

International Study of Adult Skills and Learning
(ISASL) [Program for the International Assessment
of Adult Competencies (PIAAC) Cycle II]
2022 Field Test

Supporting Statement Part B

OMB # 1850-0870 v.7

Submitted by
National Center for Education Statistics
U.S. Department of Education

revised October 2019
September 2019

As in Cycle I, a user-friendly name for PIAAC Cycle II was created – the International Study of Adult Skills and Learning (ISASL) – to represent the program to the public, and will be used on all public-facing materials and reports. As this international program is well-known within the federal and education research communities, we continue to use "PIAAC" in all internal and OMB clearance materials and communications, and use the "PIAAC" name throughout this submission. However, as seen in Appendix E, all recruitment and communication materials refer to the study as ISASL.

TABLE OF CONTENTS

<u>Section</u>		<u>Page</u>
B	COLLECTIONS OF INFORMATION EMPLOYING STATISTICAL METHODS.....	1
	B.1 Respondent Universe.....	1
	B.2 Procedures for the Collection of Information.....	1
	B.3 Maximizing Response Rates.....	5
	B.4 Tests of Methods and Procedures	9
	B.5 Individuals Consulted on Study Design.....	9
	References.....	9

B. COLLECTIONS OF INFORMATION EMPLOYING STATISTICAL METHODS

B.1 Respondent Universe

The PIAAC Cycle II Field Test target population consists of non-institutionalized adults who at the time of the survey reside in the U.S. (whose usual place of residency is in the country) and who at the time of interview are between the ages of 16 and 74 years, inclusive. Adults are to be included regardless of citizenship, nationality, or language. The target population excludes persons not living in households or non-institutional group quarters, such as military personnel who live in barracks or bases, or persons who live in institutionalized group quarters, such as jails, prisons, hospitals, or nursing homes. The target population includes full-time and part-time members of the military who do not reside in military barracks or military bases, adults in other non-institutional collective dwelling units, such as workers' quarters or halfway homes, adults living at school in student group quarters, such as a dormitory, fraternity, or sorority. Persons who are temporarily in the U.S., such as those individuals on a work visa, will be included in PIAAC if they have lived in the sampled dwelling unit for 6 months or more of the last 12 months or if they have just moved into the dwelling and expect it to be their "usual residence" moving forward (where they will reside for 6 months or more per year). Adults who are unable to complete the assessment because of a hearing impairment, blindness/visual impairment, or physical disability are in-scope; however because the assessment does not offer accommodations for physical disabilities, they are excluded from response rate computations. Based on the Census Population Projections for 2021 (U.S. Census Bureau, 2018), there are expected to be about 245 million people in the target population at the time of the PIAAC Cycle II data collection.

The overall sample size goal for the Field Test is to reach 1,550 16-74 year olds, of which 1,500 are 16-65 year olds and 50 are 66-74 year olds. The reason for the two groups is that the international PIAAC comparison group is 16-65 year olds, and 66-74 is of special interest to the U.S. The target sample size for ages 16-65 are driven by the needs of the psychometric analysis, including validating items for the assessment. We plan to achieve an extra 50 completed cases among 66-74 year olds in order to test the screening mechanism that will be used in the PIAAC Cycle II Main Study.

Three prior U.S. PIAAC data collections have occurred to date in Cycle I. The United States completed its PIAAC Cycle I 2012 Main Study along with the 23 countries who participated internationally. It included a sample of 5,010 adults in 80 Primary Sampling Units (PSUs). The survey components included a screener, an in-person background questionnaire, and a computer-based or paper assessment. A second U.S. data collection, the PIAAC Cycle I 2014 National Supplement of 3,660 respondents, was conducted with the same survey components. The PIAAC Cycle I 2012 and 2014 data collection efforts together formed an enhanced 2012/2014 combined nationally representative sample. The PIAAC Cycle I 2017 National Supplement was the third data collection for PIAAC in the U.S. National estimates of proficiency in literacy, numeracy and problem solving from the 2017 national sample provided additional set of data points for evaluating trends in proficiency as compared to estimates from the 2012/2014 combined sample, and earlier adult literacy surveys. In addition, the data collected from all survey years (2012/2014/2017) were used to form sub-national model-based estimates through small area estimation (SAE) methods.

B.2 Procedures for the Collection of Information

Contacting, Screening and Interviewing on Cycle II U.S. PIAAC

When contacting sampled households, administering the screener instrument to identify eligible respondents, and conducting the interview with sampled persons, interviewers are expected to follow established procedures and protocols as described below.

- **Contacting and Screening Sampled Dwelling Units** - All interviewer contacts with sampled households should be in-person (unless otherwise requested by the household and authorized by the supervisor). If someone opens the door at first contact, the interviewer will display their study ID and ask to speak with a household member 18 years old or older, if available. If so, the interviewer will provide the household member a copy of the Advance Letter to introduce the study, and administer the Screener. If no one is at home at the time of first contact, the interviewer will record the contact attempt on the Record of Contacts for the case on their tablet, and return to the sampled address on a different day of the week or time of day to attempt to administer the Screener. The interviewer should contact the sampled household at least seven times before the case is reviewed by the supervisor to determine the best course of action.

If, at first contact, the interviewer encounters someone at the door who does not speak English, the interviewer should ask for an English-speaking person to come to the door. If that person is a household member at least 18 years old, the interviewer should conduct the screener with that person. If not, the person can serve as an interpreter during the administration of the Screener with the household member that does not speak English. If neither option is viable, the interviewer should thank the person at the door, leave, and describe the language problem on the Record of Contacts for the case.

- **Administering the Interview** - Once the Screener has sampled eligible persons in a household, the interviewer should provide the sampled person with a copy of the Advance Letter to introduce the study and administer the Case Initialization, the Background Questionnaire and the assessment at that time to the sampled person(s) if available. The background questionnaire and the assessment should be administered in one session. If the sampled person is under 18, the interviewer will explain to a parent or guardian what survey participation entails, provide the study brochure, and answer any questions the parent/guardian may have. If the parent/guardian refuses to allow the sampled youth to participate, the interviewer will attempt to convince the parent or guardian to allow the youth to participate. If unsuccessful, the case is closed as a final refusal. If the sampled person is not available, the interviewer should make an appointment for a specific date and time, or ascertain the best time to return when the sampled person is likely to be at home, then return at the agreed upon time to conduct the interview. If unable to complete the interview at that time, the interviewer should continue to contact the sampled person at least seven times before the case is reviewed by the supervisor to determine the best course of action.

If the sampled person does not speak English, the interviewer should ask the person who interpreted the Screener to assist with the administration of the Background Questionnaire. If that approach is not successful, the interviewer should thank the sampled person, leave, and describe the language problem on the Record of Contacts for the case.

Statistical Methodology

This section describes the sample design for the field test. With the aforementioned goal of completing assessments for 1,500 adults ages 16 through 65 years and 50 adults ages 66 through 74 years, the field test sample will be a household nonprobability sample, which will satisfy the objectives of the PIAAC Field Test at a reduced cost of data collection, compared to a representative sample. The sample will be diverse, including adults of various age, gender, education, race/ethnicity, and language groups found in the U.S. population.

We plan to use the electronic address lists from the PIAAC Cycle I 2017 data collection, which will significantly reduce the time and resources needed for sample selection. Coverage inadequacies are not a concern because the Field Test is not a nationally representative sample. Therefore, the PIAAC Cycle I 2017 address lists will be used without coverage updating.

Sample Selection

A four-stage sample design will be employed, where the first stage sampling frame is comprised of primary sampling units (PSUs), which are counties or groups of contiguous counties. The second stage involves secondary sampling units (SSUs) (census blocks or combinations of blocks), the third stage is dwelling units (DUs), and the fourth stage involves the random selection of one or more eligible adults per household. The sample selection approach is described for each sampling stage below in turn.

For the first stage of selection, a stratified purposive subsample of 20 PSUs will be selected from the PIAAC Cycle I 2017 sample of 80 PSUs. The PSUs for PIAAC 2017 were selected as a stratified probability-proportionate-to-measure of size (MOS) sample, where the MOS was equal to the target population. Using the 80 selected PSUs from the 2017 sample as a sampling frame allows some savings of time in not having to build a frame of PSUs and SSUs, and, as mentioned above, not having to request addresses for sample selection. The PIAAC standards allow for a nonprobability design for the Field Test, and therefore, we have opted for a purposive sample to produce a diverse set of PSUs. The 20 PSUs for the PIAAC Cycle II Field Test will be chosen to arrive at a sample that is fairly evenly distributed across categories of key variables, such as age, gender, education, income, and race/ethnicity. The diversity among demographic subgroups is a requirement for psychometric testing. Therefore, we will stratify the 80-PSU PIAAC Cycle I 2017 sample into 20 groups based on the stratification variables from the PIAAC Cycle I 2017 sample: size (whether the PSU was selected with certainty), metropolitan status (whether the PSU was part of a metropolitan area), educational attainment (percentage of the population ages 25 and over with no more than high school education), income (percentage of population who are below 150 percent poverty), ethnicity (percentage of the population age 15 to 64 that are Hispanic or non-Hispanic Black alone), and country of birth (percentage of the population who were foreign born and entered US in 2000 or later). The 20 PSUs result from purposively selecting one PSU from each of the 20 groups.

In the second stage of selection, SSUs selected for the PIAAC Cycle I 2017 data collection will constitute the frame for the SSU sample. A systematic random subsample of SSUs will be selected for the Field Test. We plan to select 8 SSUs from the SSUs in the PIAAC Cycle I 2017 sample within each of the 20 PSUs for a total of 160 SSUs. This will result in an equal interviewer workload design with an equal number of completed cases expected per SSU.

The third stage of sample selection will involve a sample of households from the PIAAC Cycle I 2017 electronic address lists in each sample SSU. To avoid overlapping with the PIAAC Cycle I 2017 sample, any households included in the PIAAC Cycle I 2017 sample will be removed from the frame first. Then about 25 dwelling units will be selected within each of the 160 sample SSUs. The list of dwelling units will be stratified by SSU and sorted within each SSU by geographic order. A systematic random sample will be selected within each SSU. As the Field Test does not require a probability sample, we will not be conducting address coverage enhancement procedures.

The fourth stage of selection involves a two-phase sampling approach. Using a screener questionnaire, we will both collect information about persons in the household and determine eligibility and stratify by age group. Using this information, we will apply different sampling rates to arrive at the target sample sizes in each age group.

The screener begins by listing the age-eligible household members (aged 16 to 74) in the selected dwelling unit during the screener interview. Once persons are enumerated in a household, two strata will be formed, 1) 16-65 year olds, and 2) 66-74 year olds. In stratum 1, one person will be selected if there are three or less in the household; otherwise two persons will be selected. This is the same rule that will be used in the Main Study. In the Main Study, selecting two persons in the larger households will help to reduce the variation due to unequal probabilities of selection, which can increase variance estimates (Kish, 1965). For stratum 2, approximately one out of every five households with at least one 66-74 year old will be flagged to select one 66-74 year old within the household. This subsampling is needed to keep the sample size of 66-74 year olds at 50. Flagged households without an individual in either of the two strata will be classified as ‘ineligible’.

The enumeration and selection of persons will be performed using the CAPI system, which will collect information via the screener instrument, including age and gender of persons in the dwelling unit, and randomly select eligible respondents.

Table 4 below provides the response rates for the U.S. PIAAC Cycle I data collections.

Table 4. PIAAC Cycle I Response Rates

Survey Year	Weighted Response Rate
2012	70.3
2014	63.0
2017	56.0

Table 5 below provides a summary of the sample sizes needed at each sampling stage to account for sample attrition due to (a) ineligibility (i.e., households without at least one 16-74 year old adult or vacant dwelling units), (b) screener nonresponse, (c) within-household selection rates, and (d) nonresponse to the BQ and the assessment. Our assumptions for each type of sample attrition are drawn from national estimates produced by the U.S. Census Bureau as well as our experience in prior PIAAC data collections. The occupancy rate is expected to be about 87 percent according to 2017 American Community Survey (ACS) estimates. The eligibility rate (the percentage of households with at least one person in the target population) is based on the Current Population Survey (CPS) March 2018 Supplemental File. For 66-74 year olds, it is also adjusted to reflect the selection from one household out of every five. The expected percentage of households with two sample persons is based on both the CPS March

2018 file and selection rules established for PIAAC. Overall, a 50 percent response rate is expected, decreased from 56 percent in the PIAAC Cycle I 2017 data collection, due to the shortened data collection period for the Field Test. There were no significant differences in response rates by age group in the PIAAC Cycle I 2017 data collection, and so the same rates are assumed for the 16-65 year olds and the 66-74 year olds in the Cycle II Field Test.

In addition, this sample will be increased to provide a reserve sample of households. The reserve sample will only be used in case there are unusual response problems or unforeseen sample losses observed during the data collection. A reserve sample of about 50 percent the size of the main sample will be selected randomly and set aside to be used in case of a shortfall in the sample.

Table 5. PIAAC Cycle II Field Test sample yield estimates

Survey and sampling stages	Eligibility and response rates	Projected rates	Expected sample size
Number of selected PSUs			20
Number of selected SSUs			160
Expected number of selected households (HHs)	Occupied dwelling unit rate	87.4%	3,928
Expected number of occupied dwelling units	Screener response rate	71.4%	3,433
Expected number of completed screeners			2,451
Expected number of eligible screeners	Eligibility rate 16-65 (66-74)	81.5% (2.9%)	1,998 (71)
Expected number of attempted BQs	Percentage of HHs with two sample persons 16-65 (66-74)	7.2% (0%)	2,142* (71)
Expected number of persons with completed BQs	BQ response rate 16-65 (66-74)	71.4% (71.4%)	1,529 (51)
Expected number of completed or partially completed assessments	Assessment completion rate 16-65 (66-74)	98.1% (98.1%)	1,500 (50)

* This number is reached by considering 100% of eligible screeners to attempt a background questionnaire and adding an additional estimated group of 144 respondents who are second eligible respondents in the same household. We estimate that 7.2% of all households will contain two potential sample members.

NOTE: Figures in parentheses are for the subgroup 66-74 year olds.

Estimation

For the Field Test, the minimum expected sample size mentioned above is required to properly estimate item parameters for all newly developed items and to test the stability of the trend item parameters for each tested language in a participating country. The ability to properly estimate the new and trend item parameters in the field test will allow for the best selection of items to maximize efficiency of the assessment's adaptive design used in the Main Study.

The estimation procedures for the PIAAC data are prescribed by and are the responsibility of the international sponsoring agency; however, the United States has reviewed and agrees with these procedures. The United States will comply with these procedures and policies by delivering masked data (note that a disclosure analysis will be conducted prior to submitting the data to the international contractor so as to comply with current federal law), and documentation of sampling variables. All data delivered to the PIAAC Consortium will be devoid of any data that could lead to the

identification of individuals.

Degree of Accuracy

As noted above, the sample size requirements are driven by the needs of the psychometric analysis. No proficiency estimates will be produced for the Field Test.

Specialized Sampling Procedures

The Field Test objectives do not require a probability-based sampling, and so a nonprobability household sample will be used, as described above.

Any Use of Periodic (Less Frequent Than Annual) Data Collection Cycles to Reduce Burden

There will be no use of periodic data collection cycles.

B.3 Maximizing Response Rates

NCES views gaining respondent cooperation as an integral part of a successful data collection effort and will invest the resources necessary to ensure that the procedures are well developed and implemented. We will build on the advance contact strategy that was successfully employed on PIAAC Cycle I. An Advance Letter and postcard will be mailed to all sampled households in advance of the data collector's initial visit (see Appendix E). The letter will inform potential respondents of NCES' authorizing legislation; the purposes for which the PIAAC data are needed; uses that may be made of the data; and the methods of reporting the data to protect privacy. An informative brochure (text provided in Appendix E) will be given to sampled participants when the interviewer visits the sampled household.

As mentioned on p. 4 of Part A, some effort was taken to brand the study in the United States in a way that would attract, rather than hinder, perspective respondents. The official international name of the study, Program for the International Assessment of Adult Competencies, uses two of the terms that program administrators try to avoid using in interactions with respondents and in outreach materials: assessment and competencies. When respondents believe that their performance is being evaluated or their knowledge is being tested, the risk of refusal is higher, and the more user-friendly name International Study of Adult Skills and Learning (ISASL) may help in that regard.

All project materials will include the study's web site address and a toll-free telephone number for respondents to obtain additional information about the study. The materials will also mention the incentives and will include the study logo for legitimacy purposes. It is very important for the data collector to establish legitimacy at the door, which can be accomplished by the use of a strong introductory statement (see Appendix E) during which the data collector shows their study ID badge and a copy of the advance materials. The study ID badge dimensions are 5.5" by 7", a size that allows the badge to be visible to someone trying to determine from inside the dwelling unit whether to open the door to the interviewer or not. Interviewers also will carry a supply of pens and screen cleaning cloths with the study logo to hand to respondents in an effort to ease the interaction at the door and help pave the ground for a return visit if needed.

A presence in social media platforms (e.g., Facebook, Twitter) will further help legitimize the study and reach respondents who are more inclined to seek information about the study via these platforms. The Field Test will serve to test new approaches and materials in preparation for the Main Study.

The PIAAC study toll-free number is another tool for gaining cooperation from respondents. Since many of the calls made to the number are from respondents wanting to refuse participation, the PIAAC staff member who answers the calls is trained on (1) trying to determine the reason for the refusal and (2) offering the respondent reasons for participating and (3) avoiding a final or hostile refusal so that an interviewer can still contact the household in person. The toll-free number also receives calls from respondents requesting or rescheduling appointments or asking for more information about the study. All the calls are logged into a database, the case identification number is associated with the calls whenever possible, and the information about the call is relayed to the field supervisor assigned to the case for follow up.

Once data collection begins, effective contact patterns are another important component of achieving response rates. Completion rates improve when interviewers attempt contact on different days of the week and at varying times of the day. As required by the international standards and as instructed at interviewer training, interviewers will make six well-timed attempts to contact a household before reviewing the case with the supervisor to identify another pattern of contact. Well-timed contact attempts are those that the interviewer makes at different times of the day and on different days of the week. Supervisors will have access to up-to-date information on their interviewers' patterns of contact attempts and will study these on a daily basis to ensure that interviewers are following the contact protocols before they reach the sixth contact attempt. Supervisors will flag deviations from the contact protocols and will follow up promptly with interviewers.

These other contact strategies may include unscripted telephone calls to households (if requested by household) to answer questions and set appointments, FedEx letters, or postcards via USPS mail. We plan to staff each PSU with two interviewers. Having multiple interviewers in a PSU is advantageous because it allows supervisors to assign the interviewer best suited to work each case based on the demographic characteristics of the neighborhood, and allows for coverage in case of data collector illness or unavailability.

Interviewers are trained to handle refusals to participate in the study in a way that leaves the door open to further contacts. Unless the refusal was hostile, in which case no further contact with the household is had, the cases of households that have refused to complete the Screener and respondents who have refused to complete the interview are discussed by supervisors and interviewers during the weekly call to determine the best approach for attempting to convert the refusal into a completed interview.

In carrying out efforts to achieve high response and participation rates, our data collection efforts follow a phased approach that allows for refusal conversion. Although that will not be a priority during the Field Test, refusal conversion efforts may be called for to reach the 50 percent response rate set in the standards for the Field Test. Refusal conversion efforts include:

- Refusal conversion online training;
- Mailing refusal letters to households or respondents;
- Reassignment of cases to another interviewer in the area; and
- Bringing in a traveling interviewer who is adept at converting refusals to work refusals cases in an area.

Refusal Conversion Online Training

About a month into the field period, all interviewers are required to complete a Refusal Conversion online training module designed to reinforce the techniques learned at the in-person training. The module is followed by a conference call during which the supervisor asks more experienced interviewers and those who have converted several refusals since the field period started to share what strategies have worked from them and role-play with newer interviewers.

Mailing of Refusal Letters

Households and respondents that refused to participate are mailed refusal letters about two weeks after the refusal occurred. The letters are tailored to the particular refusal reason, if ascertained, otherwise a generic refusal letters is sent. The date when the letter is sent is recorded into the Case Management System so that the supervisor and the interviewer can decide when the best time to follow up is based on the interaction had at the door. Stronger refusals need more time, weak refusals are contacted sooner. Reluctant households usually receive only one refusal letter in the course of the field period.

Reassignment of cases to local interviewer

Whenever possible, interviewers working on the same area will be instructed to trade refusal cases. In prior rounds of PIAAC, having a different interviewer contact a reluctant household or respondent often yielded better results than having the same interviewer, a fresh approach at a different sometimes makes all the difference.

Bringing in a Traveling Interviewer

A few interviewers are hired for the study without being assigned to a specific sampled area. These interviewers, known as travelers, are selected for their flexibility in meeting assignment requirements, their proven record as high-producers and refusal converters, and their ability to travel by car or plane to assigned locations on short notice. These travelers are first used to cover areas that are currently unstaffed but, as the field period advances, refusal conversion work becomes a large component of their assignment. All refusal cases are usually transferred to the traveler during their stay in an area to maximize the use of their skills in convincing households and respondents to participate.

Each interviewer will receive a tablet computer loaded with the Interviewer Management System (IMS) to launch all CAPI instruments for administration of the interview. Interviewers will also receive a study smartphone with a PIAAC application for recording information about each visit to a household that did not result in a completed interview. Whenever a refusal or breakoff is encountered, the interviewer will enter information on the PIAAC app or the IMS on the tablet about the reason for refusal and suggestions for conversion strategies. Information about contacts and refusals is available to the supervisors via automated data transmission from the smartphones and tablets to the home office, where it is displayed in the Study Management System and the dashboard (see Section A.3). Information will include: contact date and time, contact result or disposition code, appointment information, and general interviewer comments. These data are very helpful for supervisors to determine when to reassign cases, bring in a traveling interviewer, and send out refusal conversion materials, and in helping to design a more directed and effective campaign to convert the nonresponding households. Contact and decline information will be collected, coded, and included in the biweekly data collection progress report.

The PIAAC dashboard was first used in the Cycle I PIAAC 2017 data collection. The main purpose of the dashboard is to facilitate the use of paradata in the management of the field work. Supervisors, field managers and survey operations staff at Westat all have access to the dashboard. In it, information about contact patterns, cost and production, transmission of data, field work activities (routes and time traveling, completed interview tasks, administrative work) is displayed all at once in various windows for ease of access. Clicking on a particular data point on a window will drill further into the data to get more details, and will link back to the Study Management System to review the history of cases as needed. The dashboard is updated continuously via automatic transmissions from interviewers' smartphones and tablets, therefore the data displayed is up to the minute. Supervisors utilize the dashboard, for example, to ensure interviewers are working the cases as expected, that their travel routes are efficient, and that their production aligns with their hours worked.

NCES believes that frequent, open communication between all levels of field staff is required for a successful data collection effort. Supervisors will use texts and encrypted email for day-to-day communication with their staff. Scheduled weekly conference calls will also be used at all levels of supervision. All supervisory staff will be available for questions or other issues that come up every day via telephone, texts and email.

Screening households posed the greatest challenge in PIAAC Cycle I data collection with respect to nonresponse. The Cycle II Field Test will be an opportunity to gauge respondents' receptiveness to the study and current reasons for refusing to participate. We will dedicate more training time, in the home study package and during initial interviewer training, to the importance of obtaining high screener completion rates and techniques for completing screeners based on our PIAAC Cycle I experience. We will continue to focus on this throughout data collection via reviewing information about contacts and providing feedback during supervisor/interviewer conference calls. We will also pair successful interviewers with those needing to improve contacting skills. We will also target hard-to-reach-households by tailoring mailings of letters and other study outreach materials.

Response rates are not a priority for the PIAAC Cycle II Field Test, but the Field Test will be used to implement an incentive experiment with the goal of maximizing response rates in the Main Study and to test an adaptive survey design approach towards minimizing nonresponse bias as described below.

Incentive Experiment

Given results of an incentive experiment in the PIAAC Cycle I Field Test conducted in 2011 (Krenzke, Mohadjer and Hao, 2012), the PIAAC Cycle I 2012 data collection provided a \$50 incentive to the sampled person for completion of the full survey, and the 2014 data collection included an additional \$5 incentive to the household respondent after the screener. Both were promised incentives that were conditional upon completion of the survey instrument.

The literature on incentives indicates that a pre-paid (unconditional) incentive is more effective than a promised incentive (e.g., Mercer, et al., 2015), although the research is primarily focused on mail, internet, and telephone surveys. A pre-paid incentive increases the salience of the promised incentive message and the belief that the researcher will follow through.

We will include an experiment in the Field Test to test the impact of providing the incentive as a combination of pre-paid and promised incentives at three points during

the survey process (treatment) versus providing the full amount as a promised incentive after completion of the survey (control). Specifically, for a random one-half of the sample (treatment group), we will provide a \$2 prepaid cash incentive in the advance letter envelope, a \$5 cash card upon screener completion, and a \$50 cash card upon completion of the assessment. For the other half sample (control group), respondents will get a \$50 cash card for completion of the background questionnaire and the assessment. In PIAAC Cycle I, the level of nonresponse was similar at the screener and BQ/assessment stages, and so the addition of a pre-paid incentive, along with an incentive for screener completion, could be beneficial.

We will conduct the experiment at the SSU level rather than at the DU level because DU-level designs (1) have an increased chance of introducing error in administering the incentives to the respondents, and (2) introduce the risk of spreading information about different types of incentive in a single neighborhood.

Incentive types will be randomly assigned to each SSU. The assignment will be done systematically by sorting the SSUs by PSU, and percentage of population who are below 150 percent poverty of the SSU within the PSU, and then alternating the assignment of the combination payment and the promised-only payment to the SSUs. After the assignment, quality checks will be conducted to ensure that each incentive group was balanced in terms of demographic characteristics (poverty status, educational attainment, percent Hispanic, and geographic region).

The following highlights other major aspects of the incentive experiment design.

- Interviewers will be given both combination and promised-only SSUs to minimize any interviewer impact on the incentive type effect.
- Incentive assignment to SSUs with common boundaries, or in close proximity, will follow its natural probability-based assignment. Therefore, there will be no special re-allocation of the incentive groups in order to have the same incentive type for SSUs close in proximity.

We will test for differences in the screener response rate, and the assessment response rate (conditional on the screener completion). Given a sample size of 1,716 DUs in each group, with 80% power and significance level of 5%, and an expected screener response rate of 0.714, we can detect a significant increase of 0.043 due to the treatment. Given a sample size of 1,071 selected persons in each group, an expected assessment response rate of 0.700, we can detect a significant increase of 0.055 due to the treatment. These calculations assume no design effect. Some marginal increase to the detectable difference may occur because the design effect is likely small due to controlling on geography (both groups are in the same PSUs).

Adaptive survey design

We will use an adaptive survey design, a method of modifying data collection methods systematically in order to maximize response rates and reduce potential nonresponse bias (Schouten et al., 2013). In our implementation of adaptive survey design, we will use an optimization framework that uses paradata, covariates, and outcome measures from prior data collections to help provide guidance for the data collection effort, by adapting current data collection follow-up strategies using case dispositions and other information collected throughout the current data collection.

An adaptive survey design experiment was conducted during the PIAAC Cycle I 2017 data collection providing some positive results (Krenzke, et al, 2019), including indications of slightly higher response rates, lower costs per complete, and indications of lower nonresponse bias. For the upcoming PIAAC Cycle II Field Test, we will employ

the following adaptive survey design components.

1. **Sample yield projections.** The purpose of this component is to achieve the target number of completes. To do so, we will predict the expected number of completes. These projections are based on a model using paradata and data from the Cycle I. We will estimate the number of completed cases at the end of the data collection period while in the early stages of the Field Test. If the projected number of completes in the Field Test is less than expected based on this model, then a random group of dwelling units will be released from the reserve sample.
2. **Sample refreshment.** The purpose is to boost sample yield at fixed cost without introducing bias. The approach first closes out cases that are fully worked (meet the standard contact protocol) and are identified as very unlikely to be successful, as determined through statistical modeling to estimate the propensity to respond. Then a random group of dwelling units is released with size that is equal to spending the same expected number of attempts as continuing with the closed out cases.
3. **Case prioritization.** The purpose of case prioritization is to reduce nonresponse bias. This approach was developed under an optimization framework based on an “influence index” that measures each open case’s potential influence on reducing nonresponse bias (Riddles and Krenzke, 2016). The index is formed by minimizing nonresponse bias subject to achieving the target sample size, achieving the target response rate, and staying under budget. The index is a function of the estimated response propensity, sample weight, the difference of the open case’s predicted outcome and the average predicted outcome among respondents, and the difference between the average among the sample and the average among respondents. There will be two phases of this method. First, we will assign priorities among cases that have not met the contact protocols (six attempts at three different times of the week). Next we will prioritize cases that have met the contact protocols based on an influence index value: the higher the index influence value, the higher the priority. The Field Management System will display a priority flag for each priority case so that interviewers know which cases to prioritize in their contacts and supervisors can monitor priority cases closely.
4. **Extra postcard.** An extra postcard will be sent to those cases that are assigned high priority on the day the high priority flag is assigned. Interviewers follow up with a visit soon after.

B.4 Tests of Methods and Procedures

The U.S. is participating in a full Field Test for PIAAC Cycle II. The Field Test will provide an opportunity for testing several facets of sampling. The main objectives of the sampling activities are to:

1. Provide a sample of adults that will be used to validate the new and trend items to be included in the psychometric assessment;
2. Test the within-household sample selection process;
3. Train field staff in the implementation of sampling activities in the field;
4. Test the Quality Control (QC) sampling-related procedures; and
5. Test the flow of materials and the sample data from sample selection to the

delivery of the Sample Design International File (SDIF) at the end of the data collection.

B.5 Individuals Consulted on Study Design

The following are responsible for the statistical design of PIAAC:

- Holly Xie, National Center for Education Statistics;
- Stephen Provasnik, National Center for Education Statistics; and
- Kentaro Yamamoto, PIAAC Consortium/Educational Testing Service.

The following are responsible for sampling activities:

- Leyla Mohadjer, PIAAC Consortium/Westat; and
- Tom Krenzke, PIAAC Consortium/Westat.

Analysis of assessment items will be performed by:

- Kentaro Yamamoto, Educational Testing Service.

References

U.S. Census Bureau. (2018). 2017 National Population Projections Tables. Retrieved from: <https://www.census.gov/data/tables/2017/demo/popproj/2017-summary-tables.html>.

Kish, L. (1965). *Survey Sampling*. New York: John Wiley & Sons.

Krenzke, T., Mohadjer, L., and Hao, H. (2012, May). Programme for the International Assessment of Adult Competencies (PIAAC): U.S. incentive experiment. Annual Meeting of the American Association for Public Opinion Research, Orlando, FL.

Krenzke, T., Mohadjer, L., Riddles, M. and Shlomo, N. (2019). Adaptive Survey Design PIAAC Experiment and International Implementation. Presented at the European Survey Research Association Conference, Zagreb, Croatia, July 17, 2019.

Mercer A, Caporaso A, Cantor D, Townsend R. (2015). How Much Gets You How Much? Monetary Incentives and Response Rates in Household Surveys. *Public Opinion Quarterly*. 79(1):105-129.

Riddles, M., and Krenzke, T. (2016). Adapting Responsive Design Strategies in a Recurring Adult Literacy Assessment. Presented at the Joint Statistical Meetings, Chicago, Illinois, August 2016.

Schouten, B., Calinescu, M., Luiten, A. (2013). Optimizing Quality of Response through Adaptive Survey Designs. *Survey Methodology*, 39(1), 29-58.