

**OECD  
PROGRAM FOR INTERNATIONAL STUDENT  
ASSESSMENT  
(PISA 2012)**

**REQUEST FOR OMB CLEARANCE  
OMB# 1850-0755 V.10**

**SUPPORTING STATEMENT PART B**

**Prepared by:**

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

**Submitted: August 20, 2010**

TABLE OF CONTENTS

<b>B. COLLECTIONS OF INFORMATION EMPLOYING STATISTICAL METHODS.....</b>	<b>1</b>
B.1 Respondent Universe.....	1
B.2 Statistical Methodology.....	1
B.3 Maximizing Response Rates.....	6
B.4 Purpose of Field Trial and Data Uses.....	7
B.5 Individuals Consulted on Study Design.....	7

## **B. COLLECTIONS OF INFORMATION EMPLOYING STATISTICAL METHODS**

### **B.1 Respondent Universe**

PISA 2012 assesses students nearing the "end of their compulsory school experience". For international comparability, this is defined as students who are 15 years old, regardless of their grade in school. 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 test period. The universe for the selection of schools is all types of schools in all states of the 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.

### **B.2 Statistical Methodology**

The Technical Standards for main study PISA 2012 established by the international governing board include the following:

**Standard 1.8.** The student sample size must be a minimum of 4,500 assessed students, or the National Defined Target Population.

**Standard 1.9.** The school sample size must be a minimum of 150 schools or all schools that have students in the National Defined Target Population.

**Standard 1.10.** The target cluster size is typically 35 PISA eligible students, which upon agreement can be increased or reduced to a number not less than 20.

**Standard 1.11.** School response rates must be above 85 percent of sampled 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. PISA establishes three response rate zones—acceptable, intermediate, and not acceptable. "Acceptable" refers to original school response rates above 85 percent and means that the country's data will be included in all international comparisons. "Not Acceptable" refers to original response rates below 65 percent and means that the country's data will be a candidate for not being reported in international comparisons unless considerable evidence is presented that nonresponse bias is minor. "Intermediate" refers to original school response rates of

between 65 and 85 percent and means that a decision on whether or not to include the country's data in comparisons must be made while taking into account a variety of factors, such as student response rates, quality control, etc. In addition, schools with less than 50 percent participation of students are not considered participating schools and neither that school nor those students that did participate are considered in the calculation of response rates.

**Standard 1.12.** The overall student response rates must be above 80 percent of sampled students.

In addition, NCES has a standard in which student response rate should be 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 PISA target population is 15-year-old students attending education institutions located within the United States in grades 7 and higher. The plan is to implement the survey in the fall of 2012, with a field trial in the spring of 2011. 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 2012 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). For the PISA 2012 field trial, we will use the school sampling frame prepared for the National Assessment of Educational Progress (NAEP) 2011 which used the 2007-2008 CCD and PSS school data. We will avoid, to the degree possible, other NCES studies that will be present in high schools in the 2012-2013 school year.

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 Trial**

International standards do not require a formal probability sample of schools for the PISA field trial. It is sufficient that the samples of schools be representative of a broad range of schools from across the United States. The field trial requires a minimum student sample of 1,800 students. The United States plans to select a sample of 124 schools with the expectation that 80 will ultimately participate, to provide for an adequate participating student sample. This allows for school and student non-response and also for school level and within-school level exclusions.

The KeyQuest sampling software provided by ACER will be used to select the student samples in each school. The target cluster size of students per school will depend on the test option. For the P&P/CBA schools, the target cluster size is to sample a total of 42 students—34 for the standard P&P option and 8 for FL. Among the 34 standard P&P students, 21 will be selected for CBA followup. For the CBA only schools, the target is to sample 38 students to achieve an average of about 30 students per school for testing in two sessions.

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 42 students in grades 9 and 10 (where most 15-year-olds are found) excluding schools with grades 7 and 8 only, small schools, and any schools sampled for other educational studies in 2011 (such as NAEP, TIMSS, and PIRLS). 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 enrolment.. We will also use stratification to separate subgroups of schools by test option and control the target cluster size in each group. 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.

### **Field Trial Study of International Options**

One of the purposes of the field trial is to test the effects of the various international options under consideration on school and student recruitment and operations. Information gained from the field trial can be useful in weighing the value of the additional education data gained by the options against increases in cost and burden and the risk of not achieving acceptably high school and student response rates to report the data collected.

In particular, the data to be collected during the field trial should provide information to help NCES determine whether to administer both the paper-

and-pencil and computer-based mathematics and reading assessments (administering the computer-based assessments will require a second session for about half or more of the students) or to administer these assessments only in one mode, and whether to participate in the financial literacy assessment which will require an increase in the per school student sample. The possible scenarios for the main study are further described below. The field trial study could also inform how we approach and incentivize schools to participate.

The core field trial is to assess students with the paper and pencil test followed by computer-based assessments for a subsample of participating students. In the field trial, we will administer the PISA assessments to simulate the main study test environment so that NCES can decide if and how to participate in the new international options. Field trial experience will be used to assess the feasibility of administering computer-based assessments in the context of also administering pencil and paper assessments to the same students, and the efficacy of operational procedures. This information will be used to guide thinking on attainable school and student response rates if the options are implemented. Since there is no requirement for a field trial component that is just pencil and paper assessments, and since considerable additional resources would be required to add such a component (as it would have to be additional to all the other field trial assessment components), we will not use such an approach. Rather, experience from the field trial will be compared with those of the 2009 PISA assessments, and the corresponding 2008 field trial, to assess indirectly the impact that adding computer-based assessments has on school and student response rates. Additionally, a number of schools in the field trial will participate only in computer-based assessments. Although information about response rates for this group is not directly relevant to the issue at hand (since conducting computer-based only assessments is not an option for the main study), nevertheless we believe that this information will assist in ascertaining what are the key factors in determining school and student participation.

Burden rates are provided in part A.12 based on various scenarios for the field test (for students they would remain the same in all scenarios but burden rates would vary for school principals).

### **Main Study**

The international minimum number of completed assessments—for the core paper-and-pencil mathematics, science and reading assessment—is 4,500 students in 150 schools. The United States PISA sample is typically approximately 5,600 students in 165 schools. Assuming the same response level as PISA 2009, our initial target is a total sample of about 240 schools to

yield about 160 participating schools. To achieve the target final school response rate, we will use replacement schools to complete the sample.

The number of sampled students per school will depend on whether or not the United States participates in the financial literacy option. None of the other international options influence sample size (although they influence burden, see below and part A.12). The student-per-school target for the core assessment is at least 35 completed student assessments per school. The student-per-school target including financial literacy is 43. Assuming a within-school response rate of 85 percent (rates were 85 percent in 2000, 82 percent in 2003, and 91 percent in 2006), the original sample size of students within schools will be 42 for the core assessment only or 51 if financial literacy is included.

### Main Study Options

One purpose of the field trial is to determine in which international options the United States will participate. Also, internationally the feasibility of some of the components will be evaluated after the field trial. As such, which assessments will be administered in the main study is not known at the time this document is being submitted (May 2010), but we can provide the mostly likely possible scenarios. These are summarized in Table B-1.

**Table B-1. Possible scenarios for the administration of the main study**

	Paper-and-pencil mathematics, reading, and science	Computer-based problem solving	Computer-based mathematics and reading	Financial literacy (paper-and-pencil)
A. All components	X	X	X	X
B. Core components only	X	X		
C. Core + financial literacy	X	X		X
D. Paper-and-pencil assessments only	X			X
E. Paper-and-pencil assessments mathematics, reading, and science only	X			
F. Computer-based only + paper and pencil science	X (science only)	X	X	
G. Computer-based only + paper and pencil mathematics, reading, and science	X	X	X	

MSA All components (core plus all options): If we learn in the field trial that response rates will *not* likely be depressed by administering both paper-and-pencil and computer-based assessments (thus requiring two testing sessions for about half or more of the students) and a larger student sample per school because of the financial literacy

assessment, we would participate in all components offered in the main study (A). This would require student samples of 51 students per school and 6,800 overall for the country. All students would take the 2-hour paper-and-pencil assessment, which would include some financial literacy items. A subset of 18 students would participate in a second 40-minute session for computer-based assessment of reading, mathematics, or general problem-solving.

**MSB** Core components only: If the field trial study results suggest that the increased sample associated with the financial literacy assessment and the increased burden associated with the computer-based assessments of reading and mathematics would unduly raise cost, burden, or risk of not achieving acceptable school or student response rates, then we may opt to participate in only the core components. This would require student samples of 42 students per school and 5,600 overall for the country. All students would take the 2-hour paper-and-pencil assessment, which would *not* include any financial literacy items. A subset of 14 students would participate in a second 40-minute session for computer-based assessment of general problem-solving.

**MSC** Core components plus financial literacy: If the field trial study results suggest the increase in student sample associated with the financial literacy assessment would *not* unduly raise cost, burden, or risk of not reaching acceptable student or school response rates, but the increase in the number of students returning for the computer-based session associated with the computer-based assessments of reading and mathematics options would unduly raise one of these concerns, then we may opt for just the core components plus financial literacy. This would require student samples of 51 students per school and 6,800 overall for the country. All students would take the 2-hour paper-and-pencil assessment, which would include financial literacy items. A subset of 14 students would participate in a second 40-minute session for computer-based assessment of general problem-solving.

**MSD** Paper-and-pencil assessments only (including financial literacy): If the field trial study results suggest that the increase in student sample associated with the financial literacy assessment would *not* unduly raise cost, burden, or risk of not reaching acceptable student or school response rates, but the presence of a second session for computer-based assessments would unduly raise these concerns, then we may seek approval from the international PISA Governing Board to participate in only the paper-and-pencil assessments, and not participate in any computer-based assessments, including the general problem-solving assessment that is considered part of the 2012 PISA core. This would require student samples of 51 students per school and



6,800 students overall for the country. All students would take the 2-hour paper-and-pencil assessment, which would include financial literacy items.

**MSE** Paper-and-pencil assessments in reading, mathematics, and science only: If the field trial study results suggest that *both* the increase in student sample associated with the financial literacy assessment and the second computer-based session unduly raise cost, burden, or the risk of not reaching acceptable student or school response rates, then we may seek approval from the international PISA Governing Board to participate in only the paper-and-pencil assessments of reading, mathematics, and science, and not participate in any computer-based assessments, including the general problem-solving assessment that is considered part of the 2012 PISA core. This would require student samples of 42 students per school and 5,600 students overall for the country. All students would take the 2-hour paper-and-pencil assessment, which would *not* include financial literacy items.

**MSF** Computer-based assessments plus paper-and-pencil science only: If the field trial study results suggest the financial literacy and computer-based assessments unduly raise cost, burden, or risk of not reaching acceptable student or school response rates, an alternative to dropping the second computer-based session may be to seek approval from the international PISA Governing Board to drop the paper-and-pencil session. This would result in a single session that included computer-based assessments of reading, mathematics, and general problem-solving and a paper-and-pencil assessment of science (which is not offered in a computer-based version). This would require student samples of 42 students per school and 5,600 students overall for the country. All students would take a 70-minute combined computer-based (40 minutes) and paper-and-pencil assessment (30 minutes), which would *not* include financial literacy items.

**MSG** Computer-based assessments plus core paper-and-pencil assessments: If the burden caused by adding 8 students per school to the sample to undertake the Financial Literacy option is judged to outweigh the benefits of participating in this limited international assessment, then this component would be dropped from the study, while retaining the other optional components.

NCES is working on plans for methodological studies to examine the relationship between performance on PISA and other indicators. To facilitate tracing students for inclusion in future studies, NCES plans to obtain student

locator information during the main study (e.g., collecting state unique identifiers for each student from schools). No activities relative to this will take place during the field trial. The post-field trial change request to be submitted in spring 2011 will include specific information about what information will be collected and how.

### **Nonresponse Bias Analysis, Weighting, Sampling Errors**

It is inevitable that nonresponse will occur at both levels: school and student. 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 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 conflicts with state testing;
- Selecting and notifying schools at least a year in advance;
- Approaching schools directly, and notifying states and school districts;
- Assigning personal recruiters for specific schools;
- Incentives for schools, school coordinators, and students (see Section A9);
- Contact with schools and school coordinators at set intervals throughout the year preceding the assessment;
- 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 Trial and Data Uses**

Participation in the field trial is an international requirement for participating in the PISA 2012 main study. The main focus of the field trial is to collect enough assessment data to perform reliable tests of the items. However, during the field trial, 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 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. The possible designs to gather this information have been discussed.

### **B.5 Individuals Consulted on Study Design**

Many people at OECD, ACER, and other organizations around the world have been involved in the design of PISA. Some of these were previously listed in section A8. In addition, determinations about how to proceed in the main study, based on the field trial, will include the contractor awarded the PISA national contract (yet to be awarded). Overall direction for PISA is provided by Holly Xie and Dana Kelly, the PISA National Project Managers at the National Center for Education Statistics, U.S. Department of Education.