

**PROGRAM FOR INTERNATIONAL STUDENT  
ASSESSMENT (PISA) VALIDATION STUDY**

**REQUEST FOR OMB CLEARANCE  
OMB# 1850-NEW v.1**

**SUPPORTING STATEMENT PART B**

**Submitted by:**

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

**July 1, 2013  
Revised August 9, 2013**

# TABLE OF CONTENTS

---

<b>B. COLLECTIONS OF INFORMATION EMPLOYING STATISTICAL METHODS.....</b>	<b>1</b>
B.1 Potential Respondent Universe and Sampling.....	1
B.2 Methods for Maximizing Response Rates.....	2
B.3 Tests of Procedures and Methods.....	4
B.4 Individuals Consulted on Study Design.....	4

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

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

#### **1. Sample**

The sample for the main PISA Validation Study will consist of students from the national PISA 2012 sample who took PISA mathematics, reading, and science assessments and completed a Student Information Form providing their contact information. PISA 2012 recruitment materials, including materials for parents, stated that students supplying contact information may be contacted by NCES for a future study.

PISA assesses students nearing the "end of their compulsory school experience." For international comparability, the PISA target population 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 test period. The universe for the selection of schools in the PISA 2012 administration was all types of schools in all states of the United States and the District of Columbia. Within sampled schools, students were selected for participation by drawing a random sample of 50 students from among the age-eligible students (42 of these students were assigned to the mathematics, science, and reading assessment, and 8 to financial literacy).

To focus the study on the main PISA domains, this study will not include the PISA 2012 students who took financial literacy, except in the field trial. Also, in three states (Connecticut, Florida, and Massachusetts), additional samples of approximately 1,500 students per state took the PISA mathematics, science, and reading literacy assessment so that these states could obtain state-level PISA results (these three states provided funding for their supplemental samples). At this time, students in the state samples are not included in the Validation Study plans, although should states wish to partner with NCES to conduct a follow-up study with the students in the PISA 2012 state samples, the proposed study may be revised to include those students.

A total of 6,116 students were assessed in the national administration of PISA in 2012. Of these, approximately 5,810 students (95 percent) completed a Student Information Form. Among those, 1,081 were assessed in financial literacy, leaving 4,729 students to serve as the starting sample for the main PISA Validation Study.

We expect we will be able to locate 90 percent of the 5,810 students from the national PISA 2012 sample who took PISA mathematics, reading, or science assessments in fall 2012 and completed the Student Information Form providing us their contact information. Locating 90 percent of those students a year later includes responses to address updates from a portion of the sample, but also relies on intensive tracing, as described in the supporting statements, which will use cost-effective tracing techniques, such as calling directory assistance, calling contacts provided by the respondent, and using publicly available Internet searches. In line with other studies that employ locating of participants such as High School Longitudinal Study (HLS) and Beginning Postsecondary Students (BPS) Longitudinal Study, we estimate that between 15 to 20 percent of the 5,810 students will respond to the first follow-up request for address verification, and the remainder will have

to be located using tracing techniques. However, the majority of these students will still be in high school and should be relatively easy to locate. A limited number of schools may be contacted to request assistance in locating students who could not be located using all other methods.

It is expected that there will be a 10 percent loss of sample each year, resulting in 3,447 students for the 2015 study. The 1,081 students assessed in financial literacy will be traced and then a subsample of 200 students will be selected for the field test. Approximately 306 of the 6,116 students did not complete the Student Information Form. NCES will not pursue these students as part of this survey; by not completing the form they have implied that they do not wish to be contacted for future surveys.

## **2. Non-response Bias Analysis**

Non-response bias can potentially occur when respondents and non-respondents are different. A non-response bias analysis will compare participating and nonparticipating students on key characteristics that are sensitive to nonresponse bias (e.g., age/grade and sex). The analysis will have two aspects: (1) bivariate analysis, in which it is determined if there are significant differences between participating and nonparticipating groups on key characteristics, and (2) logistic regression analysis, in which the key characteristics predict participation and tests of the significance of regression coefficients are performed. Significant regression coefficients indicate possible nonresponse bias. We plan to conduct this analysis on two sets of characteristics: school level, from the PISA 2012 sampling frame, and student level, collected on participating students during the PISA 2012 assessment. Since the nonrespondents will be among students who participated in PISA 2012, we will have a very rich set of characteristics for both respondents and nonrespondents.

## **3. Weighting**

The PISA 2012 sample data file contains survey weights and replicate weights appropriate for the analysis of PISA 2012 data. However, these weights will not be fully appropriate for the analysis of the longitudinal data, since there will be additional nonresponse in the longitudinal data. Therefore, a new set of weights that will make adjustments for this additional nonresponse will be derived.

Since the additional nonrespondents will be among students who participated in PISA 2012, we will have a rich set of data concerning the characteristics of these nonrespondents. This information can be used to provide additional nonresponse adjustments that we can expect to be quite effective in limiting any additional nonresponse bias arising from the nonresponse to the longitudinal component.

A calibration, or generalized regression, approach to the nonresponse adjustments will be used. Because we will have the PISA 2012 data for all the students in the follow-up study, we can use such methods to adjust the PISA sampling weights to account for the additional nonresponse. In doing this, however, we must build into the replicate weighting procedures adjustments that account for the fact that results tabulated from the full PISA sample are themselves subject to sampling error — they are not population census figures, but estimates obtained from the PISA sample. Thus we will implement weighting adjustments that will match the respondent sample to the full PISA responding sample on a variety of student characteristics, such as gender, race/ethnicity, region, urban/rural, grade, socio-economic status, and PISA achievement. This will provide replicate weights that reflect the fact that the distributions of these characteristics have been obtained from the PISA sample and not a census of 15-year-old students.

## **B.2 Methods for Maximizing Response Rates**

### **1. Tracing and Tracking**

The first steps to obtaining a sample to participate in the Validation Study are tracing and tracking the sample.

Tracing participants is a multi-step process that involves creating a database that will pair student-provided information with on-file information for all cases and locating or establishing initial contact with each participant. Initial matching of student-provided contact information to students' PISA 2012 on-file information will provide validation of student ID, which is essential to linking students' past PISA performance to data collected for this study at age 18, and provide additional contact information than that collected on the forms during PISA 2012 (e.g., participant gender, month of birth, year of birth). Establishing initial contact is important for locating the sample, verifying and updating contact information, and building a relationship that yields high participation for data collection.

We anticipate that a significant level of work will be required to initially contact these students, locate those who have moved since their contact information was collected during the conduct of PISA 2012, and continue to maintain contact with these students as they graduate from high school and transition to college or the workforce. The PISA 2012 sample included students from 163 schools across 44 states, and although some schools are clustered within a few large states, most of the states had only a handful of schools in the sample. In addition, student clustering around their initially sampled school unravels as students change high schools, drop out of school, or graduate and transition to college or the workforce.

To accomplish this, a progressive multi-modal approach for tracing of students will be used. For students that provided email contacts on their Student Information Form, emails will be sent informing them of the PISA Young Adult Follow-up Study, and asking them to access a secure website to update their contact information. For those students that could not be reached via email, letters will be mailed with stamped return postcards for updating contact information. The mailing envelope will be stamped "Address Correction Requested" to collect address updates from the U.S. Postal Service. All returned mail (e.g., wrong address, no such address, undeliverable) will be flagged for further follow-up locating activities by field staff. Any updated address will be entered into a sample tracking database. Finally those cases that do not have a confirmed address or telephone number will be subjected to intensive tracing by the tracing staff to locate the remaining participants. Intensive tracing will use cost-effective tracing techniques, such as calling directory assistance, calling contacts provided by the respondent, and using publicly available Internet searches. A limited number of schools may be contacted to request assistance in locating students who could not be located using all other methods.

Once initial contact is established and contact information has been verified or updated, students will be provided with information about the study, including information about the study purpose, schedule, and contact information via a student welcome package.

After this initial contact is established, participant contact will be maintained using well-established tracking procedures until the data collection period begins in fall 2015. We will solicit for contact updates twice a year and provide study information when appropriate. Birthday cards will also be sent to all sampled students in their birthday month, and the envelope will be stamped "Address Correction Requested." These multiple mailings serve to obtain updated address information for the study and to identify cases that need tracing and locating activities before the trail grows cold. During the tracing activities, we will

request permission from the students to allow us to use texting as a contact method for future survey activities. See Appendix A for materials.

The National Health and Nutrition Examination Survey (NHANES) has been using texting as a mode of contact with minors aged 12 to 16 for two years. NHANES is a household survey that includes an interview and a medical screening. The target is to complete 5,000 examined participants per year. A subsample of persons are preselected for morning exams requiring participants 12 years and older to fast. The study is always trying to improve fasting compliance. In May 2011, NHANES added text messages for the morning appointments that require fasting, after which the NHANES fasting rate went up 2 percent. While NHANES cannot say the text messaging was a direct cause for increasing or helping fasting rates, as those giving permission to text may be a more compliant group, they have seen the fasting rates remain steady, found no evidence that texting negatively impacts rates, and while being inexpensive it may be helping to increase compliance. To implement the NHANES text message reminders, questions were added to the household questionnaire asking permission to text participants along with an explanation of possible fees and asking for the cell phone number of the participant if it has not already been provided. The PISA Validation Study plans to use a similar procedure in ascertaining whether to send text messages to participants.

## **2. Recruitment**

Technically, recruitment for the study begins with the first tracing activities, continuing over the course of tracking, and manifests itself in the final response rate. Though tracing and tracking procedures will continue throughout the study, there will be special emphasis and procedures unique to the period immediately preceding and throughout data collection. It is this time that can be thought of as the functional recruitment period. For example, extensive phone and other case-level contact may be necessary in the month prior to and during the data collection window in order to prepare students for the survey. It may be necessary to provide additional materials specifically targeted for recruitment purposes and disseminated during this important time. These materials will be provided in the next full package that will begin its clearance process in February 2014.

Another task specific to recruitment will be refusal or nonresponse conversion. The voluntary nature of participating in this study is protected, and participants may drop out of the study for any reason at any time. Though tracing and tracking procedures will minimize attrition to the greatest degree possible prior to data collection, actually getting respondents to complete the assessment and questionnaire will pose unique challenges. For example, it is feasible that participants may begin but not complete the instruments. To address this challenge, we will use a monitoring system so that timely follow-ups and intervention are possible. Telephone contact by trained field staff will be used to prompt participants to respond.

A different foreseeable challenge is motivating respondents to begin the instruments. Based on experience in PISA and other volunteer studies, a monetary incentive does well to provide this motivation. As in OECD's International Assessment of Adult Skills (PIAAC), a \$50 incentive (cash card) will be offered to respondents for completing the assessment and questionnaire (see Part A for more detail).

### **3. Data Collection Procedures**

The web-based version of the PIAAC data collection instruments – called the Education and Skills Online (ESO) assessment and questionnaire – will be administered during a 12-week period in fall 2015. Study participants will be invited to log on to a secure website during the data collection window to complete the assessment and questionnaire. Telephone contact will be used to prompt respondents to log on to take the assessment and questionnaire or remind respondents to complete the instruments. Telephone interviewers will also be able to facilitate the completion of the questionnaire. Following standardized processes, data collection quality will be maintained by evaluating data as they are collected, the generation and review of weekly sample production reports, and Westat processes used to build and maintain quality.

#### **B.3 Tests of Procedures and Methods**

The ESO assessment and questionnaire are being piloted by the OECD in 2013 and will, therefore, have established psychometric properties by the time they are used in this study. While it will not be necessary to conduct a trial of the instruments for the Validation Study, NCES will conduct a field trial of data collection procedures. We intend to trace students who participated in the financial literacy assessment in PISA 2012 and supplied contact information. These students are not the target of this study (the students who took the mathematics, science, and reading assessment are) but are available to draw for a trial of procedures. We intend to administer the trial to 200 students over a four-week period in fall 2014.

#### **B.4 Individuals Consulted on Study Design**

Many people been involved in the design of PISA and PIAAC. Some of the lead people are listed in section A8. Overall direction for the Validation Study is provided by Dana Kelly, of NCES, U.S. Department of Education.