SAMPLING IN PISA

Doc.:
CY7_NPM(1603)AB_SMP_SamplinginPISA_1.docx

First Meeting of the PISA 2018 National Project Managers

14 – 18 March 2016

Prague, Czech Republic

Produced by Westat, Core C Contractor

SAMPLING IN PISA

Section 1 - Target Population and Sampling Standards

The operations in PISA are based on meeting the PISA standards.

Several of the standards refer to the PISA target population and to matters of sampling. Those for the 2018 Main Survey are found in the Appendix. The related sections following this one discuss the population definition, exclusions, and response rates.

Section 2 - Overview of the Main Survey

The sampling plan for PISA begins with the definition of the PISA Target Population. The Target Population for PISA is students between 15 years and 3 (completed) months and 16 years and 2 (completed) months at the beginning of the testing period, attending educational institutions located within the country, and in grade 7 or higher.

Note that since the largest part of the PISA target population is made up of 15-year-olds, then "15-year-olds" is the term often used when referring to the PISA target population. **Where you see the term "15-year-olds" in this document, think "15-year olds and 16-year olds which are part of the PISA birth date population definition".**

Historically the age definition for PISA arises from operational considerations for the 2000 assessment. It was desired by the OECD and the participating countries that the assessments should take place in about April of 2000. For ease of implementation it was decided that the population to be surveyed in April 2000 would be of students born in 1984. This was the basis of the PISA definition of "15-year-olds", and the relationship between the birth dates of eligible students, and the timing of the assessment.

The student birth date definition needs to be adjusted according to the chosen time of testing so that the target population is between 15 years and 3 months and 16 years and 2 months of age at the beginning of the time of testing. This means that if the assessment is to be conducted throughout the month of April 2018, for example, the eligible population is defined as students born during 2002. If the testing is to take place in June 2018, the population is students born between March 2002 and February 2003 inclusive.

Variation of up to one month in this age definition is permitted so long as the birth date definition is maintained as a 12 month period. In particular, if the testing period is any 6 week period for PBA countries or 8 week period for CBA countries, between March 1, 2018 and May 31, 2018, the birth date population may be defined as students born in 2002. If no local factors dictate to the contrary, countries are encouraged to test within this time period and to use this population definition.

You will need to be alert to ensure that possible drift in the assessment period does not lead to an unacceptable birth date population definition. For example, a National Project Manager (NPM) might propose to test during the month of May, students born during 2002. This would be acceptable. But if in fact the testing period slips to become May to mid-June, the population will need to be changed to students born February 2002 to January 2003 (or March 2002 to February 2003 if preferred), as it is not

acceptable to have testing in June with a birth date definition of students born in 2002. The FT Sampling Guidelines has more details on this.

Note that the MS PISA population definition and testing time should be decided first, and then the FT PISA population and testing time defined secondly, ensuring that there is no overlap between FT and MS students.

2.1. Exclusions

Usually, practical reasons are invoked for excluding schools and students, such as increased survey costs, increased complexity in the sample design and/or difficult test conditions.

Exclusions can occur at the school level, i.e., entire schools are excluded, or within schools, i.e., specific sampled students within sampled schools are excluded. All such exclusions, at both the school level, and the within-school level, will need to be described and quantified for approval.

In PISA, exclusions from the target population are to be kept to a minimum. After all exclusions (school and within-school), the resultant population is required to cover at least 95% of the desired target population.

Difficulties raised by the existence of small schools, (i.e., those with few PISA-eligible students), or other schools where it is difficult or costly or not feasible to conduct assessments, are mainly addressed by modifying the sample design to reduce the number of such schools selected, rather than by exclusion.

Similarly, the exclusion of special education students and students with insufficient assessment language experience is to be kept to a minimum.

2.1.1. School-Level Exclusions

A school may be excluded if it provides instruction only to students in the excluded categories defined under “within-school exclusions” (described in the next section), such as schools for the blind. The percentage of PISA-eligible students enrolled in such schools should be less than 2 percent.

An additional 0.5% of students may be excluded in schools as agreed upon with Westat, Core C, such as those in remote areas, as one example. Students in a very small language group provide another example.

2.1.2. Within- School Student l-Level Exclusions

Within schools, all PISA-eligible students as defined by the population birthdates should first be listed. Sampled students who are deemed as excluded **will need to be retained**, and a variable maintained to briefly describe the reason for exclusion (see the PISA *School Coordinator* and *Test Administrator* manuals and the *Within-School Sampling* manual). Using this method, the size of the within-school exclusions can be well-estimated from the sample data, which is a requirement.

International within-school student exclusion rules are specified as follows:

- The student is functionally disabled in such a way that he/she cannot participate in the PISA testing situation. Functionally disabled students are those with a moderate to severe permanent physical disability. Functionally disabled students who can respond should be included in the testing.

- The student has a cognitive, behavioural or emotional disability such that in the opinion of qualified staff, he/she cannot participate in the PISA testing situation. These are students who are cognitively, behaviourally or emotionally unable to follow even the general instructions of the assessment. Students should not be excluded solely because of poor academic performance or normal discipline problems.

- The student has insufficient assessment language experience to take the PISA test. Students who have insufficient assessment language experience are those who meet all the following three criteria:
  - they are not native speakers in the assessment language,
  - they have limited proficiency in the assessment language, and
  - they have received less than one year of instruction in the assessment language.

- The student is not assessable for some other reason as agreed upon.

- There are no test materials available in the student's language of instruction.

The exact extent of within-school exclusions will not be known until the within-school sampling frames (student lists) have been returned from the participating schools and sampling weights computed.

Estimates of the extent of within-school exclusions are therefore required from the NPM to ensure that the national defined target population will cover at least 95 percent of the national desired target population.

At the time of school sampling frame creation, the NPM will be asked to provide his/her best estimate of the extent of within-school exclusions by type of exclusion and report the estimates on the Sampling Task 7b form.

Although it is expected that the four defined categories should cover all types of within-school exclusions, it may be the case that a particular country has one additional category. This needs to be limited to special circumstances as defined by the NPM and approved by the Contractors via the process of agreeing to adaptations to manuals.

2.2. Response Rates

2.2.1. School-Level Response Rates

PISA requires a minimum weighted participation (response) rate of 85% of originally sampled schools. However, nonparticipating sampled schools may be substituted with "replacement schools" identified at the time of school sampling to meet sample size and response rate requirements. The use of replacement schools does not guarantee that potential biases have been reduced. Therefore, you are encouraged to persuade as many original sampled schools as possible to participate; only a high participation rate of original sampled schools will minimize the potential for response bias.

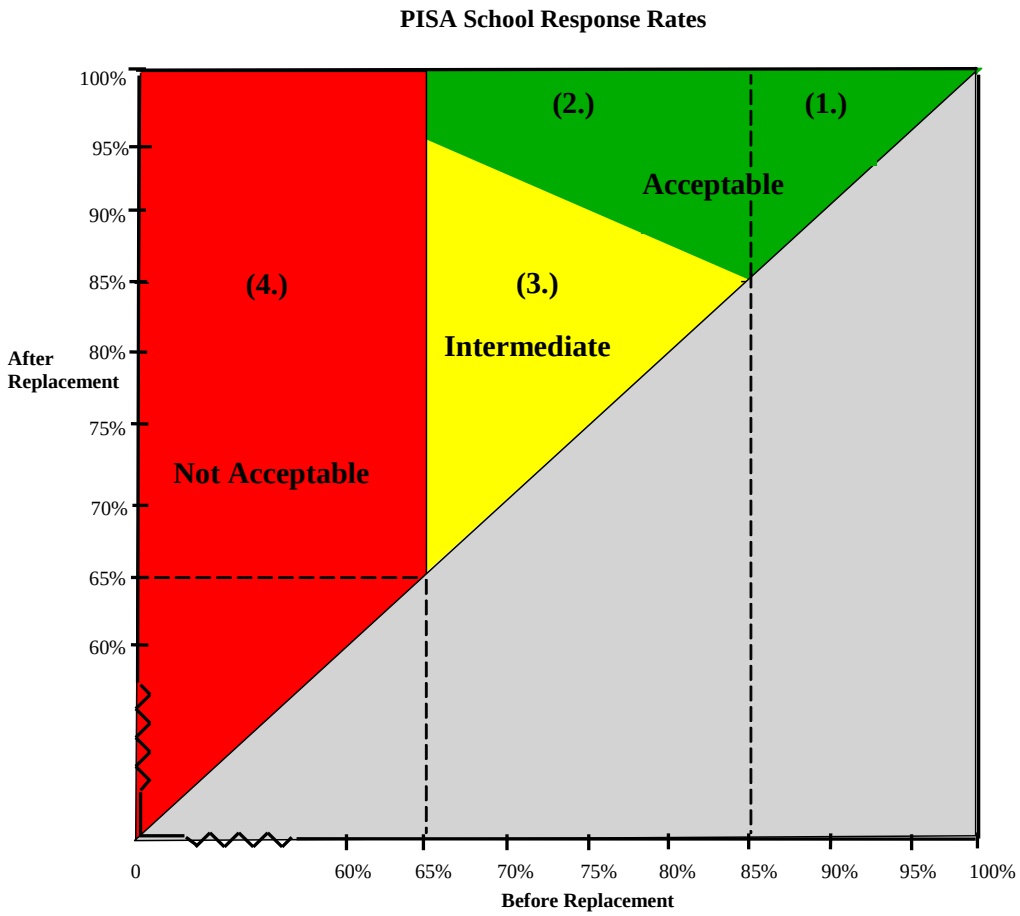
Note that raising participation/response rates through the use of replacements improves yield, but that a given response rate achieved through the use of replacements is not as good as that same rate achieved without replacements.

That is, a school participation rate of x% before replacement has less potential for biases than the same response rate of x% achieved only after replacement.

Therefore, all other factors being equal, acceptability of the country's data in international comparisons will be relative to both the school participation rate of originally sampled schools and the response rate achieved with the use of replacements.

Three response rate zones -- acceptable, intermediate and not acceptable –- are defined. "Acceptable" means that the country's data will be included in international comparisons. "Not Acceptable" means that the country's data will be a candidate for not being reported in international comparisons, and will be included only if the NPM provides considerable evidence that nonresponse bias is likely to be minor. Finally, the "Intermediate" zone means that a decision on whether or not to include the country's data in international comparisons needs to be made while taking into account various other factors.

These zones are a function of the response rate before replacement, and the response rate after replacement. There are two sets of combinations of before and after response rates that are noted as being in the "acceptable" zone. Area (1.) in the chart below indicates the case where the response rate is acceptable based on the original sample only. Area (2.) indicates where replacement schools were needed in order to achieve an acceptable rate.



Note that a school with less than 50% participation among the selected eligible and non-excluded students will not be considered as a participating school.

If such a school has less than 25% participation among the selected eligible and non-excluded students, then the students in such schools will not be included in analysis.

If such a school has more than 25% and less than 50% participation among the selected eligible students, the students in these schools will be retained for analysis, even though the school is considered a non-participant for the purposes of monitoring response rates.

2.2.2. Student-Level Response Rates

PISA also requires, nationally, a minimum weighted participation rate (weighted) of 80 percent of students within participating schools (originally sampled and replacement).

Follow-up sessions may be required in schools where too few students participated in the originally scheduled test sessions. Guidelines for determining when follow-up sessions are appropriate are found in the PISA National Project Managers Manual. The task of determining whether follow-up sessions are required will be delegated to the School Coordinators and Test Administrators. As they oversee the actual test sessions, they will be able to promptly determine whether follow-up sessions are required at each school.

Student participation rates are calculated over all participating schools, whether originally sampled schools or replacement schools, and from the participation of students at the originally scheduled session and any follow-up session that may be required. The student participation rate requirement needs to be met at the national level, not necessarily for each participating school.

2.3. Sample Design

PISA uses a stratified two-stage sample design, where schools are sampled using probability proportional to size (the school enrolment of 15-year-olds) sampling, and students are sampled with equal probability within schools. Sampled students receive a final weight which indicates how many other students from the population are represented. The final student weight incorporates both the school weight (the inverse of the school's probability of selection) and the within-school student weight (the inverse of the student s probability of selection).

Why are schools stratified (divided into groups)? In sampling theory, the sampling variance can be reduced from that obtained by a simple random sample of schools, if schools within strata are thought to be fairly alike, and schools from different strata are thought to be quite different.

Why are schools sampled as the first stage unit and students sampled as the second stage unit? One of the main reasons for sampling students in this way is that often countries do not have a national list of enrolled PISA students available. The following points indicate additional reasons for this method of student sampling.

Consider a population of 400 students in 10 schools.. Suppose we select a sample of 40 students from the 400, ignoring which schools they are in. Table 2 shows what the sample could look like.

Table 1

School ID	# of students selected
1	4

2	9
3	0
4	1
5	0
6	2
7	4
8	10
9	0
10	10
Total	40

Suppose that out of the 10 schools seen in Table 1, school 8 and 10 both have all students sampled, and that five other schools have varying numbers of students sampled. The disadvantages of this design are that 7 out of 10 schools would have to be visited, and several schools would need to be visited for very few students while others would need to be visited to see all their students.

To remedy one of these disadvantages, a pre-determined number of schools is sampled first in PISA thereby controlling the number of schools that need to be visited to reach the required student sample size. Within each sampled school, an equal number of PISA students are sampled, usually 35 students for PBA countries and 42 students for CBA countries. This controls the second disadvantage noted above that in some schools all or most students would need to be assessed (think about very large schools).

As noted earlier, schools are sampled with probability-proportional-to-size sampling while students are sampled with equal probability within schools. This method of sampling has many nice qualities. In a perfect world, all sampled students would have the same final weight, thus helping to minimize sampling variability. Additionally, the sum of the final student weights should equal the number of students in the population. If schools were sampled with equal probability, neither of these nice properties would occur. Thus, the combination of probability-proportional-to-size sampling of schools, together with equal probability sampling of a fixed number of students within each school, is designed to yield a sample in which all categories of students are represented in proportion to their size in the population.

For PISA 2018, the required sample size needs to aim for at least 6300 assessed students for CBA countries and for 5250 assessed students for PBA countries from a minimum of 150 sampled and participating schools. In small countries that do not have this many students, a census of students should be conducted.

Any countries that had larger than average sampling variances in PISA 2015, or which did not achieve the minimum required number of assessed students will have additional work done on their sample design so that these situations can be remedied in PISA 2018. This may mean selecting more than the minimum number of students.

2.4. School Sampling

2.4.1. The Frame (School List)

The international desired target population is intended to provide full coverage of all PISA-eligible students attending an educational institute within the country. All NPMs are required to construct a school sampling frame covering their target population. It is from this list that the school sample will be selected.

This means that countries are to include PISA-eligible students who attend on only a part-time basis, are in vocational training or other non-general types of programmes, or any other related type of educational programme, or who are in foreign or International schools within the country, **even if they are not included in other international or national studies**.

Therefore, all schools located within a country **with the potential to have one or more** 15-year-old students in grades 7 or above at the time of testing should be listed on the school frame for sampling:

- urban and rural, domestic and foreign schools, national and international, and tribal schools;
- public and private (including charter, parochial, parish schools);
- grade 7 and above (including middle and secondary schools, graded and ungraded schools);
- mainstream, vocational training, special education, or any other special types of programmes or tracks;
- schools for boys and girls or single sex schools; and
- schools for full-time or part-time students and for students enrolled in day/night/evening programmes.
- Non-educational institutions are not part of the study population. Some examples of institutions NOT eligible for PISA are homebound schools, correspondence schools, or weekend driving schools.

A well-constructed school sampling frame is one that provides complete coverage of the target population without being contaminated by incorrect entries, duplicate entries, incomplete entries, or entries that refer to elements that are not part of the defined target population.

Initially, this list should include **any school that could have any PISA-eligible students**, even those that may later be excluded, or those which currently do not have any PISA-eligible students.

The construction of a school sampling frame depends to a great extent on the availability of appropriate information about schools and students.

Each school entry on the frame should include at minimum:

- School identification information, such as a unique national ID, and contact information such as name, address, phone number, etc.

- A suitable measure of the approximate enrolment of PISA-eligible students. In order of suitability, examples of this are:
  - Current school enrolment of 15-year-olds; Current enrolment data, however, is rarely available at the time schools are sampled. If none of the types of enrolment listed below are available, or if the available enrolment data are too out of date, schools may have to be selected with equal probabilities, which may in turn require an increased school sample size.

  - If 15-year-olds tend to be enrolled in two or more grades, and the proportions of students who are 15 in each grade are approximately known, the 15-year-old enrolment can be estimated by applying these proportions to the corresponding grade level enrolments; For example, suppose that grades 9 and 10 have the majority of PISA students, and 40% of PISA age students are in grade 9 and 50% of PISA age students are in grade 10. If you have for school X on the school frame, that the school has 185 students in grade 9 and 250 in grade 10, then your estimate for PISA enrolment for school X, if done using this method, would be (0.40 x 185)+(0.50x250) = 199.

  - The grade enrolment of the modal grade for 15-year-olds; and

  - Total student enrolment, divided by the number of grades in the school.

- Coded information about the school, such as region of country, school type (public or private), urban/rural classification, etc. which may have some effect on the assessment outcomes, and can be used as stratification variables (variables used for dividing the population into mutually exclusive groups so as to improve the precision of sample-based estimates).

The quality of the sampling frame has a direct effect on the survey results, and thus frame construction should be treated with extreme care. As noted, an approximate enrolment of PISA-eligible students (i.e.15-year-olds) needs to be associated with each school listed in the sampling frame. The quality of a sampling frame will, to a large extent, depend on the accuracy of this data because school selection probabilities are based on this quantity.

Westat, Core C will assess the extent of missing or inaccurate data used for stratification and for creating the measure of size needed for the school sampling method, as these will increase sampling error.

2.4.2. School Sample Selection

As in the past, the school selection will be done by Westat, Core C and its sub-contractor, ACER, based on a design agreed upon between each country's NPM and the Westat, Core C.

School sampling by the contractors was first implemented for PISA 2003 at no extra cost to countries and was also done in this way for PISA 2006, PISA 2009, PISA 2012, and PISA 2015. This is also the plan for PISA 2018.

From the Contractor's point of view, there have been obvious benefits from sampling schools in this way in terms of efficiency, timeliness, and quality.

Again in PISA 2018 there will be no additional cost to countries unless a country's national options result in a sample design too complex for this to be easily done.

This school selection procedure, shared between Westat, Core C and ACER, will require that you keep Westat, Core C and ACER closely informed about when you need to have your school samples. Meeting each country's timing requirements can only be done if all required information and an adequate sampling frame (school list) of schools are submitted in a timely and accurate fashion.

2.4.3. Replacement Schools

As noted earlier, you need to make every effort to get as many of the sampled schools to participate as possible. However, it is not always possible to obtain the participation of all sampled schools.

In order to avoid the resulting sample size losses, a mechanism to identify **in advance** replacement schools for non-participating sampled schools is needed. A second, perhaps more important, reason for identifying replacement schools in advance is to avoid the haphazard use of alternate schools as replacements, which may actually amplify response biases.

Although the approach used is no guarantee of avoiding non-response biases, it will at least tend to minimise the potential for bias. The technique of identifying replacements should lead to less non-response bias than the alternative of drawing a larger sample initially, in anticipation of non-response.

Each sampled school in the main survey will be assigned two replacement schools, if possible, on the sampling frame.

Replacement schools should only be used if the originally sampled school refuses to participate or otherwise does not participate for a reason OTHER than ineligibility or exclusion. Ineligible or excluded originally sampled schools should not be replaced.

2.4.4. Treatment of Small Schools

Small schools in the sample can result in a reduced sample size of students for the national sample, below the desired target. Alternatively, the sample may contain many small schools, which is an administrative burden.

In order to minimise these problems, the small schools in the sampling frame may require undersampling (relative to other schools in the population) depending on the proportion of students who are in these schools.

Additionally, as in previous PISA cycles, extremely small schools with exactly one or two students on the school sampling frame will be retained on the frame for sampling (as opposed to being excluded from the frame). If any of these extremely small schools are sampled, they can be excluded in the field, IF the school has only 1 or only 2 PISA students, and IF the expected level of such exclusions is less than 0.5% of the population and IF this would not result in an overall level of exclusion in excess of 5% and IF this is approved by Westat, Core C.

2.5. Within-School Sampling

As in the past, students within sampled schools are to be sampled with equal probability using the Contractor software, *KeyQuest*.

Within sampled schools, students will need to be sampled from a complete list of all PISA students in the school, no matter what classroom they belong to, nor what grade or gender they belong to, nor what programme or track or shift/attendance session they belong to.

That is, student sampling for PISA has always been, and continues to be, **direct student sampling.** Similar to the previous PISA cycles, class based sampling will not be acceptable for obtaining a sample of 15-year-olds since this would adversely affect the ability of the study to provide variance components within and between schools.

For all countries, but especially for any country implementing any International and/or national options, using *KeyQuest* becomes crucial because of higher reliance on the output from *KeyQuest* for the weighting of the student samples. This applies even to the class-grade sample option and in fact, more so for that option.

Usually the number of students selected per school (called the target cluster size (TCS)) is 35 for PBA countries and 42 for CBA countries, but this can vary among countries, and even among strata within a country. The number of students selected in each school needs to be at least 25. In all sampled schools that have fewer than TCS PISA students, all PISA students need to be sampled, even if there is only one student.

2.6. Submission of Required Information to Westat, Core C

For PISA 2018, as for previous PISA cycles some required information will need to be uploaded as sampling forms to the PISA Portal.

Section 3 - Field Trial for PISA 2018

The Field Trial should occur in the period March – June 2017 with appropriate birth date definitions. The Field Trial birth date definition should not overlap the definition to be used in the Main Survey, so as to eliminate the possibility that an individual student could be sampled for both studies.

For countries with multiple language groups, a Field Trial should occur in a language if that language group represents more than 5 percent of the total eligible population.

KeyQuest will need to be used to select the within-school student samples.

The number of schools required will depend on the language composition of the country as well as on any International and/or national options that may be implemented. A few schools in addition to the required number (25 for PBA countries and 28 for CBA countries) should be sampled as replacement schools since some may not respond.

If a country will implement any International or national option for the Main Survey which involves within-school sampling, these options need to be tested in the Field Trial.

If the option under consideration is grade sampling, the country should be fairly certain that it will implement grade sampling in the Main Survey before testing this in the Field Trial, as testing this option in the Field Trial increases the undesirable risk of having the same students in the Field Trial sample and the Main Survey sample.

The sample for the Field Trial is a convenience sample of schools chosen by each country. Although not a probability sample, there are guidelines that need to be followed.

In particular, the sample should cover different tracks (vocational, business, general, etc.) in the school system where any significant proportion of the 15-year-olds are in school.

The sample should cover schools with different grades that contain any significant proportion of the 15-year-olds.

The sample should cover different demographic and socio-economic groups in the population of enrolled 15-year-olds (e.g. different geographic regions, urban and rural regions, public and private schools, etc.).

Various Sampling Task forms need to be submitted to and approved by Westat, Core C.

Appendix

Standard 1.1 The *PISA Desired Target Population* is *agreed upon* through negotiation between the National Project Manager and the international contractors within the constraints imposed by the definition of the *PISA Target Population*.

Standard 1.2 Unless *otherwise agreed upon* only *PISA-Eligible students* participate in the test.

Standard 1.3 Unless *otherwise agreed upon*, the *testing period*:

- is no longer than eight consecutive weeks in duration for computer-based testing participants,
- is no longer than six consecutive weeks in duration for paper-based testing participants,
- does not coincide with the first six weeks of the academic year, and
- begins exactly three years from the beginning of the *testing period* in the previous PISA cycle

Standard 1.4 Schools are sampled using *agreed upon*, established and professionally recognised principles of scientific sampling.

Standard 1.5 Student lists should not be collected more than 8 weeks prior to the start of data collection, unless otherwise *agreed upon*.

Standard 1.6 Students are sampled using *agreed upon*, established and professionally recognised principles of scientific sampling and in a way that represents the full population of *PISA-Eligible students*.

Standard 1.7 The *PISA Defined Target Population* covers 95% or more of the *PISA Desired Target Population*. That is, *school-level exclusions* and *within-school exclusions* combined do not exceed 5%.

Standard 1.8 The student sample size for the **computer-based mode** is a minimum of 6300 assessed students, and 2100 for *additional adjudicated entities*, or the entire *PISA Defined Target Population* where the *PISA Defined Target Population* is below 6300 and 2100 respectively. The student sample size of assessed students for the **paper-based mode** is a minimum of 5250.

Standard 1.9 The school sample size needs to result in a minimum of 150 participating schools, and 50 participating schools for *additional adjudicated entities*, or all schools that have students in the *PISA Defined Target Population* where the number of schools with students in the *PISA Defined Target Population* is below 150 and 50 respectively. Countries not having at least 150 schools, but which have more students than the required minimum student sample size, can be permitted, *if agreed upon* , to take a smaller sample of schools while still ensuring enough sampled PISA students overall.

Standard 1.10 The final weighted school response rate is at least 85% of sampled eligible and non-excluded schools. If a response rate is below 85% then an acceptable response rate can still be achieved through *agreed upon* use of replacement schools.

Standard 1.11 The final weighted student response rate is at least 80% of all sampled students across responding schools.

Standard 1.12 The final weighted sampling unit response rate for any optional cognitive assessment is at least 80% of all sampled students across responding schools.

Standard 1.13 Unless *otherwise agreed upon*, the international contractors will draw the school sample for the Main Survey.

Standard 1.14 Unless *otherwise agreed upon*, the National Centre will use KeyQuest to draw the student sample, using the list of eligible students provided for each school.

Note 1.1 Standards 1.1 through 1.14 apply to the Main Survey but not the Field Trial.

Note 1.2 Data from schools where the student response rate is greater than 25% will be included in the PISA dataset.

Note 1.3 For the purpose of calculating school response rates, a participating school is defined as a sampled school in which more than 50% of sampled eligible, non-excluded students respond.

Note 1.4 Acceptable response rates obtained through the use of replacement schools are described in detail in the School Sampling Preparation Manual.

Note 1.5:

Guidelines for acceptable exclusions that do not affect standard adherence, are as follows:

- School level exclusions that are exclusions due to geographical inaccessibility, extremely small school size, administration of PISA would be not feasible within the school, and other agreed upon reasons and whose students total to less than 0.5 % of tl PISA Desired Target Population,

- School level exclusions that are due to a school containing only students that would be within-school exclusions and that total to less than 2.0 % of the PISA Desired Target Population, and
- Within-school exclusions that total to less than 2.5 % of the PISA Desired Target Population – these exclusions could include, for example, students not able to do the test because of a functional disability.

• Note 1.6 Principles of scientific sampling include, but are not limited to:

- The identification of appropriate stratification variables to reduce sampling variance and facilitate the computation of non-response adjustments.
- The incorporation of an agreed target cluster size of PISA-Eligible students: The minimum acceptable target cluster size is 25 students. In determining the target cluster size for a given country, or stratum within a country, it is necessary to ensure that the minimum sample size requirements for both schools and students will be met.

Note 1.7 Any exceptional costs associated with verifying a school sample taken by the National Centre, or a student sample selected other than by using KeyQuest will be borne by the National Centre.

Note 1.8 Agreement with the international contractor on alternative methods of drawing samples will be subject to the principle that the sampling methods used are scientifically valid and consistent with PISA's documented sampling methods. Where a PISA participating country chooses to draw the school sample, the National Centre provides the international contractor with the data and documentation required for it to verify the correctness of the sampling procedures applied. Where a PISA participating country chooses not to use KeyQuest to draw the student sample, the National Centre provides the international contractor with the data and documentation required for it to verify the correctness of the sampling procedures applied.