

**PROGRAM FOR INTERNATIONAL STUDENT
ASSESSMENT (PISA 2018) FIELD TEST AND
RECRUITMENT FOR MAIN STUDY**

OMB# 1850-0755 v.18

SUPPORTING STATEMENT PART B

Submitted by:

**National Center for Education Statistics (NCES)
U.S. Department of Education
Institute of Education Sciences
Washington, DC**

**April 2016
Revised June 2016**

TABLE OF CONTENTS

B. COLLECTIONS OF INFORMATION EMPLOYING STATISTICAL METHODS

| | |
|--|---|
| B.1 Respondent Universe..... | 1 |
| B.2 Statistical Methodology..... | 1 |
| B.3 Maximizing Response Rates..... | 4 |
| B.4 Purpose of Field Test and Data Uses..... | 5 |
| B.5 Individuals Consulted on Study Design..... | 5 |

B. COLLECTIONS OF INFORMATION EMPLOYING STATISTICAL METHODS

B.1 Respondent Universe

PISA 2018 assesses students nearing the "end of their compulsory school experience." For international comparability, this is defined as students who are 15 years old, in grades 7 or higher. A range of exact birthdates is specified by the international coordinating committees based on the months in which the data will be collected. However, students must be between the ages of 15 years and 3 completed months and 16 years and 2 completed months at the beginning of the testing period. The universe for the selection of schools is all types of schools in all states of the 50 United States and the District of Columbia. Within sampled schools, students will be selected for participation by drawing a random sample among the 15-year-old students. For the Puerto Rico sample, the universe for the selection of schools is all types of schools, both public and private. The same age-range criteria for eligible students will be implemented in Puerto Rico.

B.2 Statistical Methodology

The Technical Standards for main study PISA 2018 established by the international governing board include the following:

Standard 1.8 The student sample size for the computer-based mode is a minimum of 6,300 assessed students, and 2,100 for additional adjudicated entities, or the entire PISA Defined Target Population where the PISA Defined Target Population is below 6,300 and 2,100 respectively. The student sample size of assessed students for the paper-based mode is a minimum of 5,250.

Standard 1.9 The school sample size needs to result in a minimum of 150 participating schools, and 50 participating schools for additional adjudicated entities, or all schools that have students in the PISA Defined Target Population where the number of schools with students in the PISA Defined Target Population is below 150 and 50 respectively. Countries not having at least 150 schools, but which have more students than the required minimum student sample size, can be permitted, if agreed upon, to take a smaller sample of schools while still ensuring enough sampled PISA students overall.

Standard 1.10 The final weighted school response rate is at least 85 percent of sampled eligible and non-excluded schools. If a response rate is below 85 percent then an acceptable response rate can still be achieved through agreed upon use of replacement schools.

Standard 1.11 The final weighted student response rate is at least 80 percent of all sampled students across responding schools.

Standard 1.12 The final weighted sampling unit response rate for any optional cognitive assessment is at least 80 percent of all sampled students across responding schools. In addition, NCES has a standard in which student response

rate should be at least 85 percent, and the sampling design described below is based on that rate.

Overview

The design for this study will be self-weighting, stratified, consist of two stages, and will use probability proportional to size (PPS). There will be no oversampling of schools or students. Schools will be selected in the first stage with PPS and students will be sampled in the second stage yielding overall equal probabilities of selection.

Target Populations

The national PISA target population is 15-year-old students attending education institutions located within the United States in grades 7 and higher. The target population for Puerto Rico is the same. The plan is to implement the main survey in the fall of 2018, with a field test in the spring of 2017. The specific definition of age eligibility that will be used in the survey is "...between 15 years and 3 (completed) months to 16 years and 2 (completed) months at the beginning of the testing window."

Sampling Frame of Schools

The population of schools for PISA 2018 is defined as all schools containing any 15-year-olds in grades 7 through 12. As in previous PISA cycles, the school sampling frame will be developed from the most up-to-date NCES Common Core of Data (CCD) and Private Schools Survey (PSS) datasets. For the national PISA 2018 field test, we will use the school sampling frame prepared for the National Assessment of Educational Progress (NAEP) 2017 which uses the 2014-2015 CCD and the 2013-2014 PSS school data. We will avoid, to the degree possible, NAEP, TALIS, and TIMSS which will be in the field in high schools in the 2016-2017 school year. For the Puerto Rico field test, we will utilize a school sampling frame prepared by the Puerto Rico Department of Education (PRDE).

The grade structure of the school is a key stratification variable designed to reduce sampling error, but this is especially so in PISA because data analyses have shown that achievement is highly related to grade. Other stratification variables may include public/private, region of the country, location (urban/suburban/town/rural, etc.), and enrollment by race/ethnicity.

Field Test Sampling

International standards do not require a formal probability sample of schools for the PISA field test. It is sufficient that the samples of schools be representative of a broad range of schools from across the United States (and Puerto Rico). The

national field test requires a minimum student sample of 2,400 students. The United States plans to select a sample of 70 schools each with two substitute schools, with the expectation that 70 schools will ultimately participate, to provide for an adequate participating student sample. Among the 70 schools, 64 will be public schools and 6 will be private schools. This allows for school and student non-response and also for school level and within-school level exclusions. In Puerto Rico, the field test requires a student sample of 2,100 students. Puerto Rico will select a sample of 60 schools each with two substitute schools, with the expectation that 60 schools will ultimately participate. Among these 60 schools, 54 will be public schools and 6 will be private schools.

The KeyQuest sampling software provided by the consortium will be used to select the student samples in each school. The target cluster size of students per school will be 50 students with the goal of assessing at least 42 students per school (after refusals and student ineligibility). The target cluster size of 50 includes 42 students sampled for the core PISA assessment and 8 students sampled for the financial literacy assessment.

To obtain a school sample that is broadly representative of schools across the United States, we will target a convenience sample of schools with grade 9 and above and enrollment of at least 50 students in grades 9 and 10 (where most 15-year-olds are found) excluding schools with grades 7 and 8 only, small public schools, and schools sampled for other educational studies in 2017 (such as TALIS). We will use the sample stratification characteristics used in previous PISA cycles including census region, locality (city/urban fringe/town/rural MSA), school type (public/private), grade span, and minority enrollment. The sample will be a stratified systematic sample, with sampling probabilities proportional to measures of size, where the measure of size is the estimated number of 15-year-olds. The school sample for Puerto Rico will be selected in a similar manner.

Field Test Instrumentation and Design

Cognitive Items. There are a total of 66 forms in the field test containing 4 clusters for reading, mathematics, science, and global competence, which will be administered in a 2-hour session. Students will receive one form with a combination of clusters depending on the form. The forms are combined and organized in three distinct groups in order to ensure adequate coverage of newly developed items and to examine the psychometric properties of the items. There is an expectation to have multistage adaptive testing for the main study (reading component only) in 2018. The field trial design will thus include variable unit positioning within clusters and will investigate the effects of variable unit positioning versus fixed positions in preparation for the main study, the hypothesis being that item parameter invariance is only supported when using intact clusters.

Cognitive items to be administered in the field trial consist of the following subjects and number of clusters (groups of items/units):

Science = 6 trend clusters

Mathematics = 6 trend clusters

Reading = 6 trend clusters; 12 new clusters

Global Competence = 4 new clusters

The field test assessment design utilizes 6 trend clusters and 12 new clusters of reading items, 6 trend clusters each of science and math, and 4 new clusters of global competence items. These clusters are organized in a rotation within three groups of students. Within a school, sampled students will be assigned to each of the three groups.

Group 1 will receive 2 trend clusters of combinations of science and mathematics, mathematics and reading, or science and reading. These clusters will be fixed unit order. Data from this group will describe the degree of invariance between 2015 and 2018 in reading and mathematics. The science items, which use trend science items, will provide information about variability between 2015 and 2018 and the impact of different ordering in 2018. Group 1 is expected to yield 128 responses per item.

Group 2 will receive new and trend items in reading. These items will apply variable unit ordering within clusters. The design will provide variations in unit ordering within clusters and can be examined relative to Group 1. Each of the 24 Group 2 forms contains a combination of one to six trend reading clusters and 3 of the 12 new reading clusters. Every trend cluster will be paired with a new cluster once and appears once or twice in each position. The design is expected to yield 108 responses per trend item and 162 responses per new item.

Group 3 contains new reading clusters and is based on a fixed order of units to provide a basis of comparison to the varying unit orders in Group 2. Each form will be administered to 32 students (768 students total).

For countries administering the Global Competence option, as in the case of the United States, these items are included in Group 3 with the new reading clusters.

Financial Literacy

The United States is again participating in the optional financial literacy assessment in 2018. In PISA 2015 students were subsampled to participate in financial literacy from the core assessment group within each school and these subsampled students returned for an additional hour of financial literacy. The 2018 assessment design is similar to the one used in PISA 2012, however, when an expanded sample was used to assess financial literacy in the same session that mathematics, science, and reading were assessed. That is, students sampled for financial literacy in 2018 will not be required to return for a second session. Approximately 8 students will be selected to participate in the financial literacy assessment in addition to 42 students selected for the core assessment. Each student sampled for financial literacy will receive two clusters of math or reading and two clusters of financial literacy and will also be asked to respond to a set of financial literacy-specific background

questionnaire items. The design for financial literacy is based on a yield of 384 assessed students. These students will be selected separately from the main assessment, but administered the assessment in the same session as the students taking the main assessment and will be included in Group 1 (see explanation above). The financial literacy instrument will contain 3 clusters with trend items from 2012 and 2015 as well as new interactive items.

Background Questionnaire Instruments. The questionnaires have been developed to address the questionnaire framework developed for PISA 2018. The framework defines 14 modules across the school, student, and teacher questionnaires comprising student background characteristics, teaching and learning practices, professional development of teachers, school governance, and non-cognitive/metacognitive constructs dealing with reading-related outcomes, attitudes, and motivational strategies. In addition, the questionnaires include items that have been included in multiple cycles of PISA, allowing the investigation of patterns and trends over time.

School questionnaire. A representative from each participating school will be asked to provide information on basic demographics of the school population and more in-depth information on one or more specific issues (generally related to the content of the assessment in the major domain, which is reading in 2018). Basic information to be collected includes data on school location; measures of socio-economic context of the schools' student population, including location, school resources, facilities, and community resources; school size; staffing patterns; instructional practices; and school organization. The in-depth information is designed to address a very limited selection of issues that are of particular interest and that focus primarily on the major content domain, reading. For both the field test and main study, it is anticipated that the school questionnaire will take approximately 45 minutes. It will be available to respondents online.

Teacher questionnaire. The teacher questionnaire will be offered online and is estimated to take approximately 45 minutes to complete in the field test, with a goal of around 30 minutes in the main study. Within a school, a total of up to 25 teachers who are eligible to teach the modal grade (grade 10) will be selected. Up to ten teachers will be English (or, more precisely, English/Language Arts) teachers (teachers who are eligible to teach grade 10 in an English subject) and up to 15 teachers will be non-English/Language arts teachers (teachers who are eligible to teach grade 10, but in subjects other than ELA). The teachers and students data are not linked; that is, the teachers are not necessarily teachers of the sampled students. The sampling selection for teacher and students are independent of one another. The teacher questionnaire is used to gather school-level contextual information about the structural and process characteristics of schools from a teacher's perspective (e.g. teaching practices and learning opportunities in classrooms, leadership and school policies for professional development, vertical and horizontal differentiation of the school system) and will be analyzed alongside data received through the school questionnaire to provide a context for the student achievement scores.

Student questionnaire. Participating students will be asked to provide information pertaining primarily to the major assessment domain in 2018, reading. Information to be collected includes demographics (e.g., age, gender, language, race, and ethnicity); socio-economic background of the student (e.g., parental education, economic background); student's education career; and access to educational resources and their use at home and at school, which have been standard questions in PISA since the earliest rounds. Domain-specific information will include instructional experiences and time spent in school, as perceived by the students, and student attitudes towards reading. The goal is for the student questionnaire to take approximately 30 minutes to complete in the main study. In the field test there will be multiple forms of the questionnaire in order to try different items and item formats. The main study may or may not use multiple forms. The core student questionnaire in the field trial is expected to take approximately 30 minutes to complete.

Information and Communication Technology [ICT] Familiarity Module. The ICT questionnaire aims to examine students' ICT activities and domain-specific attitudes including access to and use of ICT at home and at school, students' attitudes towards and self-confidence in using computers, self-confidence in doing ICT tasks and activities; and navigation indices extracted from log-file data (number of pages visited, number of relevant pages visited). The ICT questionnaire for students is expected to take approximately 15 minutes to complete.

Financial Literacy (FL) Module. The FL questionnaire aims to examine students' experience with money matters, such as having savings accounts, debit or prepaid cards, as well as whether they have experienced financial-related lessons in their school careers. Many of the items in the FL questionnaire were previously administered in 2012 and 2015, with a handful of new items being piloted in the field trial. The FL questionnaire for students is expected to take approximately 15 minutes to complete.

Main Study

The international minimum number of completed assessments—for the core computer-based assessments in reading, mathematics, science, and global competence—is 6,300 students in 150 schools. In PISA, the United States typically assesses between 5,600 and 5,900 students in 165 schools when sampling 42 students per school. To achieve a larger number of students assessed in 2018 as well as to account for anticipated nonparticipation and student ineligibility, the number of students sampled within schools will be increased to 50 students because of the added sample for financial literacy (42 students for the core assessment + 8 students for financial literacy). Assuming the same response level as PISA 2015, the initial target is a total sample of about 285 schools to yield about 205 participating schools (assuming a 72 percent participation rate among schools). To achieve the target final school response rate, we will use replacement schools to complete the sample, as allowed under the international sampling standards.

The student-per-school target for the core assessment is at least 36 completed student assessments per school. Assuming a within-school response rate of 85 percent (rates were 85 percent in 2000, 82 percent in 2003, 91 percent in 2006, 86 percent in 2009, 89 percent in 2012 and 89 percent in 2015), the original sample size of students within schools will be 50. Should any states participate in the 2018 assessment, each state would have a sample of between 50 and 60 schools and 2,700 students to yield 2,430 assessed students.¹ Like the states, Puerto Rico will also have a sample size of between 50 and 60 schools and 2,700 students. As the main study plans for states and subnational jurisdictions is finalized, this information will be updated in the burden table.

Nonresponse Bias Analysis, Weighting, Sampling Errors

It is inevitable that nonresponse will occur at the school and student level. We will analyze the nonrespondents and provide information about whether and how they differ from the respondents along dimensions for which we have data for the nonresponding units, as required by NCES statistical standards. After the international contractor calculates weights, sampling errors will be calculated for a selection of key indicators incorporating the full complexity of the design, that is, clustering and stratification.

B.3 Maximizing Response Rates

Our approach to maximizing school and student response rates in the main study includes the following:

- Use of a fall test administration, to avoid major conflicts with state testing;
- Selecting and notifying schools at least a year in advance;
- Communicating with state officials early in the process and applying a more proactive approach with states by coordinating with NAEP State Coordinators to gain assistance with sampled schools;
- Assigning personal recruiters for specific schools;
- Incentives for schools, school coordinators, teachers, and students (see Section A9);
- Contact with schools and school coordinators at set intervals throughout the year preceding the assessment;
- A tentative plan for a summer conference for representatives from sampled schools several months before the data collection (main study only) to inform them about PISA and keep them engaged in the study; and
- Provision of school-level results on PISA.

These approaches are based on recommendations from an NCES panel and experience with previous PISA administrations.

B.4 Purpose of Field Test and Data Uses

Participation in the field test is an international requirement for participating in the PISA 2018 main study. The main focus of the field test is to collect enough

¹ This is based on an expected response rate in the states of 90 percent; response rate in the states has been historically slightly higher than the national sample. This is attributable to the fact that states recruit and manage the participation of the schools and the students.

assessment data to perform reliable tests of the items. However, during the field test, procedures for conducting the main study, including recruitment methods for obtaining school and student participation also will be evaluated. This information will be used to (a) determine the final main study design and in which international options the United States will participate, and (b) improve our recruiting strategies and materials for the main study.

B.5 Individuals Consulted on Study Design

Many people at OECD, ETS, and other organizations around the world have been involved in the design of PISA. Some of the lead people are listed in section A8. Overall direction for PISA is provided by Patrick Gonzales, the PISA National Project Manager at the National Center for Education Statistics, U.S. Department of Education.