.e., having at least one regular public school district and a private school serving grades 6–8, and having 10 or more schools serving grades 6–8). Our approach to recruiting and assessing the level of cooperation within sites for those with multiple public districts will depend on the size of the districts. If there are multiple districts—each of sufficient size to ensure that the site would be large and diverse enough without the others (i.e., the site would meet minimum requirements), we will randomly select one district (the “main” district) to recruit and use the others as replacements if the main district in the site declines to participate. On the other hand, if all districts within a multiple district site are needed to meet minimum requirements, then the site would be replaced if any declined to participate.

### **B.1.2. Selecting Schools**

---

<sup>8</sup> The variable “ULOCAL” on the CCD defines 12 levels of urbanicity for each district, ranging from large cities to remote rural areas.

<sup>9</sup> To allow for replacements, as discussed below, we would prefer to have 30 such schools, but placing such a restriction may limit the available sites too greatly.

<sup>10</sup> We can obtain ELL status for public districts from the CCD, poverty status from district estimates made by the Census (SAIPE), and other characteristics from county estimates based on data from the American Community Survey.

<sup>11</sup> Using the variables CONUM and CONAME on the CCD, we can link each district to a county, based on the location of its administrative offices.

<sup>12</sup> Lists of STEM schools in each of the 50 states are available at [www.stemschool.com](http://www.stemschool.com).

Schools for the field test will be sampled within the final set of five sites. We will recruit 50 schools for the final sample. To allow for nonparticipation at the school level,<sup>13</sup> we will select an initial sample of up to 150 schools, at least 20 of which will serve grade 5 (the number may vary from site to site, as discussed below). All schools in the initial sample will serve grade 6. Not all will serve grade 5 or grades 7-8. This sample will be large and diverse enough so that from this larger initial sample we can recruit 50 schools, of which 10–12 will have 5<sup>th</sup> grade students and which are spread as evenly as possible over the five sites. Table 1 summarizes the sample selection of sites and schools. The number of recruited schools will be large enough to support the desired student sample size while ensuring representation of many groups of students.

**Table 1. Sample of Schools for MGLS Field Test**

School Type and Grades Served	Initial Sample	Contact	Recruited (desired yield)
Schools Type <sup>a</sup>			
Private	6	4	2
STEM	9	4	3
Other Public	135	64	45
Total	150	72	50
Grades Served			
Grades 5 and 6, but not grades 7–8	18	9	6
Grades 5–8	18	9	6
Total with grade 5	36	18	12
Grades 6-8, but no grade 5	114	54	38
Total	150	72	50

<sup>a</sup> Private and STEM can be exchanged for sampling purposes.

With a few exceptions, all schools in the selected sites serving grade 6 students will have a chance of being sampled. Schools with fewer than 24 students in each of grades 6–8 may be excluded. Also any schools will be excluded if they are operated by the federal government, serve correctional facilities or other institutionalized groups only, or are schools where at least 80% of the students enrolled in the school have special education individualized education programs (IEPs).

The final sample will include an average of 10 schools (8–12) per site. To select the samples of schools within sites, we first will construct a sampling frame for public schools (from the CCD supplemented with information provided by districts) and private schools (from the PSS). We will add information from other sources as needed to identify STEM schools. We will define strata before sample selection to ensure that the samples of schools have the characteristics needed to meet the field test requirements. Stratification variables will include status as a STEM public or private school, grade levels served, ethnic composition (e.g., percentage of minority enrollment), a poverty indicator (percentage receiving free or reduced-priced meals), and language use among the student body (e.g., percentage of students who are ELLs). The number of strata and their definitions will depend on the characteristics of the five sites.

We will use probability selection methods to select the samples of schools. We will use explicit strata (with oversampling where needed) to ensure representation of key groups of schools (e.g., schools with a grade 5, STEM schools and schools with large numbers of students with disabilities). Depending on site characteristics, we may use other variables to serve as explicit stratifiers or form implicit strata.<sup>14</sup> Within the explicit strata, we will select schools by using probability proportional-to-size (PPS) methods. To implement PPS selection, we will assign each school a measure of size (MOS) that will be its estimated

<sup>13</sup> Schools will be recruited within sites where at least one public school district (the largest) is participating. We expect that better than 75% of schools will participate once districts have agreed to be part of the field test. However, there will be some variation in participation from site to site. Also, there may be different levels of participation between public and private schools. The initial sample size needs to be large enough to allow for this variability.

<sup>14</sup> Explicit stratification involves defining groups (strata) within which samples of a specified size are selected. Implicit stratification, by which the frame is sorted on the stratifying variables before sampling, is used when systematic or sequential sampling methods are used to help ensure that the sample is proportionately distributed on certain characteristics.

number of students in grades 5–8. Sample selection will rely on probability minimum replacement methods (PMR, also known as the Chromy Method). The use of PMR will allow us to implement implicit stratification by using as sorting variables those school characteristics not already used to form explicit strata.

We expect the proposed design to yield a sample in which 4-5 schools out of 50 have a grade 6 but not grades 7 or 8.<sup>15</sup> Thus, in the full field test sample, the 7<sup>th</sup> and 8<sup>th</sup> grade student samples will be spread over 45-46 schools, resulting in a small increase (a projected two assessments per grade) in the number of completed 7<sup>th</sup> and 8<sup>th</sup> grade student assessments in each school that has these grades, compared to the number we would have obtained if all 50 schools in the sample served all three grades. (The sampling of students within schools is discussed in the following section.) We will give some types of schools (e.g., STEM schools) a higher probability of selection within a site to ensure their presence in the sample. Since the initial sample will include more schools in each site than we need to recruit, we will randomly assign some sampled schools to the main sample (to be contacted first) and others to the replacement sample in case of ineligibility or noncooperation. The plan will allow for replacement by schools with similar characteristics (e.g., where possible, a private school would be replaced by another private school, and a school with a 5<sup>th</sup> grade would be replaced by another with a 5<sup>th</sup> grade). We will attempt to replace non-cooperative schools within sites, but this may not always be possible (for example if the only private or STEM school in a site refused to participate.)

### **B.1.3. Samples of Students and Their Parents and Teachers**

The MGLS:2017 field test student population includes students in grades 5–8 as well as teachers and parents of those students. Normally we will sample students from rosters provided by the individual schools, linking teachers and parents to the students chosen. Table 2 presents the anticipated sample sizes of students, parents, and teachers needed to obtain the target numbers of assessments and questionnaires. We will select samples large enough to obtain mathematics and reading assessments for 4,075 students (a basic sample large enough to yield assessments with 1,200 students in grades 6, 7, and 8; 350 students in grade 5; and an additional 125 students with disabilities in grades 6–8 so that our target of 250 assessments of students with the specific disabilities can be reached). We anticipate a completion rate of 80% for non-disabled students and 62-63% among students in the disability groups, thus our initial sample will include 5,412 students. The overall sample to be selected (6,600 students) allows for lower than expected participation – either overall or in specific schools.

To insure a large enough sample of students with disabilities in our field test, we will have to add to the basic sample in some but not all disability categories. We expect the sample yield of 3,600 students in grades 6–8 to naturally include an adequate number of students with specific learning disabilities and not require an oversample of students with this disability. However, based on current prevalence estimates for the other two disability groups of interest, we do not expect a large enough number to naturally occur in our sample yield of 3,600 students from grades 6–8 for sound psychometric analyses. As a result, we will use school and/or district records to identify and sample additional students with disabilities. Table 3 presents the sample sizes for the three groups of students with disabilities, and information about the expected number of students that will need to be oversampled.

Finally, in addition to the direct-assessments that will be administered during the field test, we will also administer several survey instruments or parts of these instruments (e.g., student questionnaire, parent questionnaire, teacher questionnaire). To avoid excessive costs and respondent burden, the full student questionnaire will not be administered to every student; instead, a subsample of students (in grades 6–8) will complete sections of it. The goal is to administer sections of the student questionnaire to 819 students, including students with disabilities.<sup>16</sup> The parent questionnaire will be given to the parents of the same

---

<sup>15</sup> These schools also are likely to have 5<sup>th</sup> grade students, as schools with only a 6<sup>th</sup> grade are rare.

<sup>16</sup> This subsample will include 300 students with disabilities and 519 students without disabilities.

sample of students who complete the student questionnaire in order to allow us to examine the correlations between sets of items in the parent and student survey (e.g., associations between different student and parent measures of student socioemotional well-being). The goal will be to complete 819 parent interviews, including the parents of students with disabilities. In selecting the subsample we assume 90% response among non-disabled students and 80% in the disability oversample groups. As with the mathematics assessment sample, the total size of the subsample selected will allow for lower than expected levels of response. There are two dimensions of the teacher questionnaire: a teacher-level and student-level dimension. We will obtain student-level teacher questionnaires from up to 450 mathematics teachers for 819 students, including students with disabilities, and from 150 special education teachers for a total of 300 students across the five disability groups.<sup>17</sup> The teachers who complete the student-level survey will also complete a teacher-level questionnaire (either Mathematics or Special Education Teacher Questionnaire).

**Table 2. Target Observations and Sample Sizes for Various Instruments**

Number	Instrument	Initial Sample	First Release	Target (desired yield)
1	Math assessment	6,600	5,215	4,075
2	Reading assessments <sup>(a)</sup>	6,600	5,215	4,075
3	Executive functioning measures <sup>b</sup>	1,142	952	819
4	Student questionnaire <sup>b</sup>	1,142	952	819
5	Parent questionnaire <sup>(b)</sup>	1,142	952	819
6	Math teacher survey	600	500	450
7	Math teacher-student report <sup>(b,c)</sup>	1,142	952	819
8	Special education teacher survey	200	167	150
9	Special education teacher report <sup>(c)</sup>	400	333	300
10	School administrator questionnaire	150	64	50

<sup>(a)</sup> Same sample as math assessments.

<sup>(b)</sup> Will include 300 with disabled students (or their parents).

<sup>(c)</sup> Number of students reported about two per teacher.

**Table 3. Distribution of Sample of Students with Disabilities<sup>a</sup>**

Group #	Group	Schools	Target Students, Including Oversample	Prevalence Rate in Schools <sup>b</sup>	Average Population per School	Total Expected Population <sup>c</sup>	Initial Sample <sup>d</sup>	Expected Completion Rate	Initial Sample Yield <sup>d</sup>	Additional Students Needed for Oversample
Specific learning disability		50	100	0.066	39.6	1,980	339	65%	220	0
Emotional disturbance		50	50	0.006	3.6	180	31	40%	12	38
3		50	100	0.005	3.0	150	26	50%	13	87
		<b>50</b>	<b>250</b>		<b>46.2</b>	<b>2310</b>	<b>396</b>	<b>60%</b>	<b>245</b>	<b>125</b>

<sup>a</sup> Assumes that sample will include schools with an average of 600 students in grades 6–8.

<sup>b</sup> Data Source: Roberta Woods, NAEP, NCES, U.S. Department of Education, email document, 8/5/2013 in “Background on MGLS Disability Augmentation to Design Contract” e-mailed by Joseph Gibbs, U.S. Department of Education, Contracts & Acquisitions Management, December 11, 2013.

<sup>c</sup> For 50 schools.

<sup>d</sup> It is expected that the initial sample will also include 87 students with other health impairments (70 of which will have ADD/ADHD), 57 students with speech or language impairment and 82 students with other types of disabilities. The numbers of assessments expected for these groups are, respectively: 57 (46), 37 and 53.

For schools that agree to participate in the field test, we will select enough students to yield 24 completed student assessments in grade 6, and 26–27<sup>18</sup> completed student assessments in grades 7 and 8, respectively. A total of 1,200 assessments per grade will be completed, plus assessments for an additional 125 students with disabilities spread across these three grades, for a total of 3,725 assessments in grades 6-8. We plan to

<sup>17</sup> The actual number of teachers completing these instruments is difficult to estimate and will depend on how the student sample is distributed across the mathematics classes in the sampled schools.

<sup>18</sup> The slightly higher targeted number of 7<sup>th</sup> and 8<sup>th</sup> grade students in each school reflects the expectation that not all sampled schools will have a 7<sup>th</sup> and 8<sup>th</sup> grade (see Table 1), therefore in schools that do serve these grades, there will need to be a slightly larger sample of students taken to insure sufficient coverage on the aggregate across all schools.



sample students from school rosters rather than selecting classrooms and taking all or a sample of students from each sampled class. In schools with a grade 5, we will select enough grade 5 students to obtain a total of 350 completed student assessments (an average of 29-30 assessments in the expected 12 schools with 5<sup>th</sup> graders). Thus we expect to administer 4,075 assessments in the field test. From among the 3,725 students in grades 6–8, we will randomly select a subsample large enough to yield a total of 819 completed student surveys, executive functioning measurements, and parent surveys (table 2).

The actual method for sampling students will depend on whether records are available through the individual school or through, for example, a school district office, and whether students with the specified disabilities are identified and their disability status indicated on a list containing all students. For example, if all 6<sup>th</sup> grade students in a school are identified on one list but disability status is on another, we would select the initial random sample of 6<sup>th</sup> grade students (enough to yield 24 assessments at an 80% response rate plus have a reserve sample to account for a lower response among students with disabilities and perhaps a lower than average response at this school) from the full list of students. The initial sample would be checked against the list of students having disabilities and we would then randomly sample additional students in the specific disability categories to allow us to meet our targets for those groups.

Between 300–500 observations will be adequate to estimate the item characteristics (using one- or two-parameter models) for scales within the questionnaires (de la Torre and Hong 2010; Linacre 1994; Hulin et al. 1982). Therefore, of the 819 completed student surveys, executive functioning measurements, and parent surveys, 300 will be students with disabilities (and their parents), and the remaining 519 will be students without disabilities (and their parents).

We will select a sample large enough to collect (online or by paper) mathematics teacher-reported information for 819 students. We will ask these teachers to complete the teacher-level and student-level parts of the Mathematics Teacher Questionnaire. With one exception, teachers will be asked to complete only three (3) student-level teacher questionnaires regardless of the number of field test children they have in their class, in order to limit the respondent burden these individuals have from the field test. As the exception to this rule, approximately 15 grade 6 teachers across five sampled schools will be asked to complete the student-level teacher questionnaire for all sampled students in their class, allowing us to test the timing and teacher burden tolerance that could be expected in the national study.

In order to examine the associations between items and scales in the student, parent, and teacher surveys, and their associations with student performance on the executive functioning measures, and the mathematics and reading assessments, a single subsample of 819 students (and their parents and teachers) will be chosen to complete this set of instruments. Thus, for this sample of students we will have full information about their performance on the different assessments and measures, and the responses of their parents and teachers. The sample will include both students with and without disabilities.

#### **B.1.4. Evaluating Mathematics Items, Reading Assessment, and Socioemotional and Executive Functions Measures**

We will use the field test to identify mathematics items with the strongest psychometric properties and those covering the range of difficulties we expect to encounter in the national study. Based on the findings from the field test, we will construct a two-stage adaptive mathematics assessment for the national MGLS:2017. We will also use the field test to evaluate an assessment of students' reading achievement, several different executive functions measures, and items and scales found in the student questionnaire, including a number of scales that measure different aspects of students' socioemotional well-being (e.g. internalizing problems, conscientiousness, theories of intelligence, hope, and motivation).<sup>19</sup> The goal is to use the field test to

---

<sup>19</sup> The field test will evaluate an assessment of students' reading achievement that is being developed specifically for use in the MGLS:2017. We will examine how the items individually and as a set work for grade 6-8 students in general and for students with different types of disabilities. A subsequent data collection OMB package will address these other uses of the field test.

identify the assessments, measures, and items that will be fielded in the national study and guide any needed adjustments prior to the launching of that study. With about only 60-90 minutes of a student's time available, we will use a spiral design to accomplish our several goals.

For a final two-stage adaptive mathematics assessment with a routing test and three to four second-stage forms, we will need roughly 60–75 items. We will test at least twice as many items as we will need for the full scale collection (approximately 150). We will group items into multiple forms, with some overlap, so that we have at least 400 observations per item, and can place the items on a single scale.

For the executive function tasks, we are targeting a yield of 819 cases (table 2). Each student will receive two of the four executive function measures.

## ***B.2. Procedures for the Collection***

Recruitment in preparation for the spring 2015 field test will begin with the sample of up to 15 sites, a roughly equal number of school districts, and 150 schools. We will assemble an experienced recruitment team; recruiters will participate in a day-long training covering the specific goals, requirements, and timeline of the MGLS:2017 field test, and review all instruments (types of questions being asked, of whom, and how they are asked), planned data requirements and incentives, and planned procedures. Training scenarios developed for in-class practice will prepare recruiters for additional questions they may encounter, detecting areas of confusion and reasons for hesitancy on the part of districts, schools, and gatekeepers and developing strategies for addressing reluctance and keeping the door open. Only recruiters experienced in recruiting districts and schools will be used. Before contacting districts, they will sit in on a call placed by the recruitment lead who directs the training and will initiate the first district recruitment call. The recruitment lead will in turn sit in on each recruiter's initial conversation with a district. This process will be repeated for school recruitment calls. A post-call debrief will provide an opportunity to discuss the contact, issues that arose, next steps, and strategies to strengthen future contact attempts.

**District recruitment.** Once training has been held, district recruitment will begin. The first contact to districts (and with private school administrative bodies, if required) will be in the form of a recruitment packet containing a letter from NCES requesting permission to recruit schools in the district to participate in the field test (see Appendix A). The packet also will include a summary of the field test and a set of frequently asked questions (Appendices C and D, respectively). The letter will indicate that a study representative will follow up in a few days to confirm permission to contact schools directly (Appendix E is a guide for the follow-up call to districts).

A few days after districts receive the recruitment packet, recruiters will make the initial personal contact to districts by telephone. During the call, the recruiter will respond to questions and offer solutions for challenges the district might pose (e.g., concern about conflicts with existing district testing schedule). The recruiter will also use this time to confirm that selected schools are operational and contain the grade levels required for the field test (grades 6–8), and will ask the superintendent or someone in the district office to confirm the email address for the school administrator in each sampled school. Members of the recruiting team will also initiate any formal applications for clearance to conduct research in the district, where required. Once permission to contact schools has been obtained, the recruiter will ask the district for a letter of permission and support that can be shared with school administrators.

**School recruitment.** Once district approval has been obtained, an information packet will be mailed to school administrators. The information packet will contain a letter from NCES requesting participation in the field test (Appendix B) along with frequently asked questions and a study summary (Appendices C and D), and where available, a letter of support from the school district (or from a diocese or other administrative unit for parochial and charter schools where applicable). A recruiter will follow up with a personal telephone call to discuss the data collection requirements, school and respondent incentives, and general time line. The recruiter will work with the school to encourage participation.

During school recruitment, recruiters will ask principals to designate a school coordinator. The school coordinator will work with the study staff to help gather current class rosters and other key information needed for sampling and assessment administration; identifying students with an IEP and the corresponding type of disability; distributing and collecting parent consent forms; and working with the data collection team to schedule testing, and assisting the study team in encouraging teachers and administrators to complete their survey instruments.

### ***B.3. Methods to Maximize Response Rates***

We will likely face many challenges in obtaining the cooperation of districts and schools. Loss of instructional time; competing demands, such as district and state testing requirements; spring break schedules; lack of teacher and parent support; and increased demands on principals all can seriously impede gaining permission to conduct research in schools. Our recruitment teams will be trained to communicate clearly to districts, dioceses, private school organizations, and schools the benefits of participating in the field test, and what participation will require in terms of student and staff time. Recruiters will be trained to address concerns that districts and schools may have about participation, while simultaneously communicating the value of the study and its key role in developing instruments that ensure high quality data focusing on middle grade students. Early engagement of districts and school administrators will be key. The study will also offer monetary incentives to schools (as discussed in Part A), which have proven effective in increasing response rates. In summary, along with offering monetary incentives, our plan for maximizing district and school administrator engagement includes the following:

- **Experienced recruiters.** The recruiting team will include staff with established records of successfully recruiting school districts and schools. To maximize district approval, senior staff will make the initial district telephone contacts. Their familiarity with the study and its future impact, as well as their experience in working with districts to overcome challenges to participation, will be crucial to obtaining district approval. Recruiters contacting schools will be equally adept at working with school administrators and providing solutions to many challenges associated with student assessments, including scheduling and space conflicts, minimizing interruption of instructional time, and obtaining teacher and parent buy-in.
- **Persuasive materials.** Key to the plan for maximizing response are developing and sharing informative materials, and making requests to participate both professional and persuasive. The importance of the study will be reflected in the initial invitations from NCES (Appendices A and B) sent with a comprehensive set of frequently asked questions (Appendix C) and a study description (Appendix D). Beginning recruitment as soon as OMB clearance is obtained will allow districts and school administrators enough time to review study materials, giving them a full understanding of the study's value, the importance of the field test, and the data collection activities required by the field test. A full understanding of these factors will be important, both to obtain cooperation and ensure that schools and districts expect the data collection requests that will follow.
- **Buy-in and support at each level.** During district recruitment, the study team will seek not just permission to contact the schools but, ideally, support from the district. This may take the form of approval of a research application or a letter from the district superintendent encouraging schools to participate. Active district support encourages school cooperation. The same is true when private school organizations are involved. Similarly, when principals are interested in the research activity, they are more likely to encourage teacher participation and put forward an effective school coordinator.
- **Avoiding refusals.** Recruiters will work to avoid direct refusals by placing the focus on strategies to solve problems or meet the challenges faced by the district or school administrator. They will endeavor to keep the door open while providing additional information or seeking other sources of persuasion.

#### **B.4. Tests of Procedures or Methods**

A main goal of district and school recruitment for the MGLS:2017 field test is to obtain the cooperation of enough schools in order to sample and collect data from students, parents, teachers, and administrators in order to evaluate the performance of assessment items and measures and questions in the survey instruments (e.g., in parent or mathematics teacher questionnaire). In addition, the field test will be used to evaluate procedures for conducting the main study, including recruitment methods for obtaining district participation. Of particular interest are procedures for identifying, sampling, and recruiting schools that enroll different groups of students with disabilities, and the role that districts may play in facilitating this.

#### **B.5. Individuals Responsible for Study Design and Performance**

The following individuals are responsible for the study design, district and school recruitment, and the collection and analysis of all field test data. Carolyn Fidelman is the Federal Contracting Officer's Representative (COR) leading the MGLS:2017 design effort. Carol Pistorino (Decision Information Resources) and Jerry West (Mathematica Policy Research) are co-project directors, leading a study design and field test team that includes John Hall (senior statistician); Sally Atkins-Burnett (mathematics, executive functions, and socioemotional assessment); Liza Malone (survey instruments); John Sabatini (reading assessment); Michelle Najarian and Donald Rock (psychometric analyses); and Sylvia Epps and Sheila Heaviside (survey operations).

#### **REFERENCES**

- de la Torre, J., & Hong, Y. (2010). Parameter Estimation With Small Sample Size A Higher-Order IRT Model Approach. *Applied Psychological Measurement, 34*(4), 267–285.
- Hulin, C. L., Lissak, R. I., & Drasgow, F. (1982). Recovery of two-and three-parameter logistic item characteristic curves: A Monte Carlo study. *Applied Psychological Measurement, 6*(3), 249–260.
- Linacre, J.M., (1994). Sample Size and Item Calibration Stability.” *Rasch Measurement Transactions, 7*(4), 328.