

Pierre Foy
Statistics Canada

4.1 OVERVIEW

The basic sample design used in TIMSS Populations 1 and 2 was a two-stage stratified cluster design.¹ The first stage consisted of a sample of schools; the second stage consisted of samples of one intact mathematics classroom from each eligible target grade in the sampled schools. The design required schools to be sampled using a probability proportional to size (PPS) systematic method, as described by Foy, Rust, and Schleicher (1996), and classrooms to be sampled with equal probabilities (Schleicher and Siniscalco, 1996). While TIMSS had a basic design for how the national representative samples of students in Populations 1 and 2 were to be drawn, aspects of the design were adapted to national conditions and analytical needs. For example, many countries stratified the school sampling frame by variables of national interest. As another example, some countries chose to sample two classrooms from the upper grade of the target population. Chapter 2 of this report documents in detail the national samples for TIMSS Populations 1 and 2.

While a multi-stage stratified cluster design greatly enhances the feasibility of data collection, it results in differential probabilities of selection; consequently, each student in the assessment does not necessarily represent the same number of students in the population, as would be the case if a simple random sampling approach were employed. To account for differential probabilities of selection due to the nature of the design and to ensure accurate survey estimates, TIMSS computed a sampling weight for each student that participated in the assessment. This chapter documents the calculation of the sampling weights for students sampled for the Populations 1 and 2 main assessment and for those students subsampled to also take part in the performance assessment.²

4.2 WEIGHTING PROCEDURES

The general weighting procedure for TIMSS required three steps. The first step for all target populations consisted of calculating a school weight. The school weight also incorporates weighting factors from any additional front-end sampling stages required

¹ The target populations are defined as follows:

Population 1: Students enrolled in the two adjacent grades where most 9-year-old students are found at the time of testing (third and fourth grades in many countries)

Population 2: Students enrolled in the two adjacent grades where most 13-year-old students are found at the time of testing (seventh and eighth grades in many countries).

² See Harmon and Kelly (1996) for details of the sampling procedures for the performance assessment.

by some TIMSS participants.³ A school-level nonresponse adjustment was applied to the school weight; it was calculated independently for each design domain or explicit stratum.

The second step consisted of calculating a classroom weight. A classroom-level nonresponse adjustment was not necessary since in most cases a single classroom was selected per school at each grade level. When only one of the sampled classrooms in a school participated, a grade-specific school-level response adjustment was used. When one of two selected classrooms in a school (when a country chose to sample two classrooms per grade) did not participate, the classroom weight was calculated as though a single classroom had been selected in the first place. The classroom weight was calculated independently for each school and grade.

The final step consisted of calculating a student weight. A student-level nonresponse adjustment was applied to the student weight. The student weight was calculated independently for each sampled classroom.

The overall sampling weight attached to each student record is the product of the three intermediate weights: the first stage (school) weight, the second stage (classroom) weight, and the third stage (student) weight.

The overall sampling weight attached to each student in the performance assessment sub-sample is the product of the first stage weight adjusted for the subsampling of schools required, the second stage weight, and the third stage weight adjusted for the subsampling of students required at this stage.

4.2.1 First-Stage (School) Weight

The first stage weight represents the inverse of the first stage selection probability assigned to a sampled school. The TIMSS sample design required that school selection probabilities be proportional to school size, with school size being enrollment in the target grades. The basic first stage weight for the i th sampled school was thus defined as

$$BW_{sc}^i = \frac{M}{n * m_i}$$

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

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

where N is the total number of schools in the stratum.

The basic first stage weight also incorporates a weighting factor or factors resulting from additional front-end sampling stages that were required by some TIMSS participants. This occurred when geographical regions were sampled before schools were se-

³ For example, the United States sampled school districts as primary sampling units (PSUs), and then schools within the sampled PSUs.

lected. The calculation of such weighting factors is similar to the first stage weight since sampling geographical regions was also done with probability proportional to size (PPS). The resulting first stage weight is simply the product of the "region" weight and the first stage weight as described earlier.

In some countries, schools were selected with equal probabilities. This generally occurred when no reliable measure of school size is available. In this case, the basic first stage weight for the i th sampled school was defined as

$$BW_{sc}^i = \frac{N}{n}$$

where n is the number of sampled schools and N is the total number of schools in the stratum. It should be noted that in this case the basic weight for all sampled schools is identical.

4.2.1.1 School-Level Response Rate (Participation Rate)

A school-level response rate, weighted and unweighted, was calculated to measure the proportion of originally selected schools that ultimately participated in the assessment. Since replacement schools were used to maintain the sample size, school-level response rates have been reported both with and without the use of replacement schools. The calculation of the response rate used the following terms, derived from the data collection:

n_{ex} = number of sampled schools that should have been excluded

n_{op} = number of originally sampled schools that participated

n_{rp} = number of replacement schools that participated

n_{nr} = number of non-responding schools (neither the originally selected schools nor their replacements participating.)

Note that the following equation holds:

$$n_{ex} + n_{op} + n_{rp} + n_{nr} = n$$

The unweighted school-level response rate is defined as the ratio of originally sampled schools that participated to the total number of sampled schools minus any excluded schools. It was calculated by the following equation:

$$R_{unw}^{sc} = \frac{n_{op}}{n_{op} + n_{rp} + n_{nr}}$$

The weighted school-level response rate is defined in a similar manner. The weight assigned to the i th sampled school for this purpose is the sampling interval used to select it, SI_{sc}^i . The weighted school-level response rate, based solely on originally selected schools, is therefore the ratio of the weighted sum of originally sampled schools that participated to the weighted sum of all sampled schools less any excluded schools. It was calculated by the following equation:

$$R_w^{sc} = \frac{\sum_{i=1}^{n_{op}} SI_{sc}^i}{\sum_{i=1}^{n_{op}} SI_{sc}^i + \sum_{i=1}^{n_{rp}} SI_{sc}^i + \sum_{i=1}^{n_{nr}} SI_{sc}^i}$$

The weighted school-level response rate, including replacement schools, was calculated by the following equation:

$$R_{w,rp}^{sc} = \frac{\sum_{i=1}^{n_{op}} SI_{sc}^i + \sum_{i=1}^{n_{rp}} SI_{sc}^i}{\sum_{i=1}^{n_{op}} SI_{sc}^i + \sum_{i=1}^{n_{rp}} SI_{sc}^i + \sum_{i=1}^{n_{nr}} SI_{sc}^i}$$

4.2.1.2 School-Level Nonresponse Adjustment

First stage weights were calculated for originally sampled schools and replacement schools that participated. Any sampled schools that were no longer eligible were removed from the calculation of this nonresponse adjustment. Examples are secondary schools included in the sampling frame by mistake and schools that no longer existed. The school-level nonresponse adjustment was calculated separately for each design domain and explicit stratum.

The school-level nonresponse adjustment was calculated as follows:

$$A_{sc} = \frac{n - n_{ex}}{n_{op} + n_{rp}}$$

and the final first stage weight for the i th school thus becomes

$$FW_{sc}^i = A_{sc} * BW_{sc}^i$$

In the event that a sampled school had participating classrooms in only one grade when both grades were in fact present, the school-level nonresponse adjustment becomes grade-specific. Such a school was considered a participant for the grade in which students were tested but as a non-participant for the grade in which no students were tested. This led also to the calculation of separate school-level response rates by grade.

4.2.2 Second-Stage (Classroom) Weight

The second stage weight represents the inverse of the second stage selection probability assigned to a sampled classroom. Classrooms were sampled in one of two ways in Population 1 and Population 2:

- Equal probability if there was no subsampling of students within a classroom
- Probability proportional to classroom size if subsampling of students within a classroom was required

The second stage weight was calculated independently for each grade within a sampled school in Population 1 and Population 2.

A nonresponse adjustment was not required for the second stage weight. Where the classroom selected in one target grade did not participate but the sampled classroom in the other target grade did, the separate first stage nonresponse adjustments were applied by target grade.

4.2.2.1 Equal Probability Weighting

For grade g within the i th school, let $C^{g,i}$ be the total number of classrooms and c^g be the number of sampled classrooms. Using equal probability sampling, the final second stage weight assigned to all sampled classrooms from grade g in the i th school was

$$FW_{cl1}^{g,i} = \frac{C^{g,i}}{c^g}$$

As a rule, c^g takes the value 1 or 2 and remains fixed for all sampled schools. In cases where c^g has the value 2 and only one of the sampled classrooms participated, a classroom-level nonresponse adjustment was applied to the second stage weight by multiplying it by the factor 2.

4.2.2.2 Probability Proportional to Size (PPS) Weighting

For grade g within the i th school, let $k^{g,i,j}$ be the size of the j th classroom. Using PPS sampling, the final second stage weight assigned to the j th sampled classroom from grade g in the i th school was

$$FW_{cl2}^{g,i,j} = \frac{K^{g,i}}{c^g * k^{g,i,j}}$$

where c^g is the number of sampled classrooms as defined earlier and

$$K^{g,i} = \sum_{j=1}^{c^g} k^{g,i,j}$$

Again, as a rule, c^g takes the value 1 or 2 and will remain fixed for all sampled schools. In cases where c^g has the value 2, and only one of the sampled classrooms participated, a classroom-level nonresponse adjustment was applied to the second stage weight by multiplying it by the factor 2.

4.2.3 Third-Stage (Student) Weight

The third stage weight represents the inverse of the third stage selection probability attached to a sampled student. If intact classrooms were sampled as specified in Foy, Rust, and Schleicher (1996), then the basic third stage weight for the j th grade g classroom in the i th school was

$$BW_{st}^{g,i,j} = 1.0$$

If, on the other hand, subsampling of students was required within sampled classrooms, then the basic third stage weight for the j th grade g classroom in the i th school was

$$BW_{st}^{g,i,j} = \frac{k^{g,i,j}}{s^g}$$

where $k^{g,i,j}$ is the size of the j th grade g classroom in the i th school, as defined earlier, and s^g is the number of sampled students per sampled classroom. The latter number usually remains constant for all sampled classrooms in a grade.

4.2.3.1 Student-Level Response Rate (Participation Rate) and Adjustment

The basic third stage weight requires an adjustment to reflect the outcome of the data collection efforts. The following terms were derived from the data collection for each sampled classroom:

$s_{ex}^{g,i,j}$ = number of sampled students that should have been excluded

$s_{rs}^{g,i,j}$ = number of sampled students that participated

$s_{nr}^{g,i,j}$ = number of sampled students that did not participate.

Note that the following equation holds:

$$s_{ex}^{g,i,j} + s_{rs}^{g,i,j} + s_{nr}^{g,i,j} = s^{g,i,j}$$

where $s^{g,i,j}$ is the number of sampled students per sampled classroom. This number should be constant if subsampling of students is done within each sampled classroom and represents the classroom size, $k^{g,i,j}$, when intact classrooms are tested.

The student-level response rate, for a given classroom, was calculated as follows:

$$R^{st} = \frac{s_{rs}^{g,i,j}}{s_{rs}^{g,i,j} + s_{nr}^{g,i,j}}$$

Excluded students (i.e., those meeting the guidelines for student-level exclusions specified in Foy, Rust, and Schleicher, 1996) were not included in the calculation of the response rate.

The student-level nonresponse adjustment was calculated as follows:

$$A_{st}^{g,i,j} = \frac{s_{rs}^{g,i,j} + s_{nr}^{g,i,j}}{s_{rs}^{g,i,j}}$$

Note that the student-level nonresponse adjustment is simply the inverse of the student-level response rate. The final third stage weight for the j th grade g classroom in the i th school thus becomes

$$FW_{st}^{g,i,j} = A_{st}^{g,i,j} * BW_{st}^{g,i,j}$$

The weighted overall student-level response rate was computed as follows:

$$R_w^{st} = \frac{\sum_{i=1}^{rs} BW_{sc}^i * BW_{cl1}^{g,i,j} * BW_{st}^{g,i,j}}{\sum_{i=1}^{rs+nr} BW_{sc}^i * BW_{cl1}^{g,i,j} * BW_{st}^{g,i,j}}$$

where the numerator is the summation of the basic weights over all responding students, and the denominator is the summation of the basic weights over all responding and nonresponding students. Weighted student response rates were reported separately by grade in the TIMSS international reports.

4.2.4 Overall Sampling Weights

The overall sampling weight is simply the product of the final first stage weight, the appropriate final second stage weight, and the appropriate final third stage weight. If intact classrooms were tested, then the overall sampling weight was

$$W^{g,i,j} = FW_{sc}^i * FW_{sc}^{g,i,j} * FW_{st}^{g,i,j}$$

If subsampling within classrooms was done, then the overall sampling weight was

$$W^{g,i,j} = FW_{sc}^i * FW_{cl2}^{g,i,j} * FW_{st}^{g,i,j}$$

It is important to note that sampling weights varied by school, grade, and classroom. However, students within the same classroom have the same sampling weights.

The use of sampling weights is critical to obtaining proper survey estimates when sampling techniques other than simple random sampling are used. TIMSS has produced a sampling weight for each student sampled for the TIMSS main (written) assessment and subsampled for the performance assessment. Secondary analysts using the TIMSS data will need to be aware of this and use the proper weights when conducting analyses and reporting results.

REFERENCES

- Foy, P., Rust, K., and Schleicher, A. (1996). Sample design. In M.O. Martin and D.L. Kelly (Eds.), *TIMSS technical report, volume I: Design and development*. Chestnut Hill, MA: Boston College.
- Harmon, M. and Kelly, D.L. (1996) Performance assessment. In M.O. Martin and D.L. Kelly (Eds.), *TIMSS technical report, volume I: Design and development*. Chestnut Hill, MA: Boston College.
- Schleicher, A. and Siniscalco, M.T. (1996). Field operations. In M.O. Martin and D.L. Kelly (Eds.), *TIMSS technical report, volume I: Design and development*. Chestnut Hill, MA: Boston College.