PISA-D Strand C
Draft Sampling Manual

July 2016

Doc.: PISA D DraftSampleDesignManual .docx

Produced by Westat

1. INTRODUCTION

This document contains a brief description of our general and preliminary plans for developing country-specific designs for the Strand C Field Trial (FT) and the Main Survey (MS) samples. The final plans will be developed after consultations with the Strand C countries, IAG, TAG, and OECD. The document also introduces the Sample Design Forms which will be used to collect information about each country's sample design.

The general sampling plan follows the objectives of Strand C; a pilot project with two main objectives for the sample:

1.   To yield a large enough sample size to test the validity of the questionnaire and the assessment items and other studies planned for this Strand, including linking to PISA-D Strand A (reporting in the same scale).

2.   To explore various approaches and evaluate various options to arrive at a recommendation for identifying and assessing a nationally representative sample of out-of-school 15-year-olds (as defined in section 2) in future cycles of PISA-D.

The method of data collection will be in-person interviews, most likely in households or other similar locations. Generally, costs associated with in-person interviewing are much higher than those in school samples, especially when the target sample is a rare group, as is the case for Strand C. Therefore, a critical component of this pilot project is to arrive at sampling plans that minimize costs to countries, given the objectives of Strand C. To satisfy these objectives, we have developed a general sample design consisting of two major strata and two components within each stratum, as described below. This design will be piloted in FT, and then modified/revised for MS based on lessons learned from FT.

In developing the final sample designs, we will work closely with each country to arrive at a sampling design that is most suitable and optimal for the country and at the same time follows the standards and guidelines of PISA-D Strand C. We will schedule individual sessions with countries to collect more detailed information on various aspects related to sampling and any other information that could be helpful for sampling out-of-school children. For example, it is important to find out if the country has an existing sample that could be used for Strand C. In addition, we will consult with the countries to find out if any large-scale household survey is planned for the same period as PISA-D, and whether the national team can arrange to use their listing operation to locate out-of-school 15-year-olds.

2. DEFINITIONS

The term "out-of-school 15-year-olds" has been and continues to be used throughout the Strand C discussions; however, the actual definition for this term is now considered to be the following.

The current discussions are focused around defining the Strand C population as the complement of the Strand A target population. For Strand A, the primary population consists of 15-year-olds in-school at grades 7 and above. Strand A countries have opted out of the option to include the 15-year-olds in-school at lower than 7th grade. Therefore, the Strand C target population is now considered to include 15-year-olds in-school at grades 6 and lower, and 15-year-olds who are out-of-school. This population is estimated to be roughly about 2/3 of all 15-year-olds, on average, across the PISA-D countries. However, the division between out-of-school 15-year-olds and those in grades 6 and below is vastly different across the countries. For example, in Table 2a of Carr-Hill (2015), the proportion of 15-year-olds who are out-of-school is 37 percent in Guatemala, 48 percent in Senegal, 12 percent in Zambia, and the proportion of 15-year-olds in-school at lower than 7th grade is 25 percent in Guatemala, 13 percent in Senegal and 52 percent in Zambia. In total, the proportions of 15-year-olds in the Strand C target populations are 62 percent, 61 percent, and 64 percent, for Guatemala, Senegal and Zambia, respectively. Similar information on Honduras, Panama and Paraguay were not available in Table 2a of Carr-Hill (2015). Country sample designs will be developed after more detailed and up-to-date information is collected about Strand C target population in each country.

Although references are made to 15-year-olds, the sample will include 14- to 16-year-olds. Therefore, the target population for Strand C will consist of 14- to 16-year-olds who are out-of-school, or in-school but in grades 6 or below.

3. GENERAL FEATURES OF STRAND C SAMPLE DESIGN

The challenge with selecting a representative sample of out-of-school 15-year-olds is that a substantial amount of screening would be required to locate this small subgroup. The following general sample design is aimed at reducing the cost of locating and interviewing the required number of out-of-school 15-year-olds, and at the same time evaluate various options for developing representative samples for future cycles of PISA-D.

The ultimate goal of future cycles of PISA-D Strand C is to produce representative samples of the target population. Therefore, countries will pilot the selection of a probability-based sample as the primary goal of this cycle, with nonprobability sampling used to supplement the sample size for analysis purposes. Under a probability sample, all sample units must have a non-zero probability of selection and all released cases must be worked to the standard number of attempts. In future cycles, probability-based scientific sampling will be essential for two main reasons. First, probability sampling encompasses a set of designs that leads to a variety of unbiased sampling approaches that allow analysts to generalise the results to the target population. Second, measures of precision related to survey estimates (i.e. standard errors, margins of error, confidence intervals) can be computed only under a probability design. Hence, statistical tests for differences between survey estimates are possible only for the portion of the pilot study sample selected under a probability-based design.

For this pilot project, our plan is to divide the country into two parts, called Stratum H and Stratum L. Stratum H includes the areas of the country that have high concentrations of out-of-school youth (e.g., rural areas, urban slums, etc.), while Stratum L consists of areas with low concentrations. In each of these strata, two samples will be designed, a probability sample and a nonprobability sample, as summarized in Figure 1. The nonprobability component is optional in the MS, but countries are required to test it in the FT. The following sections describe the sampling plans and selection approaches (mentioned in Figure 1) for probability and nonprobability samples in each stratum. The details of the design will be developed after consultation with countries and the experts in the area of sampling out-of-school youth.

Figure 1. Overview of sample design

Major Strata	Sample Type	Selection Approach
Within each major stratum (H, L)	Probability	Household Area Sample
	Non-probability	Options 1 Household Area Sample with link-tracing 2 School frame approach

3.1   Probability sample

Applying probability sampling in the high concentrated areas (Stratum H) can greatly reduce the cost of locating out-of-school 15-year-olds. In the low concentrated areas (Stratum L), countries will be asked to select a limited area probability sample mainly used for testing the selection of probability samples across the country. The scope of this work will be limited to reduce costs to the countries.

Generally, the most efficient sample design for a household survey with in-person interviews is a multi-stage stratified cluster design. Such designs give a higher probability of selection to areas with the largest target population size. Clustering reduces cost at expense of increasing variances associated with estimates. Alternative designs might be considered given the specific situations in various countries. The sampling plan will be flexible, and adaptive to each country's best sampling scenario.

In Strand A, the sample designs are essentially standard, where a sampling frame of schools is provided to Westat to select schools, and then students within schools are selected using standard procedures. For household studies, sampling frames are sometimes not readily available or are more difficult to create. Because of these differences, sampling approaches need to be variable across participating countries. Therefore, prior to specifying the details of the design, Westat will gather relevant input from the OECD Secretariat and TAG, as well as national authorities in participating countries. At the same time, Westat will communicate with NPMs and sampling statisticians assigned to Strand C to gather information about the country's approaches to probability sampling and to gain an understanding of the geographic spread to identify areas with large concentrations of out-of-school 15-year-olds. The final sample design will be tailored to national requirements and constraints while satisfying the comparability requirements of Strand C.

In some countries, Census Enumeration Areas (EAs) are already formed and have information available that may be useful for sampling purposes. We will consult with the national statistical institutes about their census data and propose an approach that is tailored to each country's unique situation.

Once areas (clusters) are selected, either a mini-census will be conducted to list persons in the target population (14- to 16-year-olds out-of-school, and those in-school but in grades lower than 7th grade) in the areas or a listing of dwelling units or collective dwelling units will be conducted, as described in Appendix A. Variations are expected since the type of sampling frames within countries could range from having person or address registries to not having any complete lists for use in area sampling. The general approach will be to allow for flexibility in the sample design, conduct a thorough assessment of the quality of sampling frames through a traditional approach, and prepare to be adaptive to each country's situation.

If a listing is conducted, then a screener questionnaire would be used to list persons within the dwelling, to identify persons in the target population, and to select all eligible youth for interview and assessment. Once selected during the screening, the ideal would be to interview and assess the eligible youth in the same visit, if possible. However, determining when the assessment should be administered needs to be discussed with the country representatives, as some countries may prefer to conduct the data collection in two visits. In addition, we will draw from the lessons presented in the 2014 IAG workshop in Paris, the 2014 Technical Workshop on out-of-school 15-year-olds in Montreal, Canada, and from experiences relating to surveys conducted on out-of-school youth through the UIS/UNICEF OOSCI initiative.

3.2   Nonprobability sample

Nonprobability samples are used especially in field trials to capture data at a substantially lower cost than their probability-based counterpart. Under ideal circumstances, the non-probability sample would identify all eligible youth in the cluster that were not obtained through the probability sample, resulting in complete coverage of the target population within the cluster. However, this is unlikely to be realised in practice, and the disadvantage of nonprobability designs is that sampling theory does not hold for making generalisations to the population. Two options will be offered to the countries: household area sampling with link-tracing and a school frame approach.

Link-tracing from households - Examples of link-tracing techniques are respondent-driven sampling (RDS) (Heckathorn, 2007) and snowball sampling (Goodman, 1961). Under some strict assumptions, Heckathorn and Goodman claim that RDS and snowballing can produce national estimates. However, these assumptions are rarely satisfied in practical situations. Therefore, our current plan is to use link-tracing to ensure we reach the required number of completed cases in Strand C. Meanwhile, we will conduct a thorough review of these procedures to arrive at alternative plans that could produce samples that satisfy the conditions of probability sampling. Link-tracing from households may be most applicable in the low concentration strata, but will also work in other areas as well.

A referral approach to link-tracing will be used for PISA-D Strand C. The referral approach is when the respondent refers the interviewer to other individuals (up to two) who may be eligible (age 14 to 16 and either out-of-school or in grade 6 or below). Link-tracing can be triggered by the screener questionnaire in all selected and cooperating households from the probability component, whether or not they include any individuals in the Strand C target population, to gather information to locate and interview other eligible individuals in the area. The goal is to identify eligible youth within the same cluster. However, the respondent is likely unfamiliar with cluster boundaries and so can be asked about the neighbourhood. The relevant screener questions are shown in Exhibit 1.

The link-tracing will be limited to one wave, meaning that the referred youth will not be asked for additional referrals. This has the advantage of reducing the amount of duplication, in other words, reducing the number of times that the same youth is identified through multiple sources. It also decreases the number of cases obtained through link tracing so that the sample yield does not exceed available resources and the non-probability portion does not dominate the sample. The disadvantage is that it is less likely that full coverage of the target population will be obtained in the sampled area.

One note of caution is that the non-probability component should be a fairly evenly distributed and diverse sample across areas. Field management should be aware that if the data collection starts in only a few areas, the amount of non-probability sample from those areas should not be excessive and should be monitored and managed so that those areas do not dominate the total non-probability sample for the Field Trial or the Main Survey.

Exhibit 1. Referral Questions from the PISA-D Strand C Screener

Do you know any {other} 14-16 year olds who are out of school or below grade 7 in your neighborhood?

[HELP SCREEN: BY NEIGHBORHOOD WE MEAN {COUNTRY DEFINITION}]

1. YES
2. NO [GO TO END]

980   DON'T KNOW [GO TO Q.END]
990   REFUSED [GO TO END]

I need to collect their name and address so we can contact them to participate in the study.

1. Please tell me the youth's name _____.
2. Please give me their address or directions to where the youth lives.
    _____
3. Any others?

RECORD NAME AND ADDRESS FOR EACH REFERRAL (UP TO TWO) ON SEPARATE BLANK HOUSEHOLD FOLDERS

980   DON'T KNOW
990   REFUSED

School frame approaches - The second option is to identify eligible 15-year-olds through schools. This can be applied in either Stratum H or Stratum L. The first step is to obtain a list of Strand A schools or schools that serve the selected Strand C geographic areas.

Then out-of-school 14 to 16 year olds within the neighbourhood or within a certain distance (e.g., 20 Kms) can be identified and located by the link-tracing referral approach, by teachers or students in the schools. Once addresses are established, the assessment would be conducted where the individual resides. As with the household-linking approach, this is limited to one wave.

Students in grades lower than 7th grade can be identified through administrative lists in schools with 5th and/or 6th grade that either serve the selected areas or are associated with the Strand A schools. Students will be assessed where they live or on a one-by-one basis at the school if more suitable.

The use of Strand A schools for this option has two disadvantages. First, the schools should be contacted during the Strand A field period. Given the short amount of time until the Strand A FT, and since the Strand A FT is the first opportunity to test PISA in the PISA-D countries, it seems advisable to postpone this Strand C option to the MS to avoid any potential impact on piloting Strand A activities.

Second, the Strand C interviewers may not be in the area of the Strand A schools. Therefore, if countries want to implement this option in the FT, we would recommend using schools serving the selected areas.

Street children - Sampling street children will be considered as an option under the nonprobability component after gathering information about this group in each country. Their inclusion in the sample will depend on the distribution and the proportion in each country. Attempts at conducting assessments in some manner in selected areas for this population is encouraged. The link-tracing approach may be useful in identifying youth who reside in places other than households. In addition, a special operation could be conducted in specific locations with a high concentration of youth who sleep in places other than households.

ID assignment – An ID structure will be established so that the source of each sampled or referred youth will be captured. A SourceType variable will be created to indicate whether the individual came from the probability sample, household-link tracing, school link-tracing, or school administrative lists. A unique ID (IDPERS) will be assigned to each sampled or referred person. Finally, a linking ID (IDLINK) will be used to link the referred person to the household or student/teacher that provided the referral.

4. CAUTIONARY NOTES

The goal of our general sample design, as described above, is to satisfy both objectives of Strand C. Probability sampling will be used to evaluate whether feasible options are available for selecting a nationally representative sample of out-of-school 15-year-olds in future cycles of PISA-D. Nonprobability sampling will be used to yield a large enough sample size for both FT and MS. Nonprobability samples capture data at a substantially lower cost than their probability-based counterpart. However, the disadvantage of nonprobability designs is that sampling theory does not hold for making generalisations to the population. Therefore, any conclusions are limited to the sample itself.

The generalisability of the overall MS sample to the target population (or a subset of the population) for Strand C will depend on the results of the FT and the MS. The hope is that the probability portion of the MS will produce a sample representative of the area it covered with sample sizes large enough to support reporting goals for Strand C. It might be possible to develop some modelling approaches that attempt to model the nonprobability sample to some population (based on some assumptions made about the population). However, there will be concerns about biases (misreporting) when such data is used for any reporting purposes. The experiences acquired and results gathered from the FT will be essential in providing us the information needed to ascertain sufficient size of the probability sample for the MS to support reporting goals, as well as the utility of the nonprobability samples. Also, the FT outcome will dictate the balance between probability and nonprobability MS samples across countries.

Currently, our understanding is that ETS will be able to use the entire data (probability and nonprobability samples combined) to support the scaling of item parameters, provided there are at least 900 responses per item (Westat will produce special sampling weights for this purpose). However, data used for main analysis and reporting purposes (for generalisation to Strand C target population) must be based solely on the representative samples with known probabilities.

In addition, countries need to be informed that the MS sample size might be limited when it comes to reporting for various subgroups of the population. Subgroup (e.g., gender, proportions out-of-school versus in grades 6 or below, etc.) sample sizes will be dependent on the reporting categories and distributions in each country. For example, if the OECD or a country is interested in reporting Strand C results by status (i.e., students below grade 7, out-of-school 14- to 16-year-olds) and further by gender by another non-schooling variable (i.e., SES, home language), then a larger sample size may be necessary.

5. SAMPLE DESIGN FORMS AND QUALITY CONTROL

The preceding sections provide a general overview of the sample design approach for PISA-D Strand C. Each country will need to adapt this approach to arrive at the optimal design for their country, while adhering to the PISA-D Strand C standards. As part of the quality control process, countries must submit their national sampling plans to the International Contractor (Westat) for review. To reduce burden, we have created a series of Sample Design forms in a question-and-answer format. The completed forms will serve as a country's sampling plans, and no additional report will be needed.

The Sample Design Forms are to be filled out as a planning document, before compiling sampling frames and implementing sample selection activities. A list of the specific forms is provided in Exhibit 1. They cover all major aspects of the sample design, including the target population, definition of a completed case, sampling frame, sample selection approach, sample sizes, and other sampling topics. The questions in the form address each of the PISA-D Strand C sampling standards, and the relevant standards are included in each form for reference.

The focus of the Sample Design Forms is on the MS, but the forms include text boxes to indicate any aspects of the sample design that differ for the FT. Sampling plans for the MS are considered draft at this point, and countries will have an opportunity to update them after their experience in the FT.

Exhibit 2. Country-Specific Sample Design Forms

Number	Content
SD-1	Target Population and Completes
SD-2PB	Probability-Based Sampling Frame
SD-2NP	Non-Probability-Based Sampling Frame
SD-3PB	Probability-Based Selection Approach
SD-3NP	Non-Probability Selection Approach
SD-4FT	Field Trial Sample Size
SD-4MS	Main Survey Sample Size
SD-5	Miscellaneous

The Sample Design Forms are one of the four submissions relating to quality control (QC) of the PISA-D Strand C sample design and selection. The other three submissions that countries must provide are:

- Sample selection quality control forms;

- Sample monitoring quality control forms; and

- Sample design international file (SDIF)

The sample selection quality control forms are to be submitted at the end of each stage of selection, to help the International Contractor (Westat) verify that the sample selection process was conducted accurately. The sample monitoring quality control forms are completed periodically throughout data

collection and can help countries identify potential shortfalls in the sample, problems in achieving the desired response rate and potential for nonresponse bias in the collected sample. The SDIF is due after data collection and contains sampling information, such as selection probabilities and linking IDs, which the International Contractor will use to perform a final QC check on the sample. In the MS, it will also serve as the input file for weighting and nonresponse bias analysis.

6. REFERENCES

Carr-Hill, R. (2015). Pisa for Development Technical Strand C: Incorporating Out-Of-School 15- Year-Olds In The Assessment. OECD Education Working Paper No. 120.

Heckathorn, D. (2007). Extensions of Respondent Driven Sampling: Analyzing Continuous Variables and Controlling for Differential Recruitment. Social. Methodology, 37, 151-207.

Goodman, L. (1961). Snowball Sampling. Ann. Math. Statist. 32, 148-170.

APPENDIX A. LISTING

When no reliable list of residents, dwelling units (DUs), or addresses is available, the following listing procedures can be used to develop a frame of DUs within selected geographic clusters. The listing effort should be finalised approximately three months prior to data collection, to allow for the sample selection process to be completed and data collection materials generated as needed.

C.1 Listing Preparatory Activities

In preparation for the listing effort, all countries should establish clear definitions of DUs and non-institutional collective DUs (or group quarters) for the listers to follow. A clear definition of the sampling units is necessary to ensure a high-quality sampling frame. As an example, the PISA-D Strand C Technical Standards provide a standard definition for "Basic Dwelling". In addition, all countries should specify the geographic areas in which data are to be collected, and therefore, the areas in which listing will be conducted. All units that satisfy the definition of a DU within the selected areas should be listed.

C.2 Listing Staffing and Management

A listing staff should be identified and used for the listing effort. The quality of the listing staff is of critical importance to the quality of the listings obtained and produced. Countries should hire at least one lister per primary sampling unit (PSU) for the listing operation, depending on the geographic size of the PSU. It is beneficial that listers live within close proximity of the assignment location for cost efficiency purposes. Listers are expected to perform the following duties:

- Attendance at lister training;
- Listing of the assigned areas, including the recording of addresses and the update of maps;
- Editing of materials; and
- Administrative procedures.

Close supervision of the listing staff is required to produce the best quality listings. Typically, there is at least one listing supervisor for every 15-17 listers during the listing operation. The listing supervisor duties include the following:

- Assist with lister training;
- Oversight of the listing operation in the assigned areas; and
- Reporting to the field director on the progress of the listing effort.

Listing supervisors oversee closely the work of their assigned listers. This can be done remotely or on location. All listing supervisors should conduct a weekly reporting telephone call with each of their listers, focusing on the status of the listing effort in the lister's designated area. The following topics are recommended for discussion in the weekly conference calls between the listing supervisor and lister:

- Percentage of area listed;
- Whether the area has experienced a significant amount of new construction or demolition; and

- Problems encountered.

C.3   Lister Training

For a successful listing process, listers should receive adequate training on the terms, procedures, and materials used in the listing effort, in order to ensure that standard listing procedures are used and the sample frame is correctly determined. It is recommended that listers receive a minimum of 2 hours of Home Study training and 10 hours of in-person lister training, with the in-person lister training conducted over 2 days.

The Home Study packet should include an overview of PISA-D Strand C and background on the purpose of listing, as well as written exercises on each of the topics contained in the Listing Manual (described below). These exercises should be collected and reviewed at the in-person lister training session. The Home Study packet should be distributed approximately two weeks prior to in-person lister training.

The necessary components for the in-person lister training include:

- Detailed instruction on the definition of a DU unit (DU), including what types of units will be listed for PISA-D Strand C;
- Instruction on the use of all materials, forms, and maps; and
- Written exercises to practice the listing procedures for a variety of areas (i.e., suburban, urban, and rural).

As part of in-person training, it is recommended that all listers receive a minimum of 3 hours of listing practice, in which listers are required to conduct an actual listing of a defined area, following the standard listing procedures and using the required materials.

Experienced project staff should serve as lead trainers for lister training. Listing supervisors can serve as assistant trainers and evaluate trainees' performance during training. Each country should develop a detailed training agenda for lister training. The agenda should include, at minimum, the topic and a brief description of each training session and the date and length of time of each session. Fully scripting the lister training sessions will ensure consistency of presentation across rooms. Each country should develop a training guide for lister training that includes all lecture scripts and presentations, handouts, written exercises with answer keys, etc. to be used in training.

All listers should receive a Listing Manual to reference throughout the field period. The manual should contain the following information:

- Definitions of listing terms as relevant to the PISA-D Strand C listing effort;
- Step-by-step listing procedures, including instructions on using the maps and completing the forms;
- Specifications for dealing with problem situations, (i.e., multi-unit structures, vacant DUs, mobile homes, new construction, military bases);
- Instructions on making inquiries in the community regarding a structure or area; and
- Administrative procedures.

The Listing Manual can be distributed to the listers in conjunction with the Home Study materials.

C.4   Lister Materials

To ensure that standard listing procedures are used and the sample frame is correctly determined, identical, well-designed, clear, and easy to read materials are needed. A listing sheet on which to record all units that fit the definition of a DU should be developed. Each DU is recorded on a separate line on the listing sheet. The listing sheet includes space for recording the appropriate identifying information for each DU (i.e., street address, characteristics, head of household, etc.). It is recommended that listers are provided with enough listing sheets to list 1.5 times the expected number of DUs. An automated application could be developed to serve as the method for recording information about the DUs, in lieu of paper listing sheets.

To alleviate concerns of government officials and other citizens, an authorisation letter should be developed. The authorisation letter, written by the Survey Institute, should introduce the lister and explain the lister's task and reason for the lister's presence in the area. Listers should carry the authorisation letters with them whenever conducting the listing effort. An adequate supply of authorisation letters should be provided to each lister, allowing them to provide copies to concerned individuals as needed. To help listers gain entrance into locked apartment buildings, it is suggested that the Survey Institute develop a letter explaining the purpose of the listing effort, to provide directly to building managers, superintendents, or governing boards, if questioned.

Additionally, the Survey Institute should develop maps of the area to be listed that reflect the current conditions of the area. The map would include the area to be listed and the surrounding areas, including features such as streets, highways, rivers, railroads, and lakes. The map of the area to be listed should include the count of the expected number of DUs in the area. A separate map, consisting of an inset of the area being listed, should be produced for special situations in which a portion of the map is too dense, has too many streets, or in which there is not enough space to print the street names or other descriptive information. The Survey Institute should consider purchasing commercial maps to help listers locate the area to be listed or distinguish boundaries.

C.5   Listing Procedures

A standardised approach to listing ensures that the same listing procedures are being implemented across the country. The following listing procedures can be used to develop an accurate listing of all DUs within the specified area(s).

The listing procedure involves listers recording all places where people live, or might live, within the boundaries of the identified area on the map. Each DU in the specified area must be accurately and completely identified. If the DU does not have a street or mailing address, the lister records the location, description, and distinguishing characteristics of the DU on the listing sheet. The head of the household or the most visible or known person in the DU can also be recorded, although it is possible this information could become outdated by the time of interviewing. Each DU unit is listed on a separate line/row of the listing sheet. If the house or apartment number of the DU is missing, the lister carefully describes the location of the DU, including drawing a sketch if needed.

In an apartment building, apartments are listed in order of number. A standardised procedure for listing apartments that do not contain numbers should be implemented. For instance, the lister should list in the following manner: bottom floors to top floors (basement to upper floors); right to left within each floor; and front to back on any given side.

The maps are updated in the field while conducting the listing effort. All maps are marked to show exactly what was found in the segment. Discrepancies between the map and what is observed while conducting the listing should be indicated directly on the map. Listers begin the listing effort for a designated area in the northwest corner, and place an 'X' on the map to indicate where they started. They then show their route of travel with arrows drawn directly on the map. The listers should record the listing sheet line numbers of the first and last street number on each street or boundary. If no DU is located on a street or boundary, an indication of this should be made on the map.

C.6   Quality Control

As part of the quality control monitoring process, countries conducting listing will be required to document the following:

- Preparatory activities in advance of the listing effort;
- Progress in staffing for the listing effort;
- Details and results of lister training;
- Materials to be used in the conduct of the listing effort;
- Procedures to be used in the conduct of the listing effort; and
- Quality control measures implemented in the conduct of the listing effort.

Throughout the listing effort, quality control of the lister's work must be conducted. Material review and field observation are discussed below.

Survey Institute review of maps and listing sheets should also be conducted during the listing effort by trained reviewers. The Survey Institute should review every completed map and listing sheet that is returned to the home office. The home office review process should involve the comparison of the maps with the listing sheets to ensure that all roads and boundaries have been canvassed and that map work and listing sheets correspond. Maps and listing sheets that cannot be reconciled as part of the review process should be reassigned to another lister to complete.

Coverage Enhancement

Prior to conducting data collection in an assigned area, interviewers are trained to conduct additional quality control efforts to check for any structures that were missed or DUs that were 'hidden' during the initial listing phase. This ensures that newly constructed DUs or DUs accidentally excluded during the listing phase are eligible for inclusion in PISA-D Strand C during the data collection effort. In the case of area sampling with a systematic selection of DUs, one valid approach for the missed structure procedure is to look for missed structures in sampled clusters where the first line number in the systematic sample in the cluster was selected. In those clusters, a complete canvas of the cluster needs to be completed and any missed DUs found during the missed structure procedure should be added to the sample. If too many are found, then a sample can be taken and the weights computed

appropriately. Another common approach is the half-open interval, for which a general description can be found on the Cross-Cultural Survey Guidelines website (Appendix C of Chapter V http://ccsg.isr.umich.edu/sampling.cfm). A hidden DU procedure should be used to look for missed DUs within a structure. The process can be completed within a structure if the first DU within a structure is selected under a systematic sampling process. A representative of the associated structure is asked if there are any hidden DUs. All hidden DUs found are then added to the sample. If too many hidden DUs are found for a particular structure, then a sample is taken.