CHAPTER 3

Sample Design in TIMSS 2019

Sylvie LaRoche
Marc Joncas
Pierre Foy

Introduction

TIMSS is designed to provide valid and reliable measurement of trends in student achievement in countries around the world, while keeping to a minimum the burden on schools, teachers, and students. The TIMSS program employs rigorous school and classroom sampling techniques so that achievement in the student population as a whole may be estimated accurately by assessing just a sample of students from a sample of schools. TIMSS assesses mathematics and science achievement at two grade levels and so TIMSS has two target populations—all students enrolled at the fourth grade and all students enrolled at the eighth grade, counting from the first year of primary schooling. Countries may assess either one or both student populations. In addition, at the fourth grade for the TIMSS 2019 cycle, countries for which the regular fourth grade mathematics assessment is too difficult have the option to administer a less difficult mathematics assessment, consisting of one third of the items from the regular assessment and two-thirds less difficult items. Countries availing of the less difficult mathematics option administer the regular fourth grade science assessment.

TIMSS 2019 marks the beginning of the TIMSS transition to computer based assessment, with countries having the option of administering the new computer-based version of the 2019 assessment, known as eTIMSS, or the paper-and-pencil version as in previous assessment cycles (paperTIMSS). Although the two versions were developed to be as similar in content as possible, inevitably there are some differences between them as a result of the two modes of administration. In order to control for mode effects while linking the two versions to the TIMSS achievement scales and to safeguard the measurement of trends from previous assessments, eTIMSS countries also provide a separate sample of "bridge" data. The bridge data result from administering the paper version of the trend items (eight blocks of items for each subject and grade that also were administered in 2015) to a separate, equivalent sample of students during the main data collection. These paper versions of the trend items are identical in most respects

to the eTIMSS versions that are administered as part of the main eTIMSS assessment, and so comparing performance on the eTIMSS versions to performance on the paper versions administered to the bridge sample provides a bridge between the two assessment modes.

The TIMSS assessments employ a two-stage random sample design, with a sample of schools drawn as a first stage and one or more intact classes of students selected from each of the sampled schools as a second stage. Intact classes of students are sampled rather than individuals from across the grade level or of a certain age because TIMSS pays particular attention to students' curricular and instructional experiences, and these typically are organized on a classroom basis. Sampling intact classes also has the operational advantage of less disruption to the school's day-to-day business than individual student sampling.

National Sampling Plan

Each country participating in TIMSS needs a plan for defining its national target population and applying the TIMSS sampling methods to achieve a nationally representative sample of schools and students. The development and implementation of the national sampling plan is a collaborative exercise involving the country's National Research Coordinator (NRC) and TIMSS sampling experts.

Statistics Canada is responsible for advising the National Research Coordinator on all sampling matters and for ensuring that the national sampling plan conforms to the TIMSS standards. In cooperation with sampling staff from IEA Hamburg, Statistics Canada works with the NRC to select the national school sample(s) and produce all supporting documentation for tracking the sampled schools. This includes ensuring that the school sampling frame (the school population list from which the school sample is drawn) provided by the NRC is complete and satisfactory; checking that categories of excluded students are clearly defined, justified, and kept to a minimum; assisting the NRC in determining the sample size and a stratification plan that will meet both international and national objectives; and drawing a national sample of schools. When sampling has been completed and all data collected, Statistics Canada documents population coverage and school and student participation rates and constructs appropriate sampling weights for use in analyzing and reporting the results.

The TIMSS & PIRLS International Study Center, in cooperation with Statistics Canada and IEA Hamburg, provides National Research Coordinators with a series of manuals to guide them through the sampling process. More specifically, *TIMSS 2019 Survey Operations Procedures Unit 1: Sampling Schools and Obtaining their Cooperation* describes the steps involved in defining the national target population and selecting the school sample, and *TIMSS 2019 Survey Operations Procedures Unit 3: Contacting Schools and Sampling Classes for the TIMSS 2019 Data Collection* describes the procedure for sampling classes within the sampled schools and making preparations for conducting the assessments. Within-school sampling procedures for the field test are documented in *TIMSS 2019 Survey Operations Procedures Unit 2: Preparing for and Conducting the TIMSS 2019 Field Test*. More information on the Survey Operations Units can be found in Chapter 6 of this volume.

The TIMSS National Research Coordinator is responsible for providing Statistics Canada with all information and documentation necessary to conduct the national sampling, and for conducting all sampling operations in the country. In particular, the NRC is expected to identify the grade(s) that correspond to the international target population(s); create a sampling frame by listing all schools in the population that have classes with students in the target grade(s); determine national population coverage and exclusions, in accordance with the TIMSS international guidelines; work with Statistics Canada to develop a national sampling plan and identify suitable stratification variables, ensuring that these variables are present and correct for all schools; contact all sampled schools and secure their participation; keep track of school participation and the use of replacement schools; and conduct all within-school sampling of classes. As described in this chapter, each NRC is required to complete a series of sampling forms documenting the completion of each of these tasks.

A crucial feature of each international meeting of National Research Coordinators is a one-to-one meeting between each NRC and sampling staff at Statistics Canada and IEA Hamburg. At these meetings, each step of the sampling process is documented and reviewed in detail, and NRCs have the opportunity to raise issues and ask questions about their national situation and any challenges they face. Statistics Canada consults with the TIMSS & PIRLS International Study Center and the International Sampling Referee, as necessary, to resolve issues and questions. Final approval of TIMSS national sampling plans is the responsibility of the TIMSS & PIRLS International Study Center, based upon the advice of Statistics Canada and the International Sampling Referee.

Defining the Target Population

As an international study of the comparative effects of education on student achievement in mathematics and science, TIMSS defines its international target populations in terms of the amount of schooling students have received. The number of years of formal schooling is the basis of comparison among participating countries. Thus, the TIMSS international target population at the fourth grade is all students in their fourth year of formal schooling, and at the eighth grade, all students in their eighth year. UNESCO's International Standard Classification of Education (ISCED) 2011 (UNESCO, 2012) provides an internationally accepted classification scheme for describing levels of schooling across countries. The ISCED system describes the full range of schooling, from preprimary (Level 0) to the doctoral level (Level 8). ISCED Level 1 corresponds to primary education or the first stage of basic education. The first year of Level 1 "coincides with the transition point in an education system where systematic teaching and learning in reading, writing and mathematics begins" (UNESCO, 2012, p. 30). Four years after this would be the target grade for fourth grade TIMSS and is the fourth grade in most countries. Similarly, eight years after the first year of ISCED Level 1 is the target grade for eighth grade TIMSS and is the eighth grade in most countries. However, given the cognitive demands of the assessments, TIMSS wants to avoid assessing

very young students. Thus, TIMSS recommends assessing the next higher grade (i.e., fifth grade for fourth grade TIMSS and ninth grade for eighth grade TIMSS) if, for fourth grade students, the average age at the time of testing would be less than 9.5 years and, for eighth grade students, less than 13.5 years.

The fourth grade and eighth grade target populations of students are defined as follows:

- **Fourth grade:** All students enrolled in the grade that represents four years of schooling counting from the first year of ISCED Level 1, providing the mean age at the time of testing is at least 9.5 years
- **Eighth grade:** All students enrolled in the grade that represents eight years of schooling counting from the first year of ISCED Level 1, providing the mean age at the time of testing is at least 13.5 years

All students enrolled in the target grade, regardless of their age, belong to the international target population and should be eligible to participate in TIMSS. Because students are sampled in two stages, first by randomly selecting a school and then randomly selecting a class from within the school, it is necessary to identify all schools in which eligible students are enrolled. Essentially, eligible schools for TIMSS are those that have any students enrolled in the target grade, regardless of type of school. All schools of all educational sub-systems that have students learning full time in the target grade are part of the international target population, including schools that are not under the authority of the national Ministry of Education.

National Target Populations

For most countries, the target grade for TIMSS is the fourth and/or eighth grade. However, because educational systems vary in structure and in policies and practices with regard to age of starting school and promotion and retention, there are differences across countries in how the target grades are labelled and in the average age of students. To ensure that the appropriate national target grades are selected, each NRC completes Sampling Form 1, which identifies the target grades, the country's name for those grades, and the average age of students in those grades at the time of data collection. An example of a completed Sampling Form 1 is presented in Exhibit 3.1.

For a variety of reasons, there are countries where students in the fifth or sixth grade are more likely to have developed the mathematics and science competencies necessary for success on the TIMSS fourth grade assessment, or in the ninth grade for the TIMSS eighth grade assessment. Such countries may choose to participate in TIMSS at either the fifth or sixth grade or in the less difficult mathematics fourth grade assessment. Similarly, some countries may choose to administer the TIMSS eighth grade assessment to their ninth grade students.

Exhibit 3.1: Example of Sampling Form 1

Sampling Form 1 — General Information

See Section 2 of TIMSS 2019 Survey Operations Procedures Unit 1

TIMSS 2019 Participant :	*Country X*
National Research Coordinator :	*Name of NRC*

1. Please indicate the assessment(s) in which your country plans to participate along with the target grade(s), name(s), and expected average age of students at the time of testing:

Grade 4 TIMSS Assessment	*Yes*

Target Grade	Name of the Target Grade	Average Age	Mode (Paper-TIMSS or eTIMSS)	Less Difficult Mathematics Item Blocks Option (Yes/No)
4	*Grade 4*	*9.7*	*eTIMSS*	*No*
			Select	*Select*
			Select	*Select*

Grade 8 TIMSS Assessment	*Yes*

Target Grade	Name of the Target Grade	Average Age	Mode (Paper-TIMSS or eTIMSS)
8	*Grade 8*	*13.7*	*eTIMSS*
			Select
			Select

2. Specify the usual start and end date(s) of the school year and the expected date(s) of testing for the field test and data collection.

	Start of school year: *(DD-MM-YYYY)*	End of school year: *(DD-MM-YYYY)*	Expected Testing Period
Field Test	*05/09/2017*	*22/06/2018*	*16 - 27 April 2018*
Data Collection	*01/09/2018*	*21/06/2019*	*13 - 24 April 2019*

4. Specify the language(s) in which the assessment(s) will be administered.

English

5. Describe the grade structure through ISCED Level 1 (primary education or the first stage of basic education) and ISCED Level 2 (basic or lower secondary education) in your country.

Grades 1 to 6 , Primary schools
Grades 7 to 9 , Lower secondary schools

6. Describe the age and birth date rules for entering ISCED Level 1 in your country.

Children must enter school (grade 1) in the autumn of the year in which they have their sixth birthday

National Coverage and Exclusions

TIMSS is designed to describe and summarize student achievement across the entire target grade (fourth or eighth), and so it is very important that national target populations aim for comprehensive coverage of eligible students. However, in some cases, political, organizational, or operational factors make complete national coverage difficult to attain. Thus, in some rare situations, certain groups of schools and students may have to be excluded from the national target population. For example, it may be that a particular geographical region, educational sub-system, or language group cannot be covered. Such exclusion of schools and students from the target population is referred to as reduced population coverage.

Even countries with complete population coverage find it necessary to exclude at least some students from the target population because they attend very small schools, have intellectual or functional disabilities, or are non-native language speakers. Such students may be excluded at the school level (i.e., the whole school is excluded) or within the school on an individual basis.

School-Level Exclusions. Although it is expected that very few schools will be excluded from the national target population, NRCs are permitted to exclude schools on the following grounds when they consider it necessary:

- Inaccessibility due to their geographically remote location
- Extremely small size (e.g., four or fewer students in the target grade)
- Offering a grade structure, or curriculum, radically different from the mainstream educational system
- Providing instruction solely to students in the student-level exclusion categories listed below (e.g., catering only to special needs students)

Student-Level Exclusions. The international within-school exclusion rules are specified as follows:

- Students with functional disabilities — These are students who have physical disabilities such that they cannot perform in the TIMSS testing situation. Students with functional disabilities who are able to perform should be included in the testing.
- 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 have intellectual disabilities or who have been tested 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. It should be noted that students with dyslexia, or other such learning disabilities, should be accommodated in the test situation if possible, rather than excluded.

- 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 in the test situation. Typically, a student who has received less than one year of instruction in the language(s) of the test should be excluded.

Because disability criteria vary from country to country, NRCs are asked to translate the TIMSS international exclusion standards into the local equivalent. Students should be considered for exclusion strictly in accordance with the international standards. If a sampled school contains a class consisting entirely of students from one of the exclusion categories, such a class is excluded prior to classroom sampling.

NRCs understand that exclusion rates must be kept to a minimum so that national samples accurately represent the national target population. Requirements for exclusion rates include the following:

- The overall number of excluded students must not account for more than 5 percent of the national target population of students in a country. The overall number includes both school-level and within-school exclusions.
- The number of students excluded because they attend very small schools must not account for more than 2 percent of the national target population of students.

To document population coverage and exclusions, each NRC completes Sampling Form 2, which lists the number of students in the national target population and the number of students excluded at both the school level and within the school for each population to be assessed. An example of a completed Sampling Form 2 is presented in Exhibit 3.2.

Exhibit 3.2: Example of Sampling Form 2

Sampling Form 2 — Coverage and Exclusions

See Section 3 of TIMSS 2019 Survey Operations Procedures Unit 1

TIMSS 2019 Participant :	*Country X*

1. This Sampling Form refers to: *TIMSS Grade 4 Assessment*

		Number of schools	Number of students
Total enrollment in the target grade:	[a]	*822*	*56,560*

2. School-level exclusions *(if applicable):*

	Description of exclusions	Number of schools	Number of students
1.	*Students taught in language other than English*	*8*	*630*
2.	*Special education schools*	*16*	*325*
3.	*Very small schools (less than 5 students in grade 4)*	*40*	*110*
4.			
5.			
TOTAL:	*(Sum of exclusions - Calculated automatically)* [b]	*64*	*1,065*

		schools	students
Percentage of school-level exclusions: *(Box [b] ÷ Box [a] x 100)*	**[1]**	***7.8%***	***1.9%***

3. Total enrollment after school-level exclusions: *(Box [c] = Box [a] - Box [b])* [c] *758* *55,495*

Totals and percentages calculated automatically

4. Within-school exclusions *(if applicable):*

	Description of exclusions	Number of students
1.	*Students with special education needs (based on TIMSS 2015)*	*640*
2.		
3.		
TOTAL:	*(Sum of exclusions - Calculated automatically)* [d]	*640*

		schools	students
Expected percentage of within-school exclusions: *(Box [d] ÷ Box [c] x 100)*	**[2]**	***0.0%***	***1.2%***

5. **Expected percentage of reduced coverage and exclusions:** *(Box [1] + (1 - Box [1]) X Box [2])* ***7.8%*** ***3.0%***

Totals and percentages calculated automatically

6. Total enrollment in the target grade in previous school years.

Years	Number of schools	Number of students
2016/2017	*856*	*58,451*
2015/2016	*890*	*61,489*

Requirements for Sampling the Target Population

TIMSS sets high standards for sampling precision, participation rates, and sample implementation in order to achieve national samples of the highest quality and survey estimates that are unbiased, accurate and internationally comparable.

Sampling Precision and Sample Size

Because TIMSS is fundamentally a study of student achievement, the precision of estimates of student achievement is of primary importance. To meet the TIMSS standards for sampling precision, national student samples should provide for a standard error no greater than .035 standard deviation units for the country's mean achievement. This standard error corresponds to a 95% confidence interval of ±7 score points for the achievement mean and of ±10 score points for the difference between achievement means from successive cycles (e.g., the difference between a country's achievement mean on TIMSS 2015 and TIMSS 2019).[1] Sample estimates of any student-level percentage estimate (e.g., a student background characteristic) should have a confidence interval of ±3.5%.

For most countries, the TIMSS precision requirements are met with a school sample of 150 schools and a student sample of 4,000 students for each target grade. Depending on the average class size in the country, one class from each sampled school may be sufficient to achieve the desired student sample size. For example, if the average class size in a country were 27 students, a single class from each of 150 schools would provide a sample of 4,050 students (assuming full participation by schools and students). Some countries choose to sample more than one class per school, either to increase the size of the student sample or to provide a better estimate of school-level effects.

Countries transitioning to eTIMSS require an additional sample of at least 1,500 tested students for the bridge data collection. This bridge sample is obtained by selecting one additional class from a subset of the sampled schools, by selecting a distinct sample of schools, or by a combination of both strategies. The most suitable approach is developed with the sampling experts from Statistics Canada during the sampling development stage.

A school sample larger than the minimum of 150 schools may be required under the following circumstances:

- The average class size in a country is so small that, even when sampling more than one classroom per school, it is not possible to reach the student sample size requirements by selecting only 150 schools.
- Previous cycles of TIMSS showed that the sampling precision requirements cannot be met unless a larger school sample is selected.

1 The TIMSS achievement scales were established in 1995 based on the combined achievement distribution of all countries that participated in TIMSS 1995, at each grade level. To provide a point of reference for country comparisons, the scale centerpoint of 500 was located at the mean of the combined achievement distribution. The scale units were chosen so that 100 scale score points corresponded to the standard deviation of the distribution. Accordingly, one standard deviation unit is approximately 100 scale score points.

- Classes within schools are tracked by student performance (more common at eighth grade than at fourth grade). This increases variation between classes in student achievement and can reduce sampling precision. In this situation, it is advisable to sample at least two classrooms per school whenever possible, in addition to sampling more schools.
- A high level of non-response is anticipated, leading to sample attrition and reduced sample size. Note that while a larger school sample helps to maintain sample size in the face of non-response, it does not compensate for non-response bias.

Field Test Sample

Although the TIMSS field test is scheduled in the school year before the year of data collection, the school sample for the field test is drawn at the same time and from the same population of schools as the full sample. The field test sample size requirement is 200 students per field test achievement booklet, and so the total field test sample size is a function of the number of achievement booklets being field tested. For TIMSS 2019, the paperTIMSS field test has five booklets per target grade and so requires a field test sample of 1,000 students at each grade. The eTIMSS field test has five item block combinations (corresponding to the five paperTIMSS booklets) requiring 1,000 students at each grade, and also three block combinations of Problem Solving and Inquiry tasks (PSIs). Each PSI block appears in two different item block combinations, so the three block combinations require a further 300 students, for a total field test sample of 1,300 students.

Participation Rates

To minimize the potential for non-response bias, TIMSS aims for 100 percent participation by sampled schools, classrooms, and students, while recognizing that some degree of non-participation may be unavoidable. For a national sample to be fully acceptable it must have either:

- A minimum school participation rate of 85 percent, based on originally sampled schools AND
- A minimum classroom participation rate of 95 percent, from originally sampled schools and replacement schools AND
- A minimum student participation rate of 85 percent, from sampled schools and replacement schools

  OR
- A minimum combined school, classroom, and student participation rate of 75 percent, based on originally sampled schools (although classroom and student participation rates may include replacement schools)

  Classrooms with less than 50 percent student participation are deemed to be not participating.

Developing and Implementing the National Sampling Plan

Although National Research Coordinators are responsible for developing and implementing national sampling plans, Statistics Canada and the IEA Sampling team work closely with NRCs to help ensure that these sampling plans fully meet the standards set by the TIMSS & PIRLS International Study Center, while also adapting to national circumstances and requirements. National sampling plans must be based on the international two-stage sample design (schools as the first stage and classes within schools as the second stage) and must be approved by Statistics Canada.

TIMSS Stratified Two-Stage Cluster Sample Design

The basic international sample design for TIMSS is a stratified two-stage cluster sample design, as follows:

First Sampling Stage. For the first sampling stage, schools are sampled with probability proportional to their size (PPS) from the list of all schools in the population that contain eligible students. The schools in this list (or sampling frame) may be stratified (sorted) according to important demographic variables. Schools for the field test and data collection are sampled simultaneously using a systematic random sampling approach. Two replacement schools are also pre-assigned to each sampled school during the sample selection process, and these replacement schools are held in reserve in case the originally sampled school refuses to participate. Replacement schools are used solely to compensate for sample size losses in the event that the originally sampled school does not participate. School sampling is conducted for each country by Statistics Canada with assistance from the IEA Sampling Team, using the sampling frame provided by the country's National Research Coordinator.

Second Sampling Stage. The second sampling stage consists of selecting one (or more) intact class from the target grade of each participating school. Class sampling in each country is conducted by the National Research Coordinator using the Within-School Sampling Software (WinW3S) developed by IEA Hamburg and Statistics Canada. Having secured a sampled school's agreement to participate in the assessment, the NRC requests information about the number of classes and teachers in the school and enters it in the WinW3S database. Classes smaller than a specified minimum size are combined into pseudo-classes prior to sampling. The software samples one or more classes with equal probability in each school. All students in each sampled class participate in the assessment. Sampled classes that refuse to participate may not be replaced.

An additional sampling step is required for eTIMSS countries that require a bridge sample. Students in the bridge sample are administered a paper version of the trend item blocks, and it is important that this sample should mirror the main eTIMSS sample as closely as possible. For operational reasons it is not possible to administer both the eTIMSS assessment and bridge assessment in the same class, so the bridge sample should consist of an extra class from a school sampled for eTIMSS or from an additional school. In schools selected for both the eTIMSS and the bridge samples, separate classes are sampled and randomly assigned to either the eTIMSS or bridge samples using the WinW3S software.

Stratification

Stratification consists of arranging the schools in the target population into groups, or strata, that share common characteristics such as geographic region or school type. Examples of stratification variables used in TIMSS include region of the country (e.g., states or provinces); school type or source of funding (e.g., public or private); language of instruction; level of urbanization (e.g., urban or rural area); socioeconomic indicators; and school performance on national examinations.

In TIMSS, stratification is used to:

- Improve the efficiency of the sample design, thereby making survey estimates more reliable
- Apply different sample designs, such as disproportionate sample allocations, to specific groups of schools (e.g., those in certain states or provinces)
- Ensure proportional representation of specific groups of schools in the sample

School stratification can take two forms: explicit and implicit. In explicit stratification, a separate school list or sampling frame is constructed for each stratum and a sample of schools is drawn from that stratum. In TIMSS, the major reason for considering explicit stratification is disproportionate allocation of the school sample across strata. For example, in order to produce equally reliable estimates for each geographic region in a country, explicit stratification by region may be used to ensure the same number of schools in the sample for each region, regardless of the relative population size of the regions.

Implicit stratification consists of sorting the schools by one or more stratification variables within each explicit stratum, or within the entire sampling frame if explicit stratification is not used. The combined use of implicit strata and systematic sampling is a simple and effective way of ensuring a proportional sample allocation of students across all implicit strata. Implicit stratification also can lead to improved reliability of achievement estimates when the implicit stratification variables are correlated with student achievement.

National Research Coordinators consult with Statistics Canada and the IEA Sampling team to identify the stratification variables to be included in their sampling plans. The school sampling frame is sorted by the stratification variables prior to sampling schools so that adjacent schools are as similar as possible. Regardless of any other explicit or implicit variables that may be used, the school size is always included as an implicit stratification variable.

To document the stratification variables used in their sampling plans, each NRC completes Sampling Form 3, which lists the variables to be used for explicit and implicit stratification, and the number of levels of each stratification variable. An example of a completed Sampling Form 3 is presented in Exhibit 3.3. Further details on the explicit and implicit stratification variables for each country can be found in the Characteristics of National Samples section in Chapter 9: Sampling Implementation.

Exhibit 3.3: Example of Sampling Form 3

Sampling Form 3	Stratification

See Section 4 of TIMSS 2019 Survey Operations Procedures Unit 1

TIMSS 2019 Participant :	*Country X*

1. This Sampling Form refers to: *TIMSS Grade 4 Assessment*

Stratification of schools

2. List and describe the variables to be used for stratification in order of importance:
   (Please note that the choice of variables used for explicit or implicit stratification will be discussed during consultations with the TIMSS sampling experts

Stratification Variables			
	Name	**Description**	**# of levels**
1	*School type*	*public, private*	*2*
2	*Socioeconomic status*	*high, medium, low*	*3*
3			
4			
5			
6			

Include additional information if necessary:

3. If applicable, describe any additional requirements for sub-national estimates, either for reporting or analysis purposes (e.g., oversampling of specific groups of the population):

would like to have reliable estimates for students from the private schools

School Sampling Frame

One of the National Research Coordinator's most important sampling tasks is the construction of a school sampling frame for the target population. The sampling frame is a list of all schools in the country that have students enrolled in the target grade, and is the list from which the school sample is drawn. A well-constructed sampling frame provides complete coverage of the national target population without being contaminated by incorrect or duplicate entries or entries that refer to elements that are not part of the defined target population.

A suitable school measure of size (MOS) is a critical aspect of the national sampling plan, because the size of a school determines its probability of selection. The most appropriate school measure of size is an up-to-date count of the number of students in the target grade. If the number of students in the target grade is not available, total student enrollment in the school may be the best available substitute.

Sampling Form 4, presented in Exhibit 3.4, provides some basic information about the school sampling frame, including the average class size at the target grade, the number of classrooms to be sampled per school, the school measure of size (MOS) to be used for school sampling, and the school year from which the frame was constructed.

Exhibit 3.4: Example of Sampling Form 4

Sampling Form 4	Classroom Information and Sampling Frame

See Section 5 of TIMSS 2019 Survey Operations Procedures Unit 1

TIMSS 2019 Participant :	*Country X*

1. This Sampling Form refers to: *TIMSS Grade 4 Assessment*

2. Specify the school measure of size (MOS) to be used.

Click in box and on right arrow to see drop down menu

	Name of the MOS variable in the school frame:
1. Number of students in the target grade (preferred)	*GR4_STD*

If "Other," please describe:

3. Specify the average class size (ACS) for the target grade in your schools. *24*

4. Specify how many classrooms you plan to sample per school.
(Click in box and on right arrow to see drop down menu)

2. More than one classroom in tracked schools

If "Other," please describe:

5. Specify the school year for which enrollment data will be used for the school MOS. *2017/2018*

6. If a frame other than a single-level sampling frame (list of all schools) is to be used, please provide a preliminary description of the information available to construct this frame.

Not applicable

The school sampling frame is usually a spreadsheet containing a single entry for each school. This entry includes a unique identification number and contact information (if appropriate given the country's privacy laws), the values of the stratification variables for the school, and the school measure of size. It is useful if the school entry also includes the number of classes in the school in the target grade because this provides a mechanism for predicting in advance the size of the eventual student sample. This predicted sample size may be compared with the eventual student sample size as a check on the sampling process.

Exhibit 3.5 provides an example of a partial sampling frame for a country conducting TIMSS 2019 at the eighth grade. In this example, region and urbanization are used as stratification variables.

Exhibit 3.5: Example of a Partial Sampling Frame

	A	B	C	D	E	F	G	H	I	J
1	School ID	Region	Urbani-zation	Grade 8 Students	Grade 8 Classes	School Name	School Address	Postal code	Town	Tel
2	15104	South	Rural	211	8	Campbell College	Jelly Bean Ave 23	01604	Dinsdale	040 / 5699
3	15113	North	Rural	176	7	Stromboli High School	Barracuda Street 5	01611	Lowrie	040 / 5666
4	15115	North	Rural	182	7	Central Park School	Wales Crescent 45	01600	Kristin	041 / 5599
5	15123	North	Urban	104	4	Obi Wan School	Wheel Crescent 23	01903	Curtain	040 / 5000
6	15933	North	Rural	228	9	Alfred Hitchcock High School	Dennis Street 45	01600	Tortilla Plains	041 / 5566
7	15937	North	Urban	186	7	Begonia High School	Morning Street 125	01614	Peacew	040 / 5644
8	15940	North	Urban	153	6	Calmar High School	Casey Crescent 1	01905	Waltington	040 / 5633
9	15942	North	Urban	169	7	Western High School	Travis Ave 54	01905	Waltington	040 / 5644
10	15944	North	Urban	8	1	Manhattan College	Launcaster Street 63	01614	Peacew	040 / 5577
11	15945	South	Rural	229	9	Karaoke High School	Bean Street 45	01614	Blue Lake	040 / 5700
12	15946	South	Rural	164	7	J. Oliver High Cuisine School	Cambridge Crescent 136	01905	Cinder	049 / 5777
13	15953	South	Urban	89	4	Douglas College	Douglas Drive 78	01619	Hawn	049 / 5762
14	15956	South	Urban	22	1	Emily Dickinson College	Phillip Glass Avenue 23	01619	Hawn	049 / 5645
15	15958	North	Urban	65	3	Tinsdale College	McGyver Crescent 49	01903	Curtain	040 / 5811
16	15968	South	Urban	34	1	Gualajara District High School	Strong Street 79	01615	Flowerburgh	040 / 5612
17	15970	South	Urban	188	8	Dry Creek School	Galloway Street 46	01615	Flowerburgh	040 / 5295
18	15974	South	Rural	6	1	Eagle College	Monday Street 123	01614	Candid	040 / 5774
19	15981	South	Rural	81	3	St John High School	Alec Baldwin Drive 75	01617	Holster	040 / 5511
20	15983	South	Rural	88	4	Kum Ba Yah High School	O'Malley Circuit 56	01901	Book Haven	049 / 5693
21	15984	South	Rural	54	2	La Giocconda College	Dodo Bank 45	01616	Kathleen River	049 / 5709
22	15985	South	Urban	45	2	Lake Titicaca College	Collin Benjamin Street 1	01900	Evans	049 / 5622
23	15986	South	Rural	213	9	Paul Bunyan High School	Heidelberg Street 100	01905	Charpwood	049 / 5767
24	15988	South	Rural	290	12	Lynn High School	Good Street 45	01601	Heintz	049 / 5639
25	[illegible]	[illegible]	[illegible]	128	5	Fruit Tree High School	[illegible]	01615	Karburetta	049 / 5611
	[illegible]	[illegible]	[illegible]	[illegible]	9	E. Cochran [illegible]	[illegible]	[illegible]	Garden Heights	049 / [illegible]

Sampling Schools

Once the school sampling frame is structured to meet all international and national requirements, Statistics Canada can draw the school sample. If the sampling frame is explicitly stratified, it is necessary to decide how the school sample is to be allocated among the explicit strata (i.e., the number of schools to be sampled in each stratum). When this has been decided, a sample of schools is selected within each explicit stratum using systematic sampling with probabilities proportional to size (PPS). The PPS technique means that the larger schools, those with more students, have a higher probability of being sampled than the smaller schools. However, this difference in the selection probabilities of larger and smaller schools is largely offset at the second stage of sampling by selecting a fixed number of classes

(usually one or two) with equal probability from the sampled school. Classes in large schools with many classes at the target grade have a lower probability of selection than classes in smaller schools that have just one or two classes. A description of the school sampling procedure is provided in Appendix 3A.

Even though the field test is scheduled in the school year before the year of data collection in most countries, the preferred approach in TIMSS is to select both samples of schools at the same time. This ensures that both the field test and data collection samples constitute random samples representative of all schools in the country, and that no school is selected for both samples.[2]

Replacement Schools. Ideally, all schools sampled for TIMSS should participate in the assessments, and NRCs work hard to achieve this goal. Nevertheless, it is anticipated that a 100 percent participation rate may not be possible in all countries. To avoid sample size losses, the sampling plan identifies, *a priori*, specific replacement schools for each sampled school. Each originally sampled school has two pre-assigned replacement schools, usually the school immediately preceding the originally sampled school on the school sampling frame and the one immediately following it. Replacement schools always belong to the same explicit stratum as the original but may come from different implicit strata if the school they are replacing is either the first or last school of an implicit stratum.

The main justification for replacement schools in TIMSS is to ensure adequate sample sizes for analysis of subpopulation differences. Although the use of replacement schools does not eliminate the risk of bias due to school nonparticipation, employing implicit stratification and ordering the school sampling frame by school size increases the chances that a sampled school's replacements would have similar characteristics. This approach maintains the desired sample size while restricting replacement schools to strata where nonresponse occurs. Since the school frame is ordered by school size, replacement schools also tend to be similar in size to the school they are designated to replace.

NRCs understand that they should make every effort to secure the participation of all of the sampled schools. Only after all attempts to persuade a sampled school to participate have failed is the use of its replacement school considered.

Common Adjustments to the TIMSS School Sampling Design

The TIMSS school sample design offers considerable flexibility to countries participating at both fourth and eighth grades to maximize or minimize the extent to which the same schools are assessed. Where fourth and eighth grade students attend the same school, some countries find it more efficient to administer TIMSS at the same school for both grades. In other cases, countries try to ensure that assessments are spread across schools and therefore prefer that TIMSS at the fourth and eighth grades are not administered at the same school and/or that TIMSS sampling avoid, when possible, selecting schools that have recently administered other national and international assessments. To provide flexibility to

2 With approval from the TIMSS & PIRLS international Study Center, the field test and full sample could be selected separately. In such cases an overlap control procedure is used to minimize the probability of selecting schools for the data collection that already had been sampled for the field test. This was the case for most eTIMSS countries due to operational constraints.

meet these requests, Statistics Canada implements modified sampling procedures—the details of which are described in Appendix 3B.

Sampling Classes

Within each sampled school, all classes with students at the target grade are listed, and one or more intact classes are selected with equal probability of selection using systematic random sampling. This procedure is implemented using the WinW3S sampling software. The selection of classes with equal probability, combined with the PPS sampling method for schools, in general results in a self-weighting student sample. If the school has multi-grade classes (i.e., the class contains students from more than one grade level), only students from the target grade are eligible for sampling.

When a country participating in eTIMSS has schools selected for both the eTIMSS and the bridging assessments, sampled classes within these schools are randomly assigned to one study or the other. This is done automatically within the WinW3S software.

Because small classes tend to increase the risk of unreliable survey estimates and can lead to reduced overall student sample size, it is necessary to avoid sampling too many small classes. Based on consideration of the size distribution of classes and the average class size, a lower class size limit or minimum class size (MCS) is specified for each country. Prior to sampling classes in a school, any class smaller than the MCS is combined with another class in the school to form a pseudo-class for sampling purposes. The procedure for sampling classes within schools is described in more detail in the Survey Operations Procedures chapter of this volume.

Sampling Weights

National student samples in TIMSS are designed to accurately represent the target populations within a specified margin of sampling error, as described previously. After the data have been collected and processed, sample statistics such as means and percentages that describe student characteristics are computed as weighted estimates of the corresponding population parameters, where the weighting factor is the sampling weight. A student's sampling weight is essentially the inverse of the student's probability of selection, with appropriate adjustments for nonresponse. In principle, the stratified two-stage sampling procedure used in TIMSS, where schools are sampled with probability proportional to school size and classes are sampled with probability inversely proportional to school size, provides student samples with equal selection probabilities. However, in practice, disproportionate sampling across explicit strata by varying the number of classes selected and differential patterns of nonresponse can result in varying selection probabilities, requiring a unique sampling weight for the students in each participating class in the study.

The student sampling weight in TIMSS is a combination of weighting components reflecting selection probabilities and sampling outcomes at three levels—school, class, and student. At each level, the weighting component consists of a basic weight that is the inverse of the probability of selection at that level, together with an adjustment for nonparticipation. The overall sampling weight for each student is the product of the three weighting components: school, class (within school), and student (within class).

Usually in TIMSS a country has only one set of sampling weights per target population (fourth and/or eighth grade). However, because of the introduction of the Problem Solving and Inquiry (PSI) tasks into the eTIMSS 2019 booklet rotation, eTIMSS countries have one set of sampling weights only for students who were assigned regular eTIMSS booklets, and a second set for all students, including those assigned PSI booklets. The first set of weights is computed in the same way for both eTIMSS and paperTIMSS countries (since the paperTIMSS countries do not use the PSI booklets) and is used for most analytic and reporting purposes in TIMSS 2019. Where necessary, these are referred to as the "TIMSS weights" to distinguish them from the second set, or "TIMSS+PSI weights."

In addition to the weights described above, countries with bridge data have a further set of weights exclusively for the bridge sample. Further details on the special weight adjustments for eTIMSS and the bridge data may be found in Chapter 9: Sampling Implementation. Regardless of whether they pertain to the regular booklet sample, the regular and PSI booklet sample, or the bridge sample, the procedure for calculating weights and nonparticipation adjustments remains the same.

School Weighting Component

Given that schools in TIMSS are sampled with probability proportional to school size, the basic school weight for the i^{th} sampled school (i.e., the inverse of the probability of the i^{th} school being sampled) is defined as:

$$BW_{sc}^{i} = \frac{M}{n \cdot m_i} \tag{3.1}$$

where n is the number of sampled schools, m_i is the measure of size for the i^{th} school, and

$$M = \sum_{i=1}^{N} m_i \tag{3.2}$$

where N is the total number of schools in the explicit stratum.[3,4]

School Nonparticipation Adjustment. If a sampled school does not participate in TIMSS and its two designated replacement schools do not participate, it is necessary to adjust the basic school weight to

3 For countries such as the Russian Federation that include a preliminary sampling stage, the basic school weight also incorporates the probability of selection in this preliminary stage. The basic school weight in such cases is the product of the preliminary stage weight and the school weight.

4 In schools selected for both the eTIMSS and the bridge samples, sampled classes are randomly assigned to either the eTIMSS or the bridge samples using the WinW3S software. If such a school has only one class, WinW3S randomly assigns the class to one of the samples (eTIMSS or bridge). In such cases, an adjustment is applied to the school weight in the corresponding explicit stratum of the non-selected sample.

compensate for the reduction in sample size. The school-level nonparticipation adjustment is calculated separately for each explicit stratum, as follows:

$$A_{sc} = \frac{n_s + n_{r1} + n_{r2} + n_{nr}}{n_s + n_{r1} + n_{r2}} \tag{3.3}$$

where n_s is the number of originally sampled schools that participated, n_{r1} and n_{r2} the number of first and second replacement schools, respectively, that participated, and n_{nr} is the number of schools that did not participate. Sampled schools that are found to be ineligible[5] are not included in the calculation of this adjustment.

Combining the basic school weight and the school nonparticipation adjustment, the final school weighting component for the i^{th} school becomes:

$$FW_{sc}^{\,i} = A_{sc} \cdot BW_{sc}^{\,i} \tag{3.4}$$

It should be noted that, as well as being a crucial component of the overall student weight, the final school weighting component is a sampling weight in its own right, and can be used in analyses where the school is the unit of analysis.

Class Weighting Component

The class weighting component reflects the class-within-school selection probability. After a school has been sampled and has agreed to participate in TIMSS, one or more classes are sampled with equal probability from the list of all classes in the school at the target grade. Because larger schools have more classes from which to sample than smaller schools, the probability of class selection varies with school size, with students in small schools more likely to have their class selected than students in large schools. This relatively greater selection probability for students in small schools offsets their lower selection probability at the first stage, where probability-proportional-to-size school sampling results in higher selection probabilities for larger schools.

The basic class-within-school weight for a sampled class is the inverse of the probability of the class being selected from all of the classes in its school. For the i^{th} sampled school, let C^i be the total number of eligible classes and c^i the number of sampled classes. Using equal probability sampling, the basic class weight for all sampled classes in the i^{th} school is:

$$BW_{cl}^{\,i} = \frac{C^i}{c^i} \tag{3.5}$$

For most TIMSS participants, c^i takes the values 1 or 2.

5 A sampled school is ineligible if it is found to contain no eligible students (i.e., no students in the target grade). Such schools usually are in the sampling frame by mistake or are schools that recently have closed.

Class Nonparticipation Adjustment. Basic class weights are calculated for all sampled classes in the sampled and replacement schools that participate in TIMSS. A class-level nonparticipation adjustment is applied to compensate for classes that do not participate or where the student participation rate is below 50 percent.[6] Such sampled classes are assigned a weight of zero. Class nonparticipation adjustments are applied at the explicit stratum level rather than at the school level to minimize the risk of bias. The adjustment is calculated as follows:

$$A_{cl} = \frac{\sum_{i}^{s+r1+r2} 1}{\sum_{i}^{s+r1+r2} \delta_i / c^i} \tag{3.6}$$

where c^i is the number of sampled classes in the i^{th} school, as defined earlier, and δ_i gives the number of participating classes in the i^{th} school.

Combining the basic class weight and the class nonparticipation adjustment, the final class weighting component, assigned to all sampled classes in the i^{th} school, becomes:

$$FW_{cl}^{i,j} = A_{cl} \cdot BW_{cl}^{i} \tag{3.7}$$

Student Weighting Component

The student weighting component represents the student-within-class selection probability. The basic student weight is the inverse of the probability of a student in a sampled class being selected.

In the typical TIMSS situation where intact classes are sampled, all students in the class are included, and so this probability is unity. However, under certain circumstances, students may be sampled within the class, and in this situation the probability is less than unity.

It should be noted that within-class student sampling is in effect when calculating the weights for the regular eTIMSS booklets for eTIMSS countries (the "TIMSS weight"). In this situation, students who were assigned a regular eTIMSS booklet are considered as being selected while students who received a PSI booklet are considered as not selected.

For an intact class with no student subsampling, the basic student weight for the j^{th} class in the i^{th} school is computed as follows:

$$BW_{st1}^{i,j} = 1.0 \tag{3.8}$$

6 When calculating the weights for the sample with regular TIMSS booklets only (without the PSI booklets), the nonparticipation criterion of below 50% is based on the full class, including the PSI booklets. Therefore, if 50% or more students from a class participated, regardless of the type of booklet received, the class is considered as participating when calculating the weights for the regular booklets only sample (the "TIMSS weight").

For classes with student subsampling, the basic student weight for the j^{th} class in the i^{th} school is:

$$BW_{st2}^{i,j} = \frac{n_{rg}^{i,j} + n_{bs}^{i,j}}{n_{rg}^{i,j}} \tag{3.9}$$

Where $n_{rg}^{i,j}$ is the number of students in the j^{th} class of the i^{th} school selected to participate in TIMSS and $n_{bs}^{i,j}$ is the number of students in the class not selected.

Calculating the eTIMSS weights for the regular booklet sample (the "TIMSS weight") involves student subsampling as described above, but with the added complication that the participation status is known for all the students in each sampled class. In this case, the basic student weight for the j^{th} class in the i^{th} school for this set of weights is given by:

$$BW_{st3}^{i,j} = \begin{cases} 1 & \text{for students who left school or were excluded,} \\ \dfrac{n_{rg'}^{i,j} + n_{bs'}^{i,j}}{n_{rg'}^{i,j}} & \text{for all other students who received a regular eTIMSS booklet} \end{cases} \tag{3.10}$$

where, $n_{rg'}^{i,j}$ and $n_{bs'}^{i,j}$ represent the number of students in the j^{th} class of the i^{th} school who received a regular eTIMSS booklet and the number of students in the j^{th} class of the i^{th} school who received a PSI booklet respectively, without counting students who either were excluded or left school after the class listing was completed.

Adjustment for Non-Participation. The student nonparticipation adjustment for the j^{th} classroom in the i^{th} school is calculated as:

$$A_{st1}^{i,j} = A_{st2}^{i,j} = A_{st3}^{i,j} = \frac{s_{rs}^{i,j} + s_{nr}^{i,j}}{s_{rs}^{i,j}} \tag{3.11}$$

where $s_{rs}^{i,j}$ is the number of participating students in the j^{th} class of the i^{th} school and $s_{nr}^{i,j}$ is the number of students sampled in this class who were expected to have assessment scores but did not participate in the assessment. For intact classes, the sum of $s_{rs}^{i,j}$ and $s_{nr}^{i,j}$ is the total number of students listed in the class, not counting excluded students or students who have left the school since class list was published. When calculating the "TIMSS weight" for eTIMSS countries (without the PSI booklets), the sum of $s_{rs}^{i,j}$ and $s_{nr}^{i,j}$ is the total number of students who received a regular eTIMSS booklet in the class, not counting excluded students or students who have left the school since class list was published

The final student weighting component for students in the j^{th} classroom of the i^{th} school is:

$$FW_{st}^{i,j} = A_{st\Delta}^{i,j} \cdot BW_{st\Delta}^{i,j} \tag{3.12}$$

where Δ equals 1 when there was no student subsampling (intact classes), 2 when a sample of students was drawn from the students in the class, and 3 when calculating the set of eTIMSS weights for only regular eTIMSS booklets.

Overall Student Sampling Weight. The overall student sampling weight is the product of the final weighting components for schools, classes, and students, as follows:

$$W^{i,j} = FW_{sc}^{i} \cdot FW_{cl}^{i,j} \cdot FW_{st}^{i,j} \tag{3.13}$$

Overall student sampling weights are only attributed to participating students, with non-participants weighted at 0. All student data reported in the TIMSS international reports are weighted by the overall student sampling weight, known as TOTWGT in the TIMSS international databases.

Participation Rates

Because nonparticipation can result in sample bias and misleading results, it is important that the schools, classes, and students that are sampled to participate in TIMSS actually take part in the assessments. To show the level of sampling participation in each country, TIMSS calculates both unweighted participation rates (i.e., based on simple counts of schools, classes, and students) and weighted participation rates based on the sampling weights described in the previous section. Unweighted participation rates provide a preliminary indicator that may be used to monitor progress in securing the participation of schools and classes, whereas weighted participation rates are the ultimate measure of sampling participation.

TIMSS reports weighted and unweighted participation rates for schools, classes, and students, as well as overall participation rates that are a combination of all three. To distinguish between participation based solely on originally sampled schools and participation that also relies on replacement schools, school and overall participation rates are computed separately for originally sampled schools only and for originally sampled together with replacement schools.

Unweighted School Participation Rate

The unweighted school participation rate is the ratio of the number of participating schools to the number of originally sampled schools, excluding any sampled schools found to be ineligible. A school is considered to be a participating school if at least one of its sampled classes has a student participation rate of at least 50 percent. The two unweighted school participation rates are calculated as follows:

R_{unw}^{sc-s} = unweighted school participation rate for originally sampled schools only

R_{unw}^{sc-r} = unweighted school participation rate, including originally sampled and first and second replacement schools

$$R_{unw}^{sc-s} = \frac{n_s}{n_s + n_{r1} + n_{r2} + n_{nr}} \tag{3.14}$$

$$R_{unw}^{sc-r} = \frac{n_s + n_{r1} + n_{r2}}{n_s + n_{r1} + n_{r2} + n_{nr}} \tag{3.15}$$

Unweighted Class Participation Rate

The unweighted class participation rate is the ratio of the number of sampled classes that participated to the number of classes sampled, as follows:

$$R_{unw}^{cl} = \frac{\sum_{i}^{s+r1+r2} c_*^i}{\sum_{i}^{s+r1+r2} c^i} \tag{3.16}$$

where c^i is the number of sampled classes in the i^{th} school, and c_*^i is the number of participating classes in the i^{th} school. Both summations are across all participating schools.

Unweighted Student Participation Rate

The unweighted student participation rate is the ratio of the number of selected students that participated in TIMSS to the total number of selected students that should have been assessed in the participating schools and classes. Classes where less than 50 percent of the students participate are considered to be not participating, and so students in such classes also are considered to be nonparticipants.[7] The unweighted student participation rate is computed as follows:

$$R_{unw}^{st} = \frac{\sum_{i,j} s_{rs}^{i,j}}{\sum_{i,j} s_{rs}^{i,j} + \sum_{i,j} s_{nr}^{i,j}} \tag{3.17}$$

Overall Unweighted Participation Rate

The overall unweighted participation rate is the product of the unweighted school, class, and student participation rates. Because TIMSS computes two versions of the unweighted school participation rate, one based on originally sampled schools only and the other including replacements as well as originally sampled schools, there also are two overall unweighted participation rates:

R_{unw}^{ov-s} = unweighted overall participation rate for originally sampled schools only

7 When calculating the "TIMSS weights" for eTIMSS countries (no PSI booklets), this 50% criteria is applied to all students regardless of the booklet they received.

R_{unw}^{ov-r} = unweighted overall participation rate, including originally sampled and first and second replacement schools

$$R_{unw}^{ov-s} = R_{unw}^{sc-s} \cdot R_{unw}^{cl} \cdot R_{unw}^{st} \tag{3.18}$$

$$R_{unw}^{ov-r} = R_{unw}^{sc-r} \cdot R_{unw}^{cl} \cdot R_{unw}^{st} \tag{3.19}$$

Weighted School Participation Rate

The weighted school participation rate is the ratio of two estimates of the size of the target student population. The numerator is derived from the measure of size of those sampled schools that participated in TIMSS and the denominator is the weighted estimate of the total student enrollment in the population. Weighted school participation rates are computed for originally sampled schools and for originally sampled and replacement schools combined, as follows:

R_{wtd}^{sc-s} = weighted school participation rate for originally sampled schools only

R_{wtd}^{sc-r} = weighted school participation rate, including originally sampled and first and second replacement schools

$$R_{wtd}^{sc-s} = \frac{\sum_{i,j}^{s} BW_{sc}^{i} \cdot FW_{cl}^{i,j} \cdot FW_{st}^{i,j}}{\sum_{i,j}^{s+r1+r2} FW_{sc}^{i} \cdot FW_{cl}^{i,j} \cdot FW_{st}^{i,j}} \tag{3.20}$$

$$R_{wtd}^{sc-r} = \frac{\sum_{i,j}^{s+r1+r2} BW_{sc}^{i} \cdot FW_{cl}^{i,j} \cdot FW_{st}^{i,j}}{\sum_{i,j}^{s+r1+r2} FW_{sc}^{i} \cdot FW_{cl}^{i,j} \cdot FW_{st}^{i,j}} \tag{3.21}$$

Summations in both the numerator and denominator are over all responding students and include appropriate class and student sampling weights. Note that the basic school weight appears in the numerator, whereas the final school weight appears in the denominator.

Weighted Class Participation Rate

The weighted class participation rate is computed as follows:

$$R_{wtd}^{cl} = \frac{\sum_{i,j}^{s+r1+r2} BW_{sc}^{i} \cdot BW_{cl}^{i,j} \cdot FW_{st}^{i,j}}{\sum_{i,j}^{s+r1+r2} BW_{sc}^{i} \cdot FW_{cl}^{i,j} \cdot FW_{st}^{i,j}} \tag{3.22}$$

where both the numerator and denominator are summations over all responding students from classes with at least 50 percent of their students participating in the study, and the appropriate student-level sampling weights are used. In this formula, the basic class weight appears in the numerator, whereas the final class weight appears in the denominator. The denominator in this formula is the same quantity that appears in the numerator of the weighted school participation rate for all schools, whether originally sampled or replacement.

Weighted Student Participation Rate

The weighted student participation rate is computed as follows:

$$R_{wtd}^{st} = \frac{\sum_{i,j}^{s+r1+r2} BW_{sc}^{i} \cdot BW_{cl}^{i,j} \cdot BW_{st}^{i,j}}{\sum_{i,j}^{s+r1+r2} BW_{sc}^{i} \cdot BW_{cl}^{i,j} \cdot FW_{st}^{i,j}} \tag{3.23}$$

where both the numerator and denominator are summations over all responding students from participating schools. In this formula, the basic student weight appears in the numerator, whereas the final student weight appears in the denominator. Also, the denominator in this formula is the same quantity that appears in the numerator of the weighted class participation rate for all participating schools, whether originally sampled or replacement.

Overall Weighted Participation Rate

The overall weighted participation rate is the product of the weighted school, class, and student participation rates. Because there are two versions of the weighted school participation rate, one based on originally sampled schools only and the other including replacement as well as originally sampled schools, there also are two overall weighted participation rates:

R_{wtd}^{ov-s} = weighted overall participation rate for originally sampled schools only

R_{wtd}^{ov-r} = weighted overall participation rate, including sampled, first and second replacement schools

$$R_{wtd}^{ov-s} = R_{wtd}^{sc-s} \cdot R_{wtd}^{cl} \cdot R_{wtd}^{st} \tag{3.24}$$

$$R_{wtd}^{ov-r} = R_{wtd}^{sc-r} \cdot R_{wtd}^{cl} \cdot R_{wtd}^{st} \tag{3.25}$$

Weighted school, class, student, and overall participation rates are computed for each TIMSS participant using these procedures.

References

Chowdhury, S., Chu, A., & Kaufman, S. (2000). Minimizing overlap in NCES surveys. In *JSM Proceedings*, Survey Research Methods Section. Alexandria, VA: American Statistical Association. 147-179. Retrieved from http://www.asasrms.org/Proceedings/papers/2000_025.pdf

UNESCO. (2012). *International Standard Classification of Education (ISCED) 2011*. Montreal, Quebec: UNESCO Institute for Statistics. Retrieved from http://uis.unesco.org/sites/default/files/documents/international-standard-classification-of-education-isced-2011-en.pdf

Appendix 3A: Sampling Schools

TIMSS employs random-start fixed-interval systematic sampling to draw the school sample, with each school selected with probability proportional to its size (PPS).

To sample schools using the PPS systematic sampling method, the schools from each explicit stratum in the sampling frame are sorted by implicit stratification variables and by their measure of size (MOS), as shown in the example in Exhibit 3.6. The MOS is accumulated from school to school and the running total (the Cumulative MOS) is listed next to each school. The cumulative MOS across the entire stratum (the Total MOS) is a measure of the size of the school population in the stratum (59,614 students in the example).

First Step: Compute the Sampling Interval

Dividing the Total MOS by the number of schools required for the sample (50 in the example) gives the sampling interval.

- **59,614 ÷ 50 = 1,192.2800**

Second Step: Generate a Random Start

Generate a random number from a uniform (0,1) distribution and multiply it by the sampling interval. The school whose cumulative MOS contains the resulting number is the first school in the sample.

- **0.5481 × 1,192.2800 = 653.4887**
- **School 1718**, with cumulative **MOS** of **690**, is the first school in the sample.

Third Step: Identify the Next School in the Sample (repeat until all schools have been sampled)

- Add the sampling interval to the number computed in the previous step.
- **653.4887 + 1,192.2800 = 1,845.7687**
- **School 0067,** with cumulative **MOS** of **1,855**, is the second school in the sample.
- Repeat until all schools have been sampled. For example, to identify the third school:
- **1,845.7687 + 1,192.2800 = 3,038.0487**
- **School 0333**, with cumulative **MOS** of **3,038**, is the third school in the sample.

Fourth Step: Identify Replacement Schools

Two replacement schools are identified for each sampled school. The first replacement (R1) is the school that immediately follows the sampled school in the sampling frame, and the second replacement (R2) the school that immediately precedes the sampled school.

Exhibit 3.6: Example of PPS Systematic Sampling—Schools

Sampling Parameters	
Total Number of Schools:	2,119
Total Measure of Size:	59,614
School Sample Size:	50
Sampling Interval:	1,192.2800
Random Start:	653.4887
First Step	
Compute the Sampling Interval:	
59,6914 ÷ 50 = 1,192.2800	
Second Step	
Generate a random start:	
0.5481 × 1,192.2800 = 653.4887	
Third Step (repeat until complete)	
Compute the next selection numbers:	
653.4887 + 1,192.2800 = 1,845.7687	
1,845.7687 + 1,192.2800 = 3,038.0487	
Fourth Step	
Identify Replacement Schools	
(R1, R2)	

School Identifier	School MOS	Cumulative MOS	Sampled Schools
0829	110	110	
0552	101	211	
1802	98	309	
1288	98	407	
2043	95	502	
0974	94	596	R2
1718	94	690	✓
1807	93	783	R1
0457	93	876	
0244	93	969	
1817	91	1,060	
1741	90	1,150	
1652	89	1,239	
0121	89	1,328	
0309	89	1,417	
0032	89	1,506	
0021	89	1,595	
0609	88	1,683	
0399	86	1,769	R2
0067	86	1,855	✓
0202	86	1,941	R1
0063	86	2,027	
1467	86	2,113	
1381	86	2,199	
1043	84	2,283	
1318	84	2,367	
0659	84	2,451	
0612	83	2,534	
1696	82	2,616	
0867	82	2,698	
0537	81	2,779	
1794	80	2,859	
0695	80	2,939	
0031	80	3,019	R2
0333	79	3,098	✓
0051	79	3,177	R1
0384	79	3,256	
1361	79	3,335	
1189	79	3,414	
0731	78	3,492	
0634	78	3,570	
1230	77	3,647	

Appendix 3B: School Sampling Design Options to Accommodate Other Samples

TIMSS provides optional modifications to its sampling design for countries that want to maximize or minimize sampling overlap between schools sampled by TIMSS at the fourth and eighth grades as well as for countries that want to minimize overlap between schools sampled for TIMSS and schools sampled for other national or international assessments.

To provide options for countries in designing their school samples, Statistics Canada implements two special sampling procedures. Method A is applied when data collection occurs simultaneously for two or more populations (as is the case in 2019 with TIMSS at fourth grade and eighth grade) and the country wants to control the overlap between the schools. Method B is used primarily to ensure that the TIMSS samples avoid schools sampled for other studies, and also used when Method A is not appropriate.

Sampling Method A: Sampling Modifications for Simultaneous Data Collection

This procedure stratifies the school population according to whether schools contain students from both populations to be sampled (fourth and eighth grades, for example), or students from one population only (fourth grade only or eighth grade only) as a way of controlling sample overlap. Each school is assigned a measure of size (MOS) based on the number of students in the two populations combined (i.e., fourth grade and eighth grade combined). Schools are sampled according to the sampling design described in this chapter. When selecting schools from strata comprising students from both populations, a country can choose to maximize or minimize the number of schools to be sampled at each grade level.

The example below in Exhibit 3.7 shows a hypothetical country participating in TIMSS at both grades. For reasons of administrative efficiency, the country wants to maximize the overlap between the fourth and eighth grade school samples. The 8,805 schools from the combined school frames (fourth and eighth grades) were first split in three strata and then a school sample of 164 was drawn as shown in the exhibit.

Exhibit 3.7: Example of Method A - Allocation of School Samples in a Country Participating at Two Grade Levels

Overlap Strata	Total Sampled Schools	Allocation	
		To TIMSS Grade 4	To TIMSS Grade 8
Grade 4 only	14	14	0
Grade 8 only	14	0	14
Grade 4 & Grade 8	136	136	136
Total	**164**	**150**	**150**

Choosing as many schools as possible from the Grade 4 & Grade 8 stratum resulted in a sample of 150 schools (136+14) for each grade level, from a total of 164 sampled schools. In this case, both studies were administered in the 136 schools selected from the Grade 4 & Grade 8 stratum.

This sampling technique was most often used for TIMSS countries and benchmarking participants that had schools with students in both fourth and eighth grade populations, where there was a strong correlation between the measure of size at both grades across these schools, and when school samples could be drawn at the same time.

Sampling Method B: Sampling Modifications for Sequential Data Collection

Method B was used to minimize overlap with another study such as a national study that also samples schools, and was also used when Method A was not appropriate (e.g., low correlation between MOS for fourth grade and eighth grade, samples not drawn simultaneously). In Method B, schools were sampled using a technique described in Chowdhury, Chu, and Kaufman (2000). As explained by the authors, the method can be used to either minimize or maximize overlap amongst several samples. This method is illustrated below with an example where the aim was to minimize the overlap between a current sample of schools S_2 and a previously selected school sample S_1. (For a complete description of the method, readers are referred to the original paper).

Let RL (Response Load) be the number of times a school was sampled from previous samples. In this example, given that there is only one previous sample, RL takes the value 1 if the school was already selected and 0 otherwise.

Given that the RL variable splits the current school frame in two distinct subsets of schools, S_1 where RL=1 and $\overline{S}_1$ where RL=0, we have the following relation:

$$P_i(S_2) = P_i(S_2|S_1) \cdot P_i(S_1) + P_i(S_2|\overline{S}_1) \cdot P_i(\overline{S}_1) \tag{3.26}$$

where $P_i(S_j)$ gives the probability that school i be selected in the sample (S_j), and $P_i(S_j|S_k)$ gives the probability that school i be selected in sample (S_j) given that school i already belongs to (S_k). The idea here is to derive the conditional probabilities in such a way that the unconditional probability of selecting a school in the current sample, $P_i(S_2)$, be equal to the expected probability (as defined by the TIMSS sample design).

Note that the first term after the equal sign in equation (3.26) is related to cases where the school response load is 1, while the last term is related to cases where the school response load is 0. Therefore, minimizing the sample overlap is equivalent to zeroing the first term. In such case, equation (3.26) becomes:

$$P_i(S_2) = 0 \cdot P_i(S_1) + P_i(S_2|\overline{S}_1) \cdot P_i(\overline{S}_1) \tag{3.27}$$

and consequently,

$$P_i(S_2|\overline{S}_1) = P_i(S_2)/P_i(\overline{S}_1) \tag{3.28}$$

In other words, in the current sample S_2, schools would be selected with the following conditional probabilities:

$$\begin{cases} 0 & \text{if school } i \text{ was already selected in the first sample,} \\ P_i(S_2)/P_i(\overline{S}_1) & \text{otherwise} \end{cases} \tag{3.29}$$

However, equation (3.26) no longer holds if expression $P_i(S_2)/P_i(\overline{S}_1)$ is greater than 1. This can be avoided by setting 1 as an upper bound. We now have the following expression:

$$P_i(S_2) = P_i(S_2|S_1) \cdot P_i(S_1) + 1 \cdot P_i(\overline{S}_1) \tag{3.30}$$

and consequently

$$\frac{P_i(S_2) - P_i(\overline{S}_1)}{P_i(S_1)} = P_i(S_2|S_1) \tag{3.31}$$

Combining these two results, the conditional probabilities to use when selecting the current sample of schools are given by:

$$\begin{cases} \text{Max} \left[0 , \dfrac{P_i(S_2) - P_i(\overline{S}_1)}{P_i(S_1)} \right] & \text{if school } i \text{ was already selected in the first sample,} \\ \text{Min} \left[\dfrac{P_i(S_2)}{P_i(\overline{S}_1)} , 1 \right] & \text{otherwise} \end{cases} \tag{3.32}$$

Note that maximizing rather than minimizing the overlap between two studies can be done by simply zeroing the last term of equation (3.26) rather than zeroing the first term, and following the above logic to get the conditional probabilities. The Chowdhury et al. (2000) method can be generalized to more than two samples as described in their paper.

Further details about the implementation of this method for the countries and benchmark participants can be found in the Sample Implementation chapter.