Chapter 5

TIMSS 2007 Sample Design

Marc Joncas

5.1 Overview

This chapter describes the TIMSS 2007 sample design, which consists of a set of specifications for the target and survey populations, sampling frames, survey units, sample selection methods, sampling precision, and sample sizes. The sample design is intended to ensure that the TIMSS 2007 survey data provide accurate and efficient estimates of national student populations. Since measuring trends is a central goal of TIMSS, the sample design also aims to provide accurate measures of changes in student achievement from cycle to cycle. In addition to the sample design, the TIMSS 2007 sampling activities also include estimation procedures for sample statistics and procedures for measuring sampling error. These other components are described in Chapters 9. The basic TIMSS sample design has two stages: schools are sampled with probability proportional to size at the first stage, and one or more intact classes of students from the target grades are sampled at the second stage.

All participants followed the uniform sampling approach specified by the TIMSS 2007 sample design with minimum deviations. This ensured that high quality standards were maintained for all participants, avoiding the possibility that differences between countries in survey results could be attributable to the use of different sampling methodologies. This uniform approach also facilitated an efficient approval process of the national designs by the international project team.

The TIMSS 2007 National Research Coordinator (NRC) of each participating country was responsible for implementing the sample design, including documenting every step of the sampling procedure for approval by the TIMSS & PIRLS International Study Center prior to implementation. To support NRCs in their sampling activities, a series of manuals: the *TIMSS 2007 School Sampling Manual, TIMSS 2007 Survey*

Operations Procedures Unit 2, and *TIMSS 2007 School Coordinator Manual* (TIMSS & PIRLS International Study Center, 2005, 2006a, 2006b) and sampling software (IEA Data Processing and Research Center, 2006) were provided. In addition to these materials, Statistics Canada and the Sampling Unit at the IEA DPC consulted with each country throughout the process.

5.2 TIMSS Target Populations

TIMSS 2007 chose to study achievement in two target populations—the fourth and eighth grade in most countries. Participating countries were free to select either population or both. The target populations can be seen as a collection of units to which the survey results apply. The main groups of interest in TIMSS are student populations (since by-products of the selection methods, schools and classes, also can be considered as populations). The formal definitions of the TIMSS target populations make use of UNESCO's International Standard Classification of Education (ISCED) (UNESCO Institute for Statistics, 1999) in identifying the appropriate target grades:

Fourth grade population. This includes all students enrolled in the grade that represents 4 years of formal schooling, counting from the first year of ISCED Level 1, provided that the mean age at the time of testing is at least 9.5 years. For most countries, the target grade should be the fourth grade or its national equivalent.

Eighth grade population. This includes all students enrolled in the grade that represents 8 years of formal schooling, counting from the first year of ISCED Level 1, provided that the mean age at the time of testing is at least 13.5 years. For most countries, the target grade should be the eighth grade or its national equivalent.

The rationale behind these definitions is as follows:

Since the aim of TIMSS is to improve student learning in mathematics and science, it is crucial to be able to link student achievement to school practices and educational policies, most of which are tied to grade levels. TIMSS grade-level results must be as directly useful as possible for educational purposes.

To be educationally useful, the amount of schooling represented by the grade assessed should be comparable across countries. Therefore, the focus should be on comparing student achievement after the same amount of schooling.

Based on previous cycles of TIMSS and PIRLS, the grades assessed in TIMSS should represent 4 years and 8 years of formal schooling.

The procedure for identifying the first grade from which to begin counting years of schooling should be based on an internationally accepted classification scheme. As mentioned above, such a scheme exists in UNESCO's ISCED.

In IEA studies, the above definitions correspond to what is known as the *international desired target populations*. All students enrolled in the appropriate target grades, regardless of their age, belong to the international desired target populations. All schools of all education subsystems that have students learning full-time in the appropriate target grades are part of the international desired target populations. Schools that do not contain the target grades are automatically excluded from the study. Each participating country was expected to define their *national desired target population*s to correspond as closely as possible to these definitions. In order to measure trends, it was critical that countries that participated in previous TIMSS cycles chose the same target grades for TIMSS 2007 that were used in the previous cycles. Information about the target grades in each country is provided in Chapter 9.

Although countries were expected to include all students in the target grades in their definitions of the population, sometimes it was not possible to include all students who fell under the definition of the international desired target populations. Consequently, based on geographic or linguistic constraints, a country's *national desired target population* excluded some sections of the population occasionally. For example, Lithuania's national desired target populations included only students in Lithuanian-speaking schools, representing respectively, 93 and 92 percent of the fourth and eighth grade international desired populations of students in the country.

Working from the national desired population, each country had to operationalize the definition of its population for sampling purposes and define their *national defined population*. While these national defined target populations ideally should coincide with the national desired target populations, in reality, there may be some regions or school types that cannot be included. All students in the desired populations who are not included in the defined populations are referred to as the excluded populations.

TIMSS participants were expected to ensure that the national defined populations included at least 95 percent of the national desired populations of students. Exclusions (which had to be kept to a minimum) could occur at the school level, within the sampled schools, or both. Although countries were expected to do everything possible to maximize coverage of the national desired populations, *school-level exclusions* sometimes were necessary. Keeping within the 95 percent limit, school-level exclusions from the sampling frame could be for the following reasons:

- Schools were geographically remote.
- They had very few students.
- The curriculum or structure at the school was different from the mainstream education system.
- Schools were specifically for students with special needs.

The difference between these school-level exclusions and those at the previous level is that these schools were included as part of the school sampling frame (i.e., the list of schools to be sampled). They then were eliminated on an individual basis if it was not feasible to include them in the testing.

In many education systems, students with special educational needs are included in ordinary classes or grouped together in special classes within ordinary schools. Due to this fact, another level of exclusions is necessary to reach an effective target population—the population of students who ultimately will be tested. These are called *within-school exclusions* and pertain to students who are unable to be tested for a particular reason but are part of a regular classroom or part of an in-scope school. There are three types of within-school exclusions, which are explained below:

- **Students with intellectual disabilities**. These are students who are considered, in the professional opinion of the school principal or by other qualified staff members, to be intellectually disabled or who have been tested psychologically as such. This includes students who are emotionally or mentally unable to follow even the general instructions of the test. Students should not be excluded solely because of poor academic performance or normal disciplinary problems.

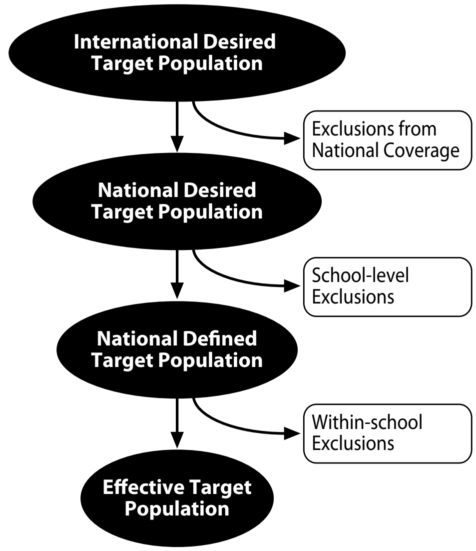
- **Students with functional disabilities.** These are students who are physically disabled in such a way that they cannot perform in the TIMSS testing situation. Functionally disabled students who are able to perform should be included in the testing.
- **Non-native language speakers**. These are students who are unable to read or speak the language(s) of the test and would be unable to overcome the language barrier of the test. Typically, a student who has received less than 1 year of instruction in the language(s) of the test should be excluded, but this definition may need to be adapted in different countries.

Students eligible for within-school exclusion were identified by staff at the schools and still could be administered the test if the school did not want the student to feel out of place during the assessment (though the data from these students were not included in any analyses). Again, it was important to ensure that these populations were as close to the national desired target populations as possible.

If combined school-level and within-school exclusions exceeded 5 percent of the national desired target population, results were annotated in the TIMSS 2007 international reports (Martin, Mullis, & Foy, 2008; Mullis, Martin, & Foy, 2008). Target population coverage and exclusion rates are displayed for each country in Chapter 9. Descriptions of each country's school-level and within-school exclusions can be found in Appendix B.

In any study that utilizes sampling, the population that ultimately participates usually differs slightly from the target population, with some portion of the target population being excluded from the study. A major objective of the TIMSS sampling strategy was to ensure that the effective target population, the population actually sampled by TIMSS, was as close as possible to the international desired population, and to document clearly all excluded populations. Exhibit 5.1 illustrates the relationship between successively more refined definitions of the target population and the excluded populations at each stage.

Exhibit 5.1 Relationship Between the Desired Populations and Exclusions



5.3 Sampling Frames and Survey Units

Once the survey populations were defined, the next step involved building the sampling frames in which all sampling units (grade 4 and/or grade 8 students) within the national defined target populations have a known probability of being sampled. In TIMSS 2007, however, it is important to note that in addition to gathering data on sampled students, a large amount of information also was gathered about their classes and schools, which required other types of sampling units. The intrinsic, hierarchical nature of these nested units necessitated the creation of a sampling frame by stages. Therefore, a two-stage stratified cluster sample design was used, with schools as the first stage and intact classes as the second stage. Because of its large population sizes, it was necessary to include a preliminary sampling stage in the Russian Federation, where regions were sampled first and then schools. Singapore also had a third sampling stage, where students were sampled within classes.

5.3.1 First Stage Sampling Units: Schools

In order to draw school samples that are representative of the student populations, NRCs were asked to provide vital information about all schools (or schools and regions in the Russian Federation) where fourth and eighth grade students could be tested. The following data were required for each school:

- *Measure of size* (MOS): for example, the student enrollment in the target grade, the average student enrollment per grade, the number of classrooms in the target grade, or the total student enrollment in the school.
- *Minimum cluster size* (MCS): the expected number of sampled students per class was required if the number of classrooms in the target grade couldn't be provided. This was calculated as the ratio of the total number of students to the total number of classes for schools having more than one class in the target grade.
- *Variables*: any variables describing school characteristics used for stratification purposes, such as type of school, degree of urbanization codes, or sex of students served by the school.
- *The school sampling probability and status*: information on whether or not that school already was sampled for a study other than TIMSS when overlapping control was required between TIMSS 2007 and other international studies.

In the Russian Federation, a MOS of the regions (preliminary sampling stage) also was required.

5.3.2 Second Stage Sampling Units: Classes

Given the nested nature of the sampling units in TIMSS, listing all classes (along with the class sizes) within sampled schools that agreed to participate in the study was the only requirement for building the class sampling frame. This list included all regular classes, as well as any types of special education classes. Note that within sampled classes, all students were listed. All TIMSS 2007 participating countries had classes as their last stage sampling units except for Singapore, where in addition to classes, students within classes also were sampled through a third sampling stage.

5.4 Sample Selection Method

The student sampling selection method used in TIMSS 2007 is a classic approach that can be found in most sampling textbooks (e.g., Cochran, 1977). The method usually is referred to as a systematic, two-stage *probability proportional-to-size* (PPS) sampling technique, where schools are first sampled and then classes within sampled (and participating) schools. This sampling method is a natural match with the hierarchical nature of the sampling units described above, with classes of students nested within schools. Stratification at the school level was used to complete this technique. Even if a country had a list from which students could be selected directly, this sampling technique, where schools are first sampled and then classes within sampled (and participating) schools, was used for all TIMSS 2007 countries. The only exceptions to this rule were the Russian Federation and Singapore, as mentioned above, which had a three-stage sampling design.

5.4.1 School Stratification

School stratification is the grouping of schools into smaller sampling frames according to information found on the initial sampling frame prior to sampling and may be employed to improve the efficiency of the sample design, to sample sections of the population at different rates, or to ensure adequate representation of specific groups in the sample. School stratification by itself can take two forms: explicit or implicit.

Explicit stratification physically creates smaller sampling frames from which samples of schools and classes ultimately will be drawn. In TIMSS, this type of stratification is used when the usual proportional allocation (i.e., students in certain regions or types of schools are represented in the sample in proportion to their distribution in the population) may not result in adequate representation of some groups of interest in the sample. For example, if a country wanted to make generalizations regarding the science achievement of private sector students, the sampling frame could be split into two strata—public and private sector schools. The sample of schools then could be allocated between the two strata to achieve the desired level of precision in each. In most countries in TIMSS 2007, the school sample allocation among strata was proportional to the number of students found in each stratum. However, it should be noted that even without any stratification, the TIMSS samples represented the different groups found in the population, on average.

Implicit stratification only requires that the school sampling frame be sorted according to some variable(s) prior to sampling and can be nested within explicit stratification. By combining the sorting of the frame with the TIMSS 2007 sampling technique, it is possible to get a sample where students (not schools) are in the same proportions as those found at the population level. When schools from the same implicit stratum tend to have similar behavior, in terms of mathematics and science achievement, implicit stratification also will produce more reliable estimates.

In the basic TIMSS 2007 sample design, all schools in the sampling frame for a country were sorted according to some MOS (see section 5.3.1). If implicit stratification was used, then the sorting by MOS was done within each stratum using a serpentine approach—high to low for the first stratum, followed by low to high for the next, etc. (see the example in Exhibit 5.2).

Exhibit 5.2 MOS Sort Order for Implicit Strata[1]

Implicit Stratum	Sort Order of MOS
1. Rural–Public	High to Low
2. Rural–Private	Low to High
3. Urban–Public	High to Low
4. Urban–Private	Low to High

This way of sorting sampling schools optimizes the chance of choosing a replacement school (see the next section), with a MOS close to that of the originally sampled school it is meant to replace.

5.4.2 Sampling Schools

Schools were sampled using systematic, random sampling with probability proportional to their measures of size. For example, if school A had a MOS value that was twice as large as school B, then School A had twice the chance of being in the sample compared to school B. In the Russian Federation, regions and then schools within sampled regions were sampled following this approach.

To implement the school sampling, schools in each explicit stratum were sorted in order by the implicit stratification variables and within these by the MOS. The measures of size are accumulated from school to school, and a running total, the cumulative measure of size, is recorded next to each school. The cumulative MOS is an indicator of the size of the population of students. Dividing the cumulative MOS by the number of schools to sample gives the sampling interval. In the Russian Federation, the same approach was used to implement the sample of regions. However, no stratification variable was used at the region level.

In order to avoid school sample overlap between TIMSS and another international study (e.g., PISA), where the other study had their sample of schools sampled first, it was necessary to modify the TIMSS school MOS

1 Please refer to the *TIMSS 2007 School Sampling Manual* (TIMSS & PIRLS International Study Center, 2005).

prior to sampling. The technique used for TIMSS is explained below and produced a quasi-PPS school sampling approach. It is a variant of the method originally proposed by Kish and Scott (1971).

Let P_{i1} be the probability of selection of the i^{th} school in sample 1 (already selected prior to TIMSS sampling), and let P_{i2} be its desired PPS probability of selection in TIMSS based on its TIMSS MOS. The i^{th} school in TIMSS with probability P_{i2}' was then selected as follows:

If the i^{th} school was already sampled for the other study,

$$P_{i2}' = Max\left[0, \left(P_{i1} + P_{i2} + 1\right) / P_{i1}\right]$$

If the i^{th} school was not already sampled for the other study,

$$P_{i2}' = Min\left[1, P_{i2} / \left(1 - P_{i1}\right)\right]$$

It is possible to show that over all possible samples, the unconditional probability of selection of the i^{th} school in TIMSS 2007 is P_{i2}. Furthermore, if all of the P_{i1} and P_{i2} are less than 0.5, no school can be sampled twice. However under this approach, the sum over all P_{i2} for a previous given sample (sample 1) is slightly different than n, the desired school sample size for TIMSS. This means that under this approach, there is no control of the sample size even if it is known that it will be n, on average. To get around this problem, an adjustment was done to the P_{i2}' to make them summed to the desired school sample size. This adjustment is given by the following:

$$\tilde{P}_{i2} = \frac{n}{\left(\sum_i P_{i2}'\right)} P_{i2}'$$

With this adjustment, it then was possible to derive a temporary MOS (given by $MOS_i' = \sum MOS_i \bullet \tilde{P}_{i2}' / n$) for each school and use a PPS sampling technique to select the TIMSS sample of schools. Although under this approach, the unconditional probability of the selection of the i^{th} school is not exactly P_{i2}, it is P_{i2} that was used to derive the student weights for TIMSS 2007 (see Chapter 9).

There were three countries and one benchmarking participant that requested control sampling overlap between studies. These are England, the Netherlands, Scotland, and the Canadian province of Alberta.

With systematic PPS sampling, it is possible for a large sampling unit to be selected more than once if its size is greater than the sampling interval. To avoid this situation, all such units were automatically selected by changing each one's MOS to the sampling interval of the associated explicit stratum.

Some schools have so few students that their selection using probability proportional to their size (MOS) becomes problematic. Since the selection of these schools depends on their size, a difference between the number of expected students when drawing the sample and the number of students actually found in the field can contribute substantially to the sampling error. To lessen the impact of this eventuality, any schools with fewer expected students than the average minimum cluster size (MCS) for the explicit stratum were sampled with equal probabilities. For example, if the MCS was 30 students and there were 28 schools with less than 30 students for a total of 476 students, the MOS of these small schools was changed to 476/28 = 17. By doing this, the overall size of the explicit stratum stayed the same, but all small schools had an equal chance of being selected.

The MCS also was used to define very small schools. Whenever a school had an expected number of students less than one quarter of the average MCS, the school was labeled as a very small school. These schools could be excluded, as long as they did not exceed 2 percent of the national desired target population and the overall exclusion rate did not exceed 5 percent.

5.4.3 Replacement Schools

Ideally, response rates always should be 100 percent, and although TIMSS 2007 participants worked hard to achieve this goal, it was anticipated that a 100 percent participation rate would not be possible in all countries. To avoid sample size losses, the TIMSS sampling plan identified, *a priori*, replacement schools for each sampled school. Therefore, if an originally selected school refused to participate in the study, it was possible to replace it with a school that already was identified prior to school sampling. Each originally selected school had up to two pre-assigned replacement schools. In general, the school immediately following the originally selected school on the ordered school sampling frame and the one immediately preceding it were designated as replacement schools. Replacement schools always belong to the same explicit stratum, although they could come from different implicit strata if the originally selected school was either the first or last school of an implicit stratum.

The main objective for having replacement schools in TIMSS 2007 was to ensure adequate sample sizes for analysis of subpopulation differences.

Although the use of replacement schools did not eliminate the risk of bias due to nonresponse, employing implicit stratification and ordering the school sampling frame by size increased the chances that any sampled school's replacements would have similar characteristics. This approach maintains the desired sample size while restricting replacement schools to strata where nonresponse occurred. Since the school frame is ordered by school size, replacement schools also tended to be of the same size as the school they meant to replace. For the field test, replacement schools were used to make sure sample sizes were large enough to validate new items, and no more than one replacement school was assigned per originally selected school.

5.4.4 Sampling Classes

For all participants in TIMSS 2007 except Singapore,[2] intact student classes were the second and final sampling stage, with no student subsampling. This means that all students within sampled classes participated in TIMSS 2007, with the exception of excluded students and students absent the day of the assessment. Classes were selected with equal probability of selection using systematic random sampling. Within each sampled school, all classes of the target grade were listed, and one or more classes were sampled using a random start (different in each sampled school). This method, combined with the PPS sampling method for schools, results in a self-weighting student sample under the following conditions: a) there is a perfect correlation between the school MOS reported in the sampling frame and the actual school size, b) the same number of classes is selected in each school, and c) the MCS is the same for all schools. Given that these conditions were never totally met, student sampling weights varied somewhat from school to school (see Chapter 9 for details about sampling weights).

Within sampled schools, some classes have so few students that it is unreasonable to go through the sampling process and end up with these small classes. Furthermore, small classes tend to increase the risk of unreliable survey estimates. To avoid these problems, a class smaller than half the specified MCS was combined with another class from the same school prior to class sampling.

2 Two classes per school were selected using PPS sampling in Singapore, and 19 students were sampled within each class.

5.5 Sampling Precision and Sample Size

Because TIMSS is fundamentally a study of mathematics and science achievement among fourth and eighth grade students, the precision of survey estimates of student achievement and characteristics was of primary importance. However, TIMSS also reports extensively on school, teacher, and classroom characteristics, so it is necessary to have sufficiently large samples of schools and classes. The TIMSS standards for sampling precision require that all student samples have an effective sample size of at least 400 students for the main criterion variable, which is mathematics and science achievement. In other words, all student samples should yield sampling errors that are no greater than would be obtained from a simple random sample of 400 students.

Given that sampling error, when using simple random sampling, can be expressed as $SE_{SRS} = S / \sqrt{n}$ where *S* gives the population standard deviation and *n* the sample size, a simple random sample of 400 students would yield a 95 percent confidence interval for an estimate of a student-level mean of ±10 percent of its standard deviation ($1.96 \bullet S / \sqrt{400}$). Because the TIMSS achievement scale has a standard deviation of 100 points, this translates into a ±10 points confidence limit (or a standard error estimate of approximately 5 points). Similarly, sample estimates of student-level percentages would have a confidence interval of approximately ±5 percentage points.

Notwithstanding these precision requirements, TIMSS required that all student sample sizes should not be less than 4,000 students. This was necessary to ensure adequate sample sizes for analyses where the student population was broken down into many subgroups. For countries involved in the previous TIMSS cycle in 2003, this minimum student sample size was set to 5,150 students in order to compensate for participaton in the TIMSS 2007 Bridging Study. Furthermore, since TIMSS planned to conduct analyses at the school and classroom level in addition to the student level, all school sample sizes were required to be not less than 150 schools, unless a complete census failed to reach this minimum. Under simple random sampling assumptions, a sample of 150 schools yields a 95 percent confidence interval for an estimate of a school-level mean that is ±16 percent of a standard deviation.

Although the TIMSS sampling precision requirements are such that they would be satisfied by a simple random sample of 400 students, sample designs such as the TIMSS 2007 school-and-class design, typically require much larger student samples to achieve the same level of precision. Because students in the same school and even more so in the same class, tend to be more like each other than like other students in the population, sampling a single class of 30 students will yield less information per student than a random sample of students drawn from across all students in the population. TIMSS uses the intraclass correlation, a statistic indicating how much students in a group are similar on an outcome measure, and a related measure known as the design effect to adjust for this "clustering" effect in planning sample sizes.

For countries taking part in TIMSS for the first time in 2007, the following mathematical formulas were used to estimate how many schools should be sampled to achieve an acceptable level of sampling precision:

$$Var_{PPS} = Deff \bullet Var_{SRS} = \frac{Deff \bullet S^2}{n} \cong \frac{\left[1+\rho(mcs-1)\right] \bullet S^2}{n} \cong \frac{\left[1+\rho(mcs-1)\right] \bullet S^2}{a \bullet mcs}$$

where $Deff$ is a compensation factor for using a sample selection method that differs from a simple random sample (also called design effect), S^2 gives the variance of the population, ρ measures the intraclass correlation between clusters, mcs corresponds to the average number of sampled students per class, and a gives the number of schools to sample. Incorporating the precision requirements described earlier into this equation, which translates into $Var_{PPS} = (0.05)^2 \bullet S^2$, gives the number of schools required as:

(1) $$a = 400 \bullet \frac{\left[1+\rho(mcs-1)\right]}{mcs}$$

For planning purposes, the intraclass correlation coefficient usually was set to 0.3 if no other information was available. For example, with a mcs of 20 students and a ρ of 0.3, equation (1) gives 134 schools.

Equation (1) is a model for determining how many schools were required for the TIMSS 2007 sample under the assumption that the standard error of the criterion variable (student mathematics and science achievement) reflects only sampling variance—the usual situation in sample surveys. However, because of its complex matrix-sampling assessment design, standard errors in TIMSS include an imputation error component in addition to the usual sampling error component (see Chapter 11). To keep the standard error within the prescribed precision limits, the number of schools determined by equation (1) has to be increased, as shown in equation (2):

(2) $$a_{irt} = (400 \bullet 0.5) / mcs$$

Continuing the example for a country with a MCS of 20 students, according to this equation (2), 10 schools would have been added to the 134 schools from equation (1), for a total of 144 schools.

For TIMSS 2007 countries that also had participated in TIMSS 2003, the standard errors computed from the 2003 data were reviewed to ensure that the student samples had been large enough to meet the precision requirements in 2003 and would be sufficiently precise to measure trends to 2007. For the several countries falling somewhat short of the sampling requirements not met in 2003, the school sample size for 2007 was increased using the relation that under similar sampling designs, sampling error is inversely proportional to the square root of the sample size. For example, if the sample size in 2003 yielded a standard error of 7 points for an estimate of a mean, the sample size in 2007 was increased by a factor of 2 to provide a standard error of 5 points ($(7/5)^2 \cong 2$). Intraclass correlation coefficients also were calculated for countries that participated in TIMSS 2003. These coefficients were provided in the *TIMSS 2007 School Sampling Manual* (TIMSS & PIRLS International Study Center, 2005).

5.6 Selecting Field-test Samples

Prior to the main data collection, which was conducted from October–November 2006 in Southern Hemisphere countries and from April–May 2007 in Northern Hemisphere countries, TIMSS 2007 conducted a full-scale field test in April 2006 in all participating countries. The field test sample size was approximately 30 schools in each country. Countries were required to draw their field test samples using the same random sampling procedures that they employed for the main samples. This ensured that field test samples closely approximated the main samples. In an attempt to reduce the burden on schools, the field test and main data collection samples of schools were drawn simultaneously, so that a school could be selected for either the field test or the main data collection, but not both. For example, if 150 schools were needed for the main data collection and another 30 schools were needed for the field test, a larger sample of 180 schools was selected using the sampling method described earlier. A systematic subsample of 30 schools then was selected from the 180 schools and assigned to the field test, leaving 150 schools for data collection.[3]

3 In countries where it was necessary to conduct a census of all schools or where the NRC believed that the sampling frame used to draw the combined sample was not appropriate for the data collection, separate sampling frames were provided for the field test and main data collection. In such situations, no attempt was made to minimize the overlap.

References

Cochran, W.G. (1977). Sampling techniques, New York: John Wiley.

IEA. (2006). *Windows Within-school sampling software (WinW3S)* [Computer software and manual]. Hamburg: IEA Data Processing and Research Center.

Kish, L. & Scott, A. (1971). Retaining units after changing strata and probabilities. *Journal of the American Statistical Association,* 66, 461–470.

Martin, M.O., Mullis, I.V.S., & Foy, P. (with Olson, J.F., Erberber, E., Preuschoff, C., & Galia, J.). (2008). *TIMSS 2007 international science report: Findings from IEA's trends in international mathematics and science study at the fourth and eighth grades*. Chestnut Hill, MA: TIMSS & PIRLS International Study Center, Boston College.

Mullis, I.V.S., Martin, M.O., & Foy, P. (with Olson, J.F., Preuschoff, C., Erberber, E., Arora, A., & Galia, J.). (2008). *TIMSS 2007 international mathematics report: Findings from IEA's Trends in International Mathematics and Science Study at the fourth and eighth grades*. Chestnut Hill, MA: TIMSS & PIRLS International Study Center, Boston College.

TIMSS & PIRLS International Study Center. (2005). *TIMSS 2007 school sampling manual.* Chestnut Hill, MA: TIMSS & PIRLS International Study Center, Boston College.

TIMSS & PIRLS International Study Center. (2006a). *TIMSS 2007 survey operations procedures unit 2: Contacting schools and sampling classes for TIMSS 2007 assessment.* Chestnut Hill, MA: TIMSS & PIRLS International Study Center, Boston College.

TIMSS & PIRLS International Study Center. (2006b). *TIMSS 2007 school coordinator manual.* Chestnut Hill, MA: TIMSS & PIRLS International Study Center, Boston College.

UNESCO Institute for Statistics. (1999). *Operational manual for ISCED-1997 (international standard classification of education)*. Paris: Author.