

Program for the International Assessment of Adult Competencies (PIAAC)

Website: <https://nces.ed.gov/surveys/piaac/>

Updated: September 2020

1. OVERVIEW

The Program for the International Assessment of Adult Competencies (PIAAC) is a cyclical household study that has been developed under the auspices of the Organization for Economic Cooperation and Development (OECD). The first cycle of PIAAC consisted of three rounds of data collection between 2011 and 2017. PIAAC's second cycle is planned to begin in 2021. For cycle one's initial study in 2011–12 ("round 1"), adults were surveyed in 24 participating countries. Nine additional countries participated in PIAAC in round 2 in 2014, and five more countries participated in PIAAC in round 3 in 2017.

The United States conducted three rounds of data collection during PIAAC first cycle. The initial U.S. study for round 1 was conducted from August 2011 through April 2012 with a nationally representative sample of 5,000 adults between the ages of 16 and 65 (officially known as the U.S. PIAAC Main Study). In 2013–14, the United States conducted a second round of data collection in order to (1) supplement the 2012 U.S. household data collection (officially known as the "National Supplement") and (2) include individuals who were detained in state, federal, or private prisons housing state or federal inmates in the United States (officially known as part of the National Supplement, but having its own name: the "U.S. PIAAC Prison Study"). The household supplement was conducted from August 2013 through April 2014 with a sample of 3,660 adults in households in order to enhance the U.S. PIAAC 2011–12 dataset by (a) oversampling young adults (age 16-34), (b) oversampling unemployed adults (age 16-65), and (c) expanding the sample to include older adults (age 66-74). The U.S. PIAAC Prison Study was conducted from February through June 2014 with a nationally representative sample of 1,270 adult inmates (age 18-74) who were detained in 98 states and federal prisons in the United States. The third round of U.S. data collection was administered from March 2017 to November 2017 to a nationally representative household sample of about 3,660 adults between the ages of 16-74. These data, released in the fall of 2019, provide a second point in time to allow comparisons with the 2012/14 results. In addition, the combined 2017 and 2012/2014 data support the calculation of small area estimates at the county and state level, which were released in April, 2020.

The goal of PIAAC is to assess and compare the basic skills and the broad range of competencies of adults around the world. The assessment focuses on cognitive and workplace skills needed for successful participation in 21st-century society and the global economy. Specifically, PIAAC measures relationships between individuals' educational background, workplace experiences and training, skill-used at work and home, occupational attainment, income, health, use of information and communications technology, and cognitive skills in the areas of literacy, numeracy, and digital problem solving.

PIAAC is a complex assessment: the data collection was conducted in multiple languages and in numerous countries with diverse populations, cultures, education and life experiences. However, in order to make results comparable all participating countries (a)

SAMPLE SURVEY OF ADULTS AGES 16 TO 65 INTERNATIONALLY AND 16 TO 74 IN THE UNITED STATES:

PIAAC collected:

- Background information: demographic information, education, training, employment, and job requirements
- Direct assessments of literacy, numeracy, reading components, and problem solving in technology-rich environments

follow the quality assurance guidelines set by the OECD Consortium, (b) closely follow all the agreed-upon standards set for survey design, implementation of the assessment and the reporting of results, and (c) administer equivalent versions of all cognitive and non-cognitive instruments.

PIAAC builds on knowledge and experiences gained from previous international adult assessments including the International Adult Literacy Survey (IALS) and the Adult Literacy and Lifeskills Survey (ALL). PIAAC enhances and expands on these previous assessments' frameworks and, at the same time, improves upon their design and methodologies.

In the United States, the PIAAC assessment was conducted in English only; however, the PIAAC survey background questions were in both English and Spanish.

Purpose

The primary objectives of PIAAC are to (1) identify and measure cognitive competencies believed to underlie both personal and societal success, (2) assess the impact of these competencies on social and economic outcomes at individual and aggregate levels, (3) gauge the performance of education and training systems in generating required competencies, and (4) help to clarify the policy levers that could contribute to enhancing competencies.

One of PIAAC's core objectives is to assess how well participants use information and communications technology to access, manage, integrate and evaluate information; construct new knowledge; and communicate with other people. In addition, PIAAC collected information on participants' use of key work skills in their jobs, a first time for an international study. In this way, PIAAC offers a more complete and nuanced picture of human capital than earlier studies.

It is important that the participating countries share a set of survey objectives, to facilitate comparisons of survey results between countries. PIAAC assessment items and background questionnaires were designed to ensure cross-cultural, cross-national, and cross-language validity.

Components

In cycle one, PIAAC collected background information on adults before administering direct assessments of literacy, numeracy, reading components, and problem solving in technology-rich environments. After the interviewer-directed background questionnaire was administered, the PIAAC assessment design was to route participants to the appropriate delivery mode to assure the most reliable, valid, and comparable assessment of skills. Two delivery modes for the assessments were available: paper-and-pencil and computer-based. The computer-based mode was administered to participants who met all of the following

criteria, established for determining basic computer competence: (1) prior computer use, (2) willingness to take the assessment on the computer, and (3) passing a basic computer test (by successfully completing four of six simple tasks, such as using a mouse and highlighting text on the screen). Participants who did not meet any one of these requirements – who reported no computer use, who were unwilling to take the assessment on the computer, or who failed the basic computer test – were routed to the paper-based assessment.

The percentage of U.S. adults routed to the paper-based assessment varied in each of the three rounds of U.S. data collection: in round 1 (2011–12), approximately 16 percent of U.S. adults in the household study received the paper-based assessment; in round 2 (2014), approximately 22 percent of adults in the household study and 36 percent of adults in the prison study; and in round 3 (2017), about 16 percent of adults in the household study. The same survey procedures, processes, and assessment instruments were used in all three rounds.

Background Questionnaire for the Household Studies.

The background questionnaire for cycle one was developed to identify (a) what skills participants use in their job and in their home life, (b) how participants acquire those skills, and (c) how those skills are distributed throughout the population. Interviewers administered the background questionnaire, asking participants about their education and training; present and past work experience; the skills they use at work; their use of specific literacy, numeracy, and information and technology (ICT) skills at work and at home; personal traits; and demographic information.

In order to obtain background information from a wide range of respondents in the United States, the background questionnaire was administered in either English or Spanish (although the direct assessment of skills was only administered in English). For linking purposes, several items from IALS and ALL were included in the PIAAC background questionnaire.

Participating countries were allowed to add up to 5 minutes of country-specific items. The United States added questions focused on country of origin, language, race and ethnicity, training courses, and health related information. A majority of the U.S. country-specific questions were adopted from the 2003 National Assessment of Adult Literacy (NAAL) background questionnaire.

The household background questionnaire was identical for all three rounds of U.S. data collection in cycle one, except that in rounds 2 and 3 a few additional U.S. country-specific items were administered that were not in round 1.

Background Questionnaire for the Prison Study. While maintaining comparability between household and prison

instruments as closely as possible, adaptations to the household questionnaire were implemented in order to adequately capture information from the prison population. These included deleting questions from the household questionnaire that would be irrelevant to respondents in prison, such as questions on earnings from their current job, as well as the addition of questions that addressed respondents' specific activities in prison, such as participation in academic programs and English as a Second Language (ESL) classes offered in prison; experiences with prison work assignments; involvement in non-academic programs, such as life skills and employment readiness classes; and educational attainment and employment prior to incarceration.

PIAAC Assessment Tasks. PIAAC both paper- and computer-based assessment tasks were drawn from real-life situations that are expected to be of importance or relevance in different contexts. Tasks' contents and questions were intended to reflect the purposes of adults' daily lives across cultures, even if they were not necessarily familiar to all adults in all countries. While the background questionnaire was administered in English or Spanish, both paper- and computer-based assessments were administered in English only. Some assessment items from IALS and ALL were included in the PIAAC assessment to facilitate the linking between the three international assessments.

Paper-based Assessment. The paper-based assessment (PBA) began with a 10-minute core of literacy/ numeracy items in paper- and- pencil format. Participants who performed at or above a minimum standard on this core section were randomly assigned to either a 30-minute cluster of literacy items or a 30-minute cluster of numeracy items. After they completed those items, they received a 20-minute assessment of reading components. Participants who performed poorly on the paper literacy/numeracy core proceeded directly to the reading components booklet.

Computer-based Assessment. The PIAAC computer-based assessment (CBA) was only administered to participants who indicated having previous experience with computers in the background questionnaire interview and were willing to take the assessment on the computer. Participants were directed to a core section that was composed of two parts: a basic computer test (ICT core), which measures skills such as highlighting, and a literacy/numeracy core which measures basic skills within these domains. Participants who performed poorly in either of these core tests were switched over to the appropriate sections of the paper-and-pencil instruments. Participants who failed the ICT core proceeded to the paper-based assessment and took the paper-based literacy/numeracy core items. Participants passing the ICT core proceeded to the computer-based literacy/numeracy core. If they did not pass the computer-

based literacy/numeracy core, participants were routed directly to the reading components section of the PBA.

Participants who performed well on both parts of the computer-based core section were randomly routed to the computer-based literacy, computer-based numeracy, or digital problem-solving domains.

Adaptive Design

One of the unique aspects of PIAAC was the adaptive design of the computer-based assessment (CBA) within the domains of literacy and numeracy. The adaptive testing process meant that respondents were directed to a set of easier or more difficult items. The choice of assessment items for each participant mainly depended on an algorithm using a set of variables that included (1) the participant's level of education; (2) the participant's status as a native or non-native English-language speaker; and (3) the participant's performance in the CBA core (as well as their performance in the CBA module as they advance through the assessment). For the digital problem-solving domain, there was no adaptive process.

The key advantage of an adaptive design is to provide a more accurate assessment of participants' abilities, while using a smaller number of assessment items than a traditional test design in which respondents must answer all questions included in the test, from easiest to most difficult.

Periodicity

Internationally, PIAAC has been envisaged as a decennial survey. The PIAAC cycle 1 assessment in the United States was conducted between 2011 and 2017 with three rounds of data collection. Internationally, three rounds of data collection involved an expansion of the original 24 countries to an additional 14 countries in the subsequent rounds. PIAAC cycle 2 will be administered in 2021–22, with currently 33 countries, including the United States, taking part in the development

Data Availability

Information on the availability of data and schedule of releases for PIAAC can be found at https://nces.ed.gov/surveys/piaac/schedule_detailed.asp.

2. USES OF DATA

The skills that are assessed in PIAAC (literacy, numeracy and problem solving in technology-rich environments) are cognitive skills that provide a foundation for successful participation in society and the global economy. Understanding the level and distribution of these skills among the adult population in participating countries, as well as the ways in which such skills are developed and maintained, and the social and economic benefits for individuals, is important for policy makers and other

stakeholders. The PIAAC project seeks to provide evidence for the following policy questions:

- *How are skills distributed?* A comparison of skill levels, skill requirements, mismatches and investments in education and training across countries, and within countries across demographic categories, regions, sectors of industry, levels and fields of schooling.
- *Why are skills important?* The relation of skills to relevant labor market outcomes such as employment opportunities, earnings, job security, and skill utilization, as well as to other outcomes such as health, civic engagement, and social trust.
- *What factors are related to skill acquisition and decline?* The relation between various learning activities (i.e., education, training, informal learning activities) and skill acquisition. The relation of experiences at work, in education and everyday life to skill decline among older individuals.

3. KEY CONCEPTS

PIAAC is designed to assess adults in different countries over a broad range of abilities, from simple reading to more complex problem-solving skills. To do this, PIAAC defines four core competency domains of adult cognitive skills seen as key to facilitating the social and economic participation of adults in advanced economies: literacy, reading components, numeracy, and problem solving in technology-rich environments. All participating countries and regions are required to assess the literacy and numeracy domains, but the reading components and problem solving in technology-rich environments domains are both optional. The U.S. PIAAC assessment measured all four domains.

Literacy. Literacy is defined as the ability to understand, evaluate, use, and engage with written text to participate in the society, to achieve one’s goals and to develop one’s knowledge and potential.

Reading components. Reading components measures literacy skills of adults at the lower end of the literacy spectrum, focusing on elements of reading that are comparable across the range of languages (reading vocabulary, sentence comprehension, and basic passage comprehension and fluency).

Numeracy. Numeracy is defined as the ability to access, use, interpret, and communicate mathematical information and ideas, in order to engage in and manage the mathematical demands of a range of situations in adult life.

Problem solving in technology-rich environments (PS-TRE). Problem solving in technology-rich environments (also referred to as “digital problem solving”) involves using digital technology, communication tools and

networks to acquire and evaluate information, communicate with others and perform practical tasks. The PIAAC problem solving assessment focuses on the abilities to solve problems for personal, work and civic purposes by setting up appropriate goals and plans, accessing and making use of information through computers and computer networks.

4. SURVEY DESIGN

The PIAAC Consortium oversaw all international PIAAC activities of cycle 1 on behalf of the OECD and provided technical support to all participating countries regarding all aspects of PIAAC. Each country was responsible for conducting PIAAC in compliance with the Technical Standards and Guidelines (TS&Gs) provided by the Consortium to ensure that the survey design and implementation yields high-quality and internationally comparable data. The standards were generally based on agreed-upon policies or best practices to be followed by all participating countries when conducting the study.

The PIAAC Consortium specified TS&Gs for all aspects of the sample design, including the identification of the target population, the creation of the sampling frame, sample size requirements, and sample selection methods. All participating countries were required to submit sample design plans detailing these aspects to the Consortium for approval several months before data collection. Also, countries were required to complete quality control sample selection forms, which collected sampling information for each stage of selection. These were designed to capture aggregated information necessary for verifying that the sample is representative of the target population and that sampling was conducted in an unbiased and randomized way.

The Consortium did not conduct quality control monitoring activities for the U.S. national supplement, prison and household, and household 2017 study. However, activities similar to those monitored during the main study were conducted throughout the data collection period, and were reported and approved by NCES.

Target Population in the United States

U.S. Round 1(2011–12) or Main Study. The PIAAC main study target population consisted of non-institutionalized adults 16 to 65 years of age who resided in the United States at the time of interview, where age was determined during the screener questionnaire. Adults were included regardless of citizenship, nationality or language. The target population included only persons living in households or group quarters; it excluded all other persons (such as those living in shelters, the incarcerated, military personnel who lived in barracks or bases, or persons who lived in institutionalized group quarters, such as hospitals or nursing homes). The target population included full-time and part-

time members of the military who did not reside in military barracks or military bases, adults in other non-institutional collective dwelling units, such as workers' quarters or halfway homes, and adults living at school in student group quarters, such as dormitories, fraternities or sororities. Adults who were unable to complete the assessment because of a hearing impairment, blindness/visual impairment or physical disability were considered to be out of scope since the assessment did not accommodate such situations.

The household respondent was asked in the screener questionnaire how many people live in the dwelling and have no usual place of residence elsewhere. Those who thought of the household as their primary place of residence, or who spent most of the year in the household even though they may have had another residence, were listed as eligible household members. The list included persons who usually stayed in the household but were temporarily away on business, vacation, in a hospital or living at school.

U.S. Round 2 (2014) or National Supplement to the Main Study (2014). The target population for the national supplement's household-based sample consisted of noninstitutionalized adults, 16 to 74 years old, who resided in the United States at the time of interview, excluding adults 35 to 65 years of age who were either employed or not in the labor force as determined by the screener interview.

Prison Study (2014). The target population of the PIAAC prison study was inmates age 16 to 74 from federal, state, and private prisons that housed federal or state inmates in the United States. Based on the recommendation of the PIAAC Prison Expert Group, the following types of facilities and institutions were excluded:

- private facilities not primarily for state or federal inmates;
- military facilities;
- Immigration and Customs Enforcement (ICE) facilities;
- Bureau of Indian Affairs facilities;
- facilities operated by or for local government, including those housing state prisoners;
- facilities operated by the United States Marshals Service;
- hospital wings and wards reserved for state prisoners;
- facilities that hold only juveniles; and
- community corrections facilities (such as halfway-houses, boot camps, weekend programs, and other entities in which individuals are locked up overnight).

U.S. Round 3 (2017) U.S. PIAAC Study. The target population of the 2017 study consisted of non-

institutionalized adults 16 to 74 years of age who resided in the United States at the time of interview, where age was determined during the screener questionnaire. The other details of the target population criteria were similar to those of Round 1 of the household data collection. Throughout the sample design process and implementation, where applicable, the OECD Technical Standards and Guidelines (TSGs) were implemented.

Sample Design in the United States

U.S. Round 1(2011–12) or Main Study. To arrive at the required minimum of 5,000 completed cases among non-institutionalized persons 16-65 years of age, a four-stage, stratified area probability sample was conducted. It involved the selection of:

- 80 primary sampling units (PSUs) consisting of counties or groups of contiguous counties;
- 901 secondary sampling units, or segments, consisting of census blocks or block groups;
- 9,468 dwelling units (DUs); and
- eligible individuals within DUs, resulting in 5,010 respondents to the survey.

A nationally representative probability sample of 9,468 U.S. households was selected. Of the 9,468 sampled households, 1,285 were either vacant or not a dwelling unit, resulting in a sample of 8,183 households; of these, there were 1,267 households without an adult age 16 to 65. A total of 5,686 of the 6,916 households with eligible adults completed the screener (up to two adults per household could be selected to complete the questionnaire); survey respondents were then selected from eligible adults completing the screener.

The design was similar to the one implemented for the 2003 ALL survey and ensured the production of reliable statistics of comparable quality to the 2003 ALL. Random sampling methods were used, with calculable probabilities of selection at each stage of selection.

For the selection of individuals within DUs, a screener interview was used to identify the eligible persons within selected dwelling units. A sampling algorithm was implemented within the Computer Assisted Personal Interviewing (CAPI) system to select one or two sample persons among those identified to be eligible. Once selected, the background questionnaire interview was completed.

Round 2 (2014) or National Supplement to the Main Study. The national supplement U.S. sample was designed to achieve three core objectives: (a) oversample young adults (age 16-34), (b) oversample unemployed adults (age 16-65), and (c) expand the sample to include older adults (age 66-74). The sample selection method, therefore, differed from the main study sample design. Given the

sample size goal for unemployed and the low prevalence of unemployed adults in the population, a dual-frame approach was implemented, which is a more efficient method of sampling rare populations. The dual-frame approach consisted of an area sample and a list sample.

Under this approach, an area sample of DUs was selected from the same PSUs and segments selected for the main study. The DU frame consisted of the PIAAC main study listings after removing the DUs previously released. One or more persons from the national supplement household sample target population was sampled within a household.

To obtain the oversample of unemployed adults, the frame was supplemented with a list of DUs from high unemployment census tracts. Within each of the PSUs, five high unemployment tracts were identified, and one was randomly selected for the national supplement list sample. The USPS address list was purchased for each of the sampled tracts, and a sample of DUs was taken from these lists. Within the sampled DUs, only those who were unemployed were eligible for selection.

Specifically, to arrive at a minimum of 3,600 completed cases for the national supplement, the four-stage, stratified area frame probability sample involved the following steps:

- 80 PSUs previously selected for the main study consisting of counties or groups of contiguous counties;
- 896 secondary sampling units (SSUs or segments) previously selected for the main study consisting of census blocks or block groups;
- 9,579 DUs; and
- 3,617 individuals within DUs resulting in 2,790 respondents to the survey.

The list sample involved the following steps:

- 80 PSUs previously selected for the main study consisting of counties or groups of contiguous counties;
- 80 SSUs consisting of census tracts;
- 6,956 DUs; and
- 951 individuals within DUs resulting in 870 respondents to the survey.

The national supplement household sample design resulted in a sample that is not stand-alone and is only nationally representative when combined with the main study.

Prison Study (2014). The prison study had a target of a minimum of 1,200 completed cases, including at least 240 females and at least 960 males. In order to achieve this goal, a two-stage, stratified sample was selected with 100 sampled prisons selected in the first stage, among which 80

were all-male or coed prisons and 20 were all-female prisons. All-female prisons were oversampled in order to permit analyses with data from incarcerated women. Due to higher than expected eligibility and response rates, 1,546 eligible inmates were selected within participating prisons, resulting in 1,319 respondents to the survey.

Round Three (2017) U.S. PIAAC Study. The 2017 U.S. sample was designed to achieve two core objectives. First, to provide a nationally representative sample of the U.S. household adult population 16-74 years old. Second, to arrive at sufficient coverage of different types of counties so that, when combined with previous samples (2012 main study and 2014 national supplement), it could improve the indirect small area county- and state-level estimates.

The sample design comprised of a stratified four-stage cluster sample resulted in 3,660 completed cases. At each stage, all sampling units had a non-zero and calculable probability of selection.

To support the second core objective of producing indirect county-level estimates, the overlap with the PIAAC 2012/2014 PSUs was minimized in order to maximize the coverage of the combined sample across demographic variables related to proficiency, such as education attainment, poverty level, minority status, and foreign-born status. That is, by adding sample cases from counties with different demographic characteristics, (related to adult proficiency) as compared to those in the PIAAC 2012/2014, allowed the combined sample to be optimized for county-level estimation, given the available sample size. In PIAAC 2017, adaptive survey design procedures were implemented with the objectives of increasing sample yield through a refreshment sample (at same cost), reducing cost (as measured by contact attempts per completion) through case prioritization, and reducing bias due to nonresponse.

Data Collection and Processing

PIAAC is a voluntary literacy assessment of adults age 16 to 65 internationally and 16 to 74 in the United States.

Reference dates. The PIAAC main study was a new data collection effort and was conducted from August 2011 through April 2012. The national supplement household data collection, round 2 (2014), began in August 2013 and finished in April 2014. The 2017 data collection, round 3, began in March 2017 and finished in November 2017. The prison study (2014) data collection began in February 2014 and finished in June 2014.

Data collection. PIAAC required in-person interviews to complete the background questionnaire, before self-administration of the direct assessments (i.e., literacy, numeracy, reading components and/or problem solving in technology-rich environments). The direct assessments were available in two modes: paper-and-pencil and

computer-administered. For the 2012 data collection, approximately 16 percent of the household respondents in the U. S. sample were directed to the paper-and-pencil path. In the 2014 data collection, about 22 percent of the household respondents were directed to the paper-and-pencil path. In the 2017 data collection, about 16 percent of the respondents were directed to the paper-and-pencil path.

The same procedures and instruments used during the main study in 2012 were employed during the subsequent two household data collections in 2014 and 2017. In 2014, the background questionnaire instrument was practically identical, with only changes in terms of periods and years referred to in the questions. In 2017 there were similar updates as in 2014, as well as addition of several new questions including items on non-degree credentials, in influencing skills and labor market outcomes, military service, and total household income.

The background questionnaire for the prison study of 2014 was specifically tailored to collect information related to the needs and experiences of incarcerated adults.

The same direct assessment was used throughout cycle 1 for both household and prison populations.

Incentives. There were no screener incentives provided in the main study (2011–12). However, a \$5 incentive was offered to each responding household in the national supplement household sample in order to screen for the subgroups of interest in 2014. Upon review of the 2014 national supplement screening results and the logistics required to track the \$5 incentive given to the thousands of 2014 national supplement households, combined with the expectation that most PIAAC 2017 households would have at least one selected participant, the \$5 incentive was eliminated for the 2017 study. No screener incentives have been offered to the prison sample (2014).

In the main study, national supplement, and 2017 study household samples, following the completion of the assessment, a monetary incentive of \$50 was paid to each respondent. The incentive was also paid to those adults who attempted to complete the assessment but were not able to complete it for reasons of language barriers or physical or mental disabilities. Respondents who refused to continue with the assessment were not compensated.

Data entry and verification. The Consortium required that data preparation and processing be performed in a uniform way within and across countries and with an acceptable quality level. Key data preparation tasks ensured this uniformity and were composed of manual data entry of scoring sheets, generation and review of edits on computer generated data files, management of coding, scoring of related files, validation of the structural consistency of the database, and delivery of the national database to the

Consortium. Consortium-provided Data Management Expert (DME) software was used to perform many of these data preparation and processing activities. The Consortium provided each country with the DME software, which was used to assemble, manage, verify, and edit each country's national database. The national DME database consisted of two parts: (1) data collected by the virtual machine's processing of the background questionnaire and the computer-based assessment items or tests administered on the interviewer laptops, and (2) scoring data entered manually and generated as the result of scoring the paper-based assessment booklets.

Estimation Methods

This section provides information for rounds 1 (main study), 2 (national supplement), and 3 (2017 study sample).

Simple formulas that assume simple random sampling for variance estimation were not appropriate with PIAAC data due to complex sample design. The properties of a sample selected through a complex sample design then, could be very different from those of a simple random sample, where every individual in the target population has an equal chance of selection, and in which the observations from different sampled individuals can be considered statistically independent of one another. One way of addressing these departures (e.g. dependent observations, probability of selection not identical for all respondents) from standard statistical properties is by using sampling weights.

Weighting. All population and subpopulation characteristics based on the PIAAC data used sampling weights in their estimation. The purpose of calculating sample weights for PIAAC was to permit inferences from sampled persons to the populations from which they were drawn and to allow tabulations to reflect estimates of the population parameters. Sample weights were produced to accomplish the following five objectives: (1) to permit unbiased estimates, taking account of the fact that all persons in the population will not have the same probability of selection; (2) to minimize biases arising from differences between cooperating and noncooperating sampled persons; (3) to utilize auxiliary data on known population characteristics in such a way as to reduce sampling errors and bring data up to the dimensions of the population totals; (4) to reduce the variation of the weights and prevent a small number of observations from dominating domain estimates; and (5) to facilitate sampling error estimation under complex sample designs.

Objective 1 was accomplished by computing base weights for the households selected for screening and, subsequently, for persons selected for the background questionnaire and assessment from the eligible participating households in the household sample. For the prison study, it was accomplished by computing base weights for the sampled

prisons and then inmates sampled in the participating prisons for the background questionnaire and assessment.

Objective 2 was accomplished through nonresponse weighting adjustments that accounted for screener nonresponse and background questionnaire nonresponse.

Objective 3, the weights for the household sample were calibrated to known totals from the 2012 American Community Survey (ACS). For the prison study, the weights were calibrated to known totals provided by the Bureau of Justice Statistics. The weights were calibrated using a raking procedure (i.e., iterative poststratification) so that numerous totals calculated with the resulting full-sample weights would agree with the ACS totals.

Objective 4 was addressed by trimming the weights. A small number of weights were reduced using an inspection approach (referred to as the $k \times$ median rule) as required by PIAAC weighting guidelines. After the trimming procedure, the weights were again calibrated to ACS totals. No trimming was conducted for the prison sample.

Finally, Objective 5 was accomplished by creating 45 replicate weights using the stratified jackknife method. Full-sample and replicate weights were calculated for each record to facilitate the computation of unbiased estimates and their standard errors. The weighting procedures were repeated for 45 strategically constructed subsets of the sample to create a set of replicate weights for variance estimation using the jackknife method. The replication scheme was designed to produce stable estimates of standard errors.

Weighting was performed separately for the household and prison samples. For the household sample, an additional goal of the weighting process was to improve the precision of estimates for unemployed persons and two groups of young adults (ages 16-24 and 25-34) by combining the main study and national supplement. Composite weights were produced so that national estimates could be generated for the combined sample. The main study sample, national supplement area sample, and national supplement list sample were weighted separately to account for nonresponse, calibrated, composited, and then recalibrated.

In addition, a set of weights for the combined PIAAC 2012/2014/2017 sample was created to allow for the creation of indirect small area county-level estimates and state-level estimates when the PIAAC 2017 sample is combined with the PIAAC 2012/2014 sample. The weights for the combined sample also allow for the production of national estimates for more detailed subgroups of the population than is possible with the separate samples.

Imputation. For the combined household sample, missing values of age category (10 cases) were imputed using the

broad age range collected in the screener. Race/ethnicity for cases missing this item (175 cases) was created by imputing ethnicity (Hispanic/not Hispanic) first, and then race. To obtain values for ethnicity, cells were formed by PSU, segment, and language spoken at the screener. Then a hotdeck procedure was used to assign the value from a random donor within the cell to the missing case. To obtain values for race, cells were formed by PSU and segment and values imputed using the hotdeck procedure. For level of education and country of birth--information that was not collected through the screener, a limited amount of imputation was performed to fill in the data for respondents.

Imputation was performed separately for the main study and national supplement but followed the same general procedure. No employment status information was collected in the screener in the main study, so a different imputation approach was needed. For the two respondents with missing values, cells were formed by PSU and segment and values imputed using the hotdeck procedure. Imputation for the literacy-related nonrespondents was done by taking a random draw from the employment distributions from the 2012 ACS. For language problems, this was based on the distribution of employment for those that speak English not well or not at all. For learning/mental disabilities, imputation used the distribution of employment for persons with cognitive difficulty.

Small area estimation (SAE). Since 2013, PIAAC has published a large volume of official statistics about the proficiency of adults in the United States. The published statistics are mainly for the nation and for major subgroups. However, policymakers, business leaders, and educators/researchers often need information about smaller geographic areas. To address this need, PIAAC has used advanced statistical modeling approaches to produce literacy and numeracy estimates for all states and counties.

The objectives of the small area estimation process are to 1) reduce the mean square error associated with the state and county estimates, and to 2) provide accurate estimates of the mean square error to allow user to understand the level uncertainty associated with the small area estimates. The mean square error is a measure of the uncertainty surrounding the small area estimates. The mean square error has two components – bias and variance. To achieve the first objective, a large number of covariates were gathered from various sources, including the American Community Survey. A small number of covariates are selected for the model that provide the most predictive power. In addition, three levels of random effects (country, state, census division) are used to account for variation between areas. For the second objective, various sources of error (e.g., sampling error, measurement error) are incorporated in a Hierarchical Bayes model to generate thousands of realizations of the model outcomes. The variation across the

realizations provides resulting mean square error. A thorough set of diagnostics is done on the model results prior to making predictions for areas that do not have PIAAC sample.

A visualization-based website will launch in 2020 that will allow users access to the small area estimates through heat maps and summary card displays. This user-friendly website will provide precision estimates and facilitate statistical comparisons among counties and states.

5. DATA QUALITY AND COMPARABILITY

This section provides information for rounds 1 (main study), 2 (national supplement), and 3 (2017 study sample).

Two broad categories of error occur in estimates generated from surveys: sampling and nonsampling errors.

Sampling Error

Sampling error is the uncertainty that exists because population estimates are based on a sample rather than a census. Clustering effects can cause additional uncertainty in estimates that cannot be handled by conventional formulas for variance estimation.

Another procedure that affects variances which is not captured by standard estimation approaches is estimation through Item Response Theory (IRT) models; because different respondents take different sets of items that could differ by level of difficulty, it is inappropriate to base the competency estimates simply on the number of correct answers obtained. The IRT model uses the item responses for each individual, regarding the latent literacy score as random, and generates several predicted (plausible) values, which also have variation. Given these complexities, the Consortium specified standards in the TS&Gs regarding the creation of special weights to facilitate computation of sampling error estimates for PIAAC. For these reasons, PIAAC provides estimates of standard errors using a stratified jackknife replication approach.

Nonsampling Error

Nonsampling error contains all sources of error besides sampling error. There are three components of nonsampling error: (1) frame error, (2) measurement error, and (3) nonresponse error, with nonresponse bias being a key indicator of the latter.

Unit nonresponse for the Main Study and National Supplement to the Main Study. The PIAAC samples were subject to unit nonresponse from the screener, background questionnaire, assessment (including reading components), and item nonresponse to background questionnaire items. Both the screener and the background questionnaire had a

unit response rate below 85 percent and thus required an analysis of the potential for nonresponse bias according to the National Center for Education Statistics statistical standards.

For the U. S., the final screener response rate was 84.7 percent weighted (main study and national supplement combined). Based on the screener data, 10,668 respondents age 16 to 65 were selected to complete the background questionnaire and the assessment; 8,670 actually completed the background questionnaire. The final response rate for the background questionnaire—which included respondents who completed it and respondents who were unable to complete it because of a language problem or mental disability—was 80.9 percent weighted.

Of the 8,670 adults age 16 to 65 who completed the background questionnaire, 8,367 completed the adult literacy assessment. The final response rate for the overall assessment—which included respondents who answered at least one question on each scale and the respondents who were unable to do so because of a language problem, mental disability, or technical problem—was 98.8 percent weighted.

The final U. S. household reporting sample—including the imputed cases—consisted of 8,670 respondents.

Unit nonresponse for the 2017 Study Sample. For the U.S., the final screener response rate was 74.9 percent weighted. Based on the screener data, 4,769 respondents age 16 to 65 were selected to complete the background questionnaire and the assessment; 3,660 actually completed the background questionnaire. The final response rate for the background questionnaire—which included respondents who completed it and respondents who were unable to complete it because of a language problem or mental disability—was 76.3 percent weighted.

Of the 3,660 adults age 16 to 65 who completed the background questionnaire, 3,406 completed the adult literacy assessment. The final response rate for the overall assessment—which included respondents who answered at least one question on each scale and the respondents who were unable to do so because of a language problem, mental disability, or technical problem—was 98.0 percent weighted.

The final U. S. household reporting sample—including the imputed cases—consisted of 3,660 respondents.

Unit nonresponse for the Prison Study. Of the 1,546 sampled inmates, 1,315 completed the background questionnaire. The final response rate for the background questionnaire, which included respondents who completed it and respondents who were unable to complete it because of a literacy-related barrier, was 85.8 percent weighted.

Of the 1,315 inmates who completed the background questionnaire, 1,274 completed the assessment. The final response rate for the overall assessment was 97.7 percent weighted.

The overall weighted response rate for the prison study was 82.2 percent (treating substitute prisons as nonresponse). The final prison reporting sample consisted of 1,319 respondents, including 1,315 respondents who completed the background questionnaire plus the 4 respondents who were unable to complete the background questionnaire for literacy-related reasons.

Nonresponse error. Nonresponse bias is a key indicator of nonresponse error, and can be substantial when two conditions hold: (1) when response rate is relatively low, and (2) when the difference between the characteristics of respondents and nonrespondents is relatively large. The nonresponse bias analyses of the PIAAC household samples in the United States revealed differences in the characteristics of respondents who participated in the background questionnaire compared with those who refused. In a bivariate unit-level analysis at the background questionnaire stage, estimated percentages for respondents were compared with those for the total eligible sample to identify any potential bias owing to nonresponse. Multivariate analyses were conducted to further explore the potential for nonresponse bias by identifying the domains with the most differential response rates.

For the main study, these analyses revealed that the subgroup with the lowest response rates for the background questionnaire had a combination of the following characteristics:

- Hispanics age 26 and older;
- with no children under age 16 in the household;
- not living in the Northeastern United States;
- living in segments with unemployment exceeding 4.8 percent; and
- living in areas (census tracts) with less than 5.1 percent of the population being linguistically isolated.

The presence of children under age 16 in the household was the dominant variable in distinguishing response rate groups. In general, younger persons were found to be more likely to participate, as were those with children age 16 and younger, and women.

For the national supplement household area sample, analyses identified that the lowest response rate was for a combination of the following characteristics:

- with no children under age 16 in the household;

- not unemployed (age 16 to 34) or older (age 66 to 74);
- living in census tracts in which the employment rate exceeds 64.53 percent;
- living in the Northeastern United States;
- living in census tracts in which more than 2.42 percent of the population is foreign born;
- persons age 25 to 34 or older than 55; and
- living in census tracts in which the unemployment rate is 4.48 percent or less.

The presence of children under age 16 in the household was the dominant variable in distinguishing response rate groups.

For the national supplement household list sample, analyses identified that the lowest response rate was for a combination of the following characteristics:

- living in a Metropolitan Statistical Area;
- female;
- living in the Western and Northeastern United States;
- living in census tracts in which less than 28.57 percent of the population has a high school education; and
- with no children under age 16 in the household.

The indicator of whether a sampled person resided in a Metropolitan Statistical Area was the dominant variable in distinguishing response rate groups.

No nonresponse bias analysis was needed for the prison study because the weighted response rates for all data collection stages and all background questionnaire items were above the 85 percent response rate requirement

The variables found to be significant in the bivariate analysis—those used to define areas with low response rates—were used in weighting adjustments. The analysis showed that weighting adjustments were highly effective in reducing the bias. The overall conclusion from the PIAAC study on nonresponse bias is that some minimal potential for nonresponse bias exists in the PIAAC estimates; however, the analysis shows that the bias is negligible.

Item nonresponse. Since all items had greater than an 85 percent response rate, the potential for bias due to item nonresponse was considered negligible.

Data Comparability

Overall trend comparisons over time can be conducted for the total adult population in the areas of literacy and numeracy. In literacy, comparisons are made between PIAAC (2012/2014, 2017) and both ALL (2003–2008) and

IALS (1994–1998). In numeracy, trend comparisons are made between PIAAC (2012/2014, 2017) and ALL (2003–2008). In both the literacy and numeracy domains,

approximately 60 percent of the items are common between PIAAC and previous international surveys to ensure the comparability of these domains.

Table PIAAC-1. U.S. weighted response rates: PIAAC 2012, 2014, and 2017

Component	Main study sample (round 1)	National supplement (round 2)			2017 study sample (round 3)
		Household area sample	Household list sample	Prison sample	
Screener	87	81	85	†	75
Background questionnaire	82	78	93	86	76
Assessment (without reading component)	99	99	99	98	98
Overall	70	63	78	82	56

† Not applicable.

SOURCE: PIAAC publication NCES 2016-036REV and NCES 2020-224; available at <https://nces.ed.gov/pubsearch/getpubcats.asp?sid=113>.

6. CONTACT INFORMATION

For content information on PIAAC, contact:

Holly Xie
 Phone: (202) 245-8481
 E-mail: Holly.Xie@ed.gov

Mailing Address

National Center for Education Statistics
 Institute of Education Sciences
 Potomac Center Plaza
 550 12th Street, SW
 Washington, DC 20202

7. METHODOLOGY AND EVALUATION REPORTS

General

Goodman, M., Finnegan, R., Mohadjer, L., Krenzke, T., and Hogan, J. (2013). *Literacy, Numeracy, and Problem Solving in Technology-Rich Environments Among U.S. Adults: Results from the Program for the International Assessment of Adult Competencies 2012: First Look* (NCES 2014-008). National Center for Education Statistics, Institute of Education Sciences, U.S. Department of Education. Washington, DC. <https://nces.ed.gov/pubs2014/2014008.pdf>.

Krenzke, T., VanDeKerckhove, W., Thornton, N., Diaz-Hoffmann, L., Hogan, J., Mohadjer, L., Li, L., Li, J., Yamamoto, K., Khorramdel, L., and Ali, U. (2019). *U.S. Program for the International Assessment of Adult Competencies (PIAAC) 2012/2014/2017: Main Study,*

National Supplement, and PIAAC 2017 Technical Report (NCES 2020-224). U.S. Department of Education. Washington, DC: National Center for Education Statistics.

<https://nces.ed.gov/pubs2020/2020224.pdf>

Provasnik, S. (2018). *Analyzing U.S. Young Adults' Skills by Student and Employment Status: Methodology for a New PIAAC Variable with Initial Results* (NCES 2018-122). U.S. Department of Education. Washington, DC: National Center for Education Statistics.

<https://nces.ed.gov/pubs2018/2018122.pdf>

Krenzke, T., Mohadjer, L., Li, J., Erciulescu, A., Fay, R., Ren, W., Van de Kerckhove, W., Li, L., and Rao, J.N.K. (2020). *Program for the International Assessment of Adult Competencies (PIAAC): State and County Estimation Methodology Report* (NCES 2020-225). U.S. Department of Education, National Center for Education Statistics. Washington, DC: U.S. Government Printing Office.

<https://nces.ed.gov/pubs2020/2020225.pdf>

Rampey, B.D., Finnegan, R., Goodman, M., Mohadjer, L., Krenzke, T., Hogan, J., and Provasnik, S. (2016). *Skills of U.S. Unemployed, Young, and Older Adults in Sharper Focus: Results From the Program for the International Assessment of Adult Competencies (PIAAC) 2012/2014: First Look* (NCES 2016-039rev). National Center for Education Statistics, Institute of Education Sciences, U.S. Department of Education. Washington, DC.

<https://nces.ed.gov/pubs2016/2016039rev.pdf>

Rampey, B.D., Keiper, S., Mohadjer, L., Krenzke, T., Li, J., Thornton, N., and Hogan, J. (2016). *Highlights from the U.S. PIAAC Survey of Incarcerated Adults: Their Skills, Work Experience, Education, and Training: Program for the International Assessment of Adult Competencies: 2014* (NCES 2016-040). National Center for Education Statistics, Institute of Education Sciences, U.S. Department of Education. Washington, DC.
<https://nces.ed.gov/pubs2016/2016040.pdf>.

Survey Design

Hogan, J., Montalvan, P., Diaz-Hoffmann, L., Dohrmann, S., Krenzke, T., Lemay, M., Mohadjer, L., and Thornton, N. (2014). *Program for the International Assessment of Adult Competencies 2012: U.S. Main Study Technical Report* (NCES 2014-047). National Center for Education Statistics, Institute of Education Sciences, U.S. Department of Education. Washington, DC.
<https://nces.ed.gov/pubs2014/2014047.pdf>.

Hogan, J., Thornton, N., Diaz-Hoffmann, L., Mohadjer, L., Krenzke, T., Li, J., VanDeKerckhove, W., Yamamoto, K., and Khorramdel, L. (2016). *U.S. Program for the International Assessment of Adult Competencies (PIAAC) 2012/2014: Main Study and National Supplement Technical Report* (NCES 2016-036REV). National Center for Education Statistics, Institute of Education Sciences, U.S. Department of Education. Washington, DC.
https://nces.ed.gov/pubs2016/2016036_rev.pdf.

OECD. (2012). *Literacy, Numeracy and Problem Solving in Technology-Rich Environments: Framework for the OECD Survey of Adult Skills*. OECD Publishing.
<http://www.oecd.org/education/literacynumeracyandproblemsolvingintechology-richenvironments-frameworkfortheoecdsofadultskills.htm>.

OECD. (November 2011). *PIAAC Conceptual Framework of the Background Questionnaire Main Survey*.
[http://www.oecd.org/skills/piaac/PIAAC\(2011_11\)MS_BQ_ConceptualFramework_1%20Dec%202011.pdf](http://www.oecd.org/skills/piaac/PIAAC(2011_11)MS_BQ_ConceptualFramework_1%20Dec%202011.pdf).

OECD. (November 2009). *PIAAC Literacy: A Conceptual Framework*.
http://www.oecd-ilibrary.org/education/piaac-literacy-a-conceptual-framework_220348414075.

OECD. (November 2009). *PIAAC Numeracy: A Conceptual Framework*.
http://www.oecd-ilibrary.org/education/piaac-numeracy-a-conceptual-framework_220337421165.

OECD. (November 2009). *PIAAC Problem Solving in Technology-Rich Environments: A Conceptual Framework*.
http://www.oecd-ilibrary.org/education/piaac-problem-solving-in-technology-rich-environments-a-conceptual-framework_220262483674.

Data Quality and Comparability

Lessler, J., and Kalsbeek, W. (1992). *Nonsampling Error in Surveys*. New York: John Wiley & Sons.
https://www.worldcat.org/title/nonsampling-error-in-surveys/oclc/1058850172&referer=brief_results.