# Appendix A

# TIMSS Design and Procedures

## Test Design

The TIMSS tests were developed through an international consensus involving input from experts in mathematics, science, and educational measurement. The TIMSS Subject Matter Advisory Committee ensured that the tests reflected current thinking and priorities within the fields of mathematics and science. Every effort was made to help ensure that the tests represented the curricula of the participating countries and that the items exhibited no bias toward or against particular countries. This involved modifying specifications in accordance with data from the curriculum analysis component, obtaining ratings of the items by subject matter specialists within the participating countries, and conducting thorough statistical item analyses of data collected in the pilot testing. The final forms of the tests were endorsed by the National Research Coordinators (NRCs) of the participating countries.

TIMSS tested students at three points in their schooling – primary school (third and fourth grades in most countries), middle school (seventh and eighth grades in most countries), and the final year of secondary school. In mathematics, the third- and fourth-grade tests included items from six content areas: whole numbers; fractions and proportionality; measurement, estimation, and number sense; data representation, analysis, and probability; geometry; and algebra. For the seventh and eighth grades, the mathematics test included items from six content areas: fractions and number sense; proportionality; measurement; data representation, analysis, and probability; geometry; and algebra. In science, the primary-school test included items from four content areas: earth science; life science; physical science; and environmental issues and the nature of science. For the seventh and eighth grades, the science test included items from five content areas: earth science; life science; chemistry; physics; and environmental issues and the nature of science.

The mathematics and science literacy tests for final-year students were designed to assess students' general knowledge and understanding of mathematical and scientific principles. The mathematics items covered number sense, including fractions, percentages, and proportionality. Algebraic sense, measurement, and estimation are also covered, as are data representation and analysis. Reasoning and social utility were emphasized in several items. A general criterion in selecting the items was that they should involve the types of mathematics questions that could arise in real-life situations and that they be contextualized accordingly. Similarly, the science items selected for use in the TIMSS literacy test were organized according to three areas of science – earth science, life science, and physical science – and included a reasoning and social utility component. The emphasis was on measuring how well students could use their knowledge in addressing real-world problems having a science component. The test was designed to enable reporting for mathematics literacy and science literacy separately as well as overall.

To maximize the content coverage of the TIMSS tests, yet minimize the burden on individual students, TIMSS used a multiple

matrix sampling design whereby subsets of items from the total item pool were administered to sub-samples of students.[11] Each student responded to a subset of the total item pool; by aggregating data across booklets, TIMSS was able to derive population estimates of mathematics and science achievement. TIMSS does not provide individual proficiency estimates. The design was nearly identical for the primary and middle school assessments, but different for the assessment of final-year students.

For the primary and middle school tests, items were assigned to 26 mutually exclusive groups or "clusters." The clusters were then assigned to eight test booklets so that one cluster appeared in all test booklets, some clusters appeared in several test booklets, and some clusters appeared in one test booklet. Each test booklet contained mathematics and science test items. The test booklets were systematically distributed to students and each student completed one. Primary-school students had 64 minutes to complete their test booklets, and middle-school students had 90 minutes.

For the final year of secondary-school assessment, there were nine test booklets containing the assessment material for mathematics and science literacy, advanced mathematics, and physics. Two of these booklets contained exclusively mathematics and science literacy items, and one booklet contained some mathematics and science literacy items. Students were assigned one of nine booklets depending upon their academic preparation; all students were eligible to receive the two mathematics and science literacy booklets. Final-year students had 90 minutes to complete their booklets.

In each test, approximately one-quarter of the items were in the free-response format, requiring students to generate and write their own answers. Designed to take up about one-third of students' response time, some of these questions asked for short answers while others required extended responses in which students needed to show their work. The remaining questions were in multiple-choice format. In scoring the tests, correct answers to most questions were worth one point. Consistent with the approach of allotting longer response times for constructed-response questions than for multiple-choice questions, responses to some of these questions (particularly those requiring extended responses) could earn partial credit, with a fully correct answer being awarded two or three points.

## Sampling

TIMSS included testing at three separate populations.

> Population 1: Students enrolled in the two adjacent grades that contained the largest proportion of 9-year-old students at the time of testing – third- and fourth-grade students in most countries.
>
> Population 2: Students enrolled in the two adjacent grades that contained the largest proportion of 13-year-old students at the time of testing – seventh- and eighth-grade students in most countries.

[11] The TIMSS test design is fully described in Adams, R.J. and Gonzalez, E.J. (1996). "TIMSS Test Design" in M.O. Martin and D.L. Kelly (Eds.), *Third International Mathematics and Science Study Technical Report, Volume I.* Chestnut Hill, MA: Boston College.

Population 3: Students in their final year of secondary education. As an additional option, countries could test two special subgroups of these students: (1) students taking advanced courses in mathematics and (2) students taking physics.

Countries participating in the study were required to test the students in the two grades at Population 2, but could choose whether or not to participate at the other levels.

The selection of valid and efficient samples is crucial to the quality and success of an international comparative study such as TIMSS. The accuracy of the survey results depends on the quality of sampling information available and on the quality of the sampling activities themselves. For TIMSS, NRCs worked on all phases of sampling with staff from Statistics Canada. NRCs were trained in how to select the school and student samples and in the use of the sampling software. In consultation with the TIMSS sampling referee (Keith Rust, Westat), staff from Statistics Canada reviewed the national sampling plans, sampling data, sampling frames, and sample execution. This documentation was used by the International Study Center in consultation with Statistics Canada, the sampling referee, and the Technical Advisory Committee to evaluate the quality of the samples. In the achievement tables presented in Chapter 1 of this report, countries are grouped according to the extent to which they met the TIMSS sampling requirements. In the remaining tables, countries that did not meet the TIMSS standards for sampling at the eighth grade were excluded from the analysis.

To achieve acceptable participation rates, countries needed to assess 85% of both the schools and students or have a combined rate (the product of school and student participation) of 75% with or without replacement schools. Countries that met the participation guidelines only after including replacement schools are annotated in Chapter 1. Countries not reaching at least 50% of school participation without the use of replacement schools, or that failed to reach the sampling participation standard even with the inclusion of replacement schools are shown in a separate category for each of the tables in Chapter 1. Countries also needed to comply with the TIMSS guidelines for grade selection and classroom sampling. In particular, several countries did not comply with the guidelines for randomly sampling classrooms and these are shown in the last category in the Chapter 1 tables.

Exhibits A.1 through A.5 present the TIMSS school and student sample sizes by gender for each of the participating countries. Exhibit A.6 presents the name of the upper grade and the number of formal years of schooling associated with that grade for each country in populations 1 and 2.

## Data Collection Procedures

Each participating country was responsible for carrying out all aspects of the data collection, using standardized procedures developed for the study. Training manuals were developed for school coordinators and test administrators that detailed procedures for receipt and distribution of materials as well as for the activities related to the testing sessions. The test administrator manuals covered test security, standardized scripts to regulate directions and timing, rules for answering students' questions, and steps to ensure that

identification on the test booklets and questionnaires corresponded to the information on the forms used to track students.

Each country was responsible for conducting quality control procedures and for describing these in a report provided to the International Study Center. In addition, the International Study Center considered it essential to establish some method to monitor compliance with standard procedures. NRCs were asked to nominate a person, such as a retired school teacher, to serve as quality control monitor for their countries, and in almost all cases the International Study Center adopted the NRCs' first suggestion. The International Study Center developed manuals for the quality control monitors and briefed them in two-day training sessions about TIMSS, the responsibilities of the national centers in conducting the study, and their own roles and responsibilities.

The quality control monitors interviewed the NRCs about data collection plans and procedures. They also selected about 10 schools to visit, where they observed testing sessions and interviewed school coordinators.[12] The results of the interviews indicate that, in general, NRCs had prepared well for data collection and, despite the heavy demands of the schedule and shortages of resources, were in a position to collect the data in an efficient and professional manner. Similarly, the TIMSS tests appeared to have been administered in compliance with international procedures throughout the activities preliminary to the testing session, those during testing, and the school-level activities related to receiving, distributing, and returning materials from the national centers.

## Scoring the Free-Response Items

Because about one-third of the written test time was devoted to free-response items, TIMSS needed to develop procedures for reliably evaluating student responses within and across countries. Scoring used two-digit codes with rubrics specific to each item. Development of the rubrics was led by the Norwegian TIMSS national center. The first digit designates the correctness level of the response. The second digit, combined with the first, represents a diagnostic code used to identify specific types of approaches, strategies, or common errors and misconceptions. Although not specifically used to estimate overall proficiency in mathematics and science, analyses of responses based on the second digit should provide insight into ways to help students better understand mathematics concepts and problem-solving approaches.

To ensure reliable scoring procedures based on the TIMSS rubrics, the International Study Center prepared guides containing the rubrics and explaining how to implement them together with example student responses for the various rubric categories. These guides, together with more examples of student responses for practice in applying the rubrics, were used as a basis for an ambitious series of regional training sessions. These sessions

[12] The results of the interviews and observations by the quality control monitors are presented in Martin, M.O., Hoyle, C.D., and Gregory, K.D. (1996) "Monitoring the TIMSS Data Collection" and "Observing the TIMSS Test Administration" both in M.O. Martin and I.V.S. Mullis (Eds.), *Third International Mathematics and Science Study: Quality Assurance in Data Collection.* Chestnut Hill, MA: Boston College.

were designed to assist representatives of national centers who would then be responsible for training personnel in their countries to apply the two-digit codes reliably.[13]

To gather and document empirical information about the within-country agreement among scorers, TIMSS developed a procedure whereby systematic subsamples of some 10% of the students' responses were coded independently by two scorers. The percentage of exact agreement between the scorers was computed for each free-response item based on both the score level (first digit) and the diagnostic code (second digit) level. A very high percentage of exact agreement at the score level was observed for the free-response items on all TIMSS tests.[14]

## Data Processing

To ensure the availability of comparable, high-quality data for analysis, TIMSS undertook a rigorous set of quality control steps to create the international database.[15] TIMSS prepared manuals and software for countries to use in entering their data so that the information would be in a standard international format before being forwarded to the IEA Data Processing Center in Hamburg. Upon arrival at the Center, the data from each country underwent an exhaustive cleaning process. That process involved several iterative steps and procedures designed to identify, document, and correct deviations from the international instruments, file structures, and coding schemes. The process also emphasized consistency of information within national data sets and appropriate linking among the many student, teacher, and school data files.

Throughout the process, the data were checked and double-checked by the IEA Data Processing Center, the International Study Center, and the national centers. The national centers were contacted regularly and given multiple opportunities to review the data for their countries. In conjunction with the Australian Council for Educational Research (ACER), the International Study Center reviewed the item statistics of each cognitive item in each country to identify poorly performing items. Usually the poor statistics (negative point-biserials for the key, large item-by-country interactions, and statistics indicating lack of fit with the model) were a result of deviations in translation, adaptation, or printing.

[13] The procedures used in the training sessions are documented in Mullis, I.V.S., Garden, R.A., and Jones, C.A. (1996) "Training for Scoring the TIMSS Free-Response Items" in M.O. Martin and D.L. Kelly (Eds.), *Third International Mathematics and Science Study Technical Report, Volume I.* Chestnut Hill, MA: Boston College.

[14] Summaries of the scoring reliability data for each test are included in the appendices of the international reports (see references in "Summary of Results" chapter).

[15] These steps are detailed in Jungclaus, H. and Bruneforth, M. (1996). "Data Consistency Checking Across Countries" in M.O. Martin and D.L. Kelly (Eds.), *Third International Mathematics and Science Study Technical Report, Volume I.* Chestnut Hill, MA: Boston College.

## IRT Scaling and Data Analysis

The mathematics and science achievement results were summarized using an item response theory (IRT) scaling method (Rasch model).[16] This scaling method produces a test score by averaging the responses of each student to the items they took in a way that takes into account the difficulty of each item. The method used in TIMSS includes refinements that enable reliable scores to be produced even though individual students responded to relatively small subsets of the total mathematics item pool. Analyses of the response patterns of students from participating countries indicated that, although the items in each TIMSS test address a wide range of mathematics or science content, the performance of the students across the items was sufficiently consistent to be usefully summarized in a single score per test.

The IRT method was preferred for developing comparable estimates of performance for all students, since students answered different test items depending upon which test booklet they received. The IRT analysis provides a common scale on which performance can be compared across countries. In addition to providing a basis for estimating mean achievement, scale scores permit estimates of how students within countries vary and provide information on percentiles of performance. For Population 1 and Population 2, each scale was standardized using students from both the grades tested. When all participating countries and grades are treated equally, the TIMSS scale average is 500 and the standard deviation is 100. Since the countries vary in size, each country was reweighted to contribute equally to the mean and standard deviation of the scale. The international averages of the Population 1 scale scores (mathematics and science) were constructed to be the averages of the 26 means of countries that were available at fourth grade and the 24 means of those at third grade. The international averages of the Population 2 scale scores (mathematics and science) were constructed to be the averages of the 41 means of countries that were available at eighth grade and the 39 means of those at seventh grade. For the Population 3 mathematics and science literacy assessment, the mathematics literacy scale and the science literacy scale were constructed using data from the 21 countries that participated in the assessment and have an average of 500 and a standard deviation of 100.

[16] The TIMSS scaling model is fully documented in Adams, R.J., Wu, M.L., and Macaskill, G. (1997). "Scaling Methodology and Procedures for the Mathematics and Science Scales" in M.O. Martin and D.L. Kelly (Eds.), *Third International Mathematics and Science Study Technical Report, Volume II.* Chestnut Hill, MA: Boston College.

## The Gender Difference Index (GDI)

This report used a statistic known as the Gender Difference Index (GDI) in order to explore differences by gender at the item level. The GDI was developed for use in a previous IEA study of reading literacy.[17] The formula for calculating the GDI is listed below:

where: $$D = -\beta X.Y * 100 = -\frac{(P_0 - P_1)}{(P_0 + P_1)(2 - P_0 - P_1)} * 100$$

$D$ = (The percentage of females among those who answered incorrectly) – (The percentage of females among those who answered correctly)

$\beta$ = Regression coefficient

$X$ = A dummy variable for gender where X = 0 for females and X=1 for males

$Y$ = A dummy variable for the outcome where Y = 0 for an incorrect answer and Y=1 for a correct answer

$P_0$ = The relative frequency of females answering correctly

$P_1$ = The relative frequency of males answering correctly

In this report, an item on which males outperformed females would yield a positive GDI. Conversely, an item on which females outperformed males would yield a negative value for the GDI. For further information on the Gender Difference Index, refer to Wagemaker (1996), Appendix L.

## Estimating Sampling Error

Because the statistics presented in this report are national estimates based on samples of schools and students rather than the values that could be calculated if every school and student in a country answered every question, it is important to have measures of the degree of uncertainty of the estimates. The jackknife procedure was used to estimate the standard error associated with each statistic presented in this report.[18] The use of confidence intervals, based on the standard errors, allows inferences to be made about the population means and proportions in a manner that reflects the uncertainty associated with the sample estimates. An estimated sample statistic plus or minus two standard errors represents a 95% confidence interval for the corresponding population result.

[17] Taube, K. and Munck, I. (1996). "Gender Differences at the Item Level" in H. Wagemaker (Ed.), *Are Girls Better Readers?: Gender Differences in Reading Literacy in 32 Countries*. Delft, Amsterdam: Eburon Publishers.

[18] The jackknife repeated replication technique for estimating sampling errors is documented in Gonzalez, E.J. and Foy, P. (1997). "Estimation of Sampling Variability, Design Effects, and Effective Sample Sizes" in M.O. Martin and D.L. Kelly (Eds.), *Third International Mathematics and Science Study Technical Report, Volume II*. Chestnut Hill, MA: Boston College.

## Exhibit A.1 School and Student Sample Sizes by Gender
### Fourth Grade

| Country | Number of Schools | Number of Students | Number of Males | Number of Females |
|---|---|---|---|---|
| *Australia* | 177 | 6492 | 3240 | 3252 |
| *Austria* | 132 | 2603 | 1341 | 1262 |
| **Canada** | 389 | 8235 | 4172 | 4063 |
| **Cyprus** | 146 | 3362 | 1705 | 1657 |
| **Czech Republic** | 187 | 3268 | 1561 | 1707 |
| **England** | 127 | 3126 | 1544 | 1582 |
| **Hong Kong** | 124 | 4388 | 2375 | 2013 |
| *Hungary* | 149 | 2936 | 1474 | 1462 |
| **Iceland** | 144 | 1809 | 880 | 929 |
| **Iran, Islamic Rep.** | 180 | 3385 | 1730 | 1655 |
| **Ireland** | 165 | 2873 | 1452 | 1421 |
| **Japan** | 142 | 4306 | 2153 | 2153 |
| **Korea** | 150 | 2812 | 1424 | 1388 |
| *Latvia (LSS)* | 125 | 2216 | 1128 | 1088 |
| *Netherlands* | 129 | 2496 | 1258 | 1238 |
| **New Zealand** | 149 | 2421 | 1183 | 1238 |
| **Norway** | 134 | 2192 | 1167 | 1025 |
| **Portugal** | 148 | 2852 | 1459 | 1393 |
| **Scotland** | 152 | 3290 | 1651 | 1639 |
| **Singapore** | 191 | 7133 | 3750 | 3383 |
| *Slovenia* | 120 | 2540 | 1258 | 1282 |
| **United States** | 182 | 7296 | 3547 | 3749 |

SOURCE: IEA Third International Mathematics and Science Study (TIMSS), 1994-95.

Countries appearing in italics either did not satisfy guidelines for sample participation rates or did not meet age/grade specifications.

## Exhibit A.2 School and Student Sample Sizes by Gender
### Eighth Grade

| Country | Number of Schools | Number of Students | Number of Males | Number of Females |
|---|---|---|---|---|
| *Australia* | 161 | 7251 | 3529 | 3722 |
| *Austria* | 123 | 2706 | 1385 | 1321 |
| **Belgium (Fl)** | 141 | 2894 | 1457 | 1437 |
| *Belgium (Fr)* | 119 | 2560 | 1269 | 1291 |
| **Canada** | 364 | 8225 | 4137 | 4088 |
| *Colombia* | 140 | 2623 | 1240 | 1383 |
| **Cyprus** | 55 | 2918 | 1494 | 1424 |
| **Czech Republic** | 149 | 3327 | 1690 | 1637 |
| **England** | 121 | 1776 | 923 | 853 |
| **France** | 124 | 2879 | 1449 | 1430 |
| *Germany* | 135 | 2833 | 1410 | 1423 |
| **Hong Kong** | 85 | 3337 | 1829 | 1508 |
| **Hungary** | 150 | 2912 | 1423 | 1489 |
| **Iceland** | 129 | 1773 | 905 | 868 |
| **Iran, Islamic Rep.** | 191 | 3680 | 2043 | 1637 |
| **Ireland** | 132 | 3076 | 1541 | 1535 |
| **Japan** | 151 | 5141 | 2646 | 2495 |
| **Korea** | 150 | 2920 | 1585 | 1335 |
| **Latvia (LSS)** | 141 | 2407 | 1148 | 1259 |
| **Lithuania** | 145 | 2525 | 1140 | 1385 |
| *Netherlands* | 94 | 1957 | 980 | 977 |
| **New Zealand** | 149 | 3683 | 1908 | 1775 |
| **Norway** | 146 | 3267 | 1633 | 1634 |
| **Portugal** | 142 | 3391 | 1728 | 1663 |
| *Romania* | 163 | 3723 | 1809 | 1914 |
| **Russian Federation** | 174 | 4022 | 1871 | 2151 |
| *Scotland* | 124 | 2815 | 1457 | 1358 |
| **Singapore** | 137 | 4641 | 2334 | 2307 |
| **Slovak Republic** | 145 | 3501 | 1716 | 1785 |
| *Slovenia* | 121 | 2705 | 1324 | 1381 |
| **Spain** | 153 | 3855 | 1848 | 2007 |
| **Sweden** | 116 | 4063 | 2084 | 1979 |
| **Switzerland** | 250 | 4854 | 2443 | 2411 |
| **United States** | 183 | 7087 | 3526 | 3561 |

SOURCE: IEA Third International Mathematics and Science Study (TIMSS), 1994-95.

Countries appearing in italics either did not satisfy guidelines for sample participation rates or did not meet age/grade specifications.

## Exhibit A.3 School and Student Sample Sizes by Gender - Mathematics and Science Literacy
### Final Year of Secondary School

| Country | Number of Schools | Number of Students | Number of Males | Number of Females |
|---|---|---|---|---|
| *Australia* | 87 | 1941 | 775 | 1166 |
| *Austria* | 169 | 1919 | 878 | 1041 |
| *Canada* | 337 | 5205 | 2672 | 2533 |
| **Cyprus** | 28 | 531 | 251 | 280 |
| **Czech Republic** | 150 | 2167 | 1115 | 1052 |
| *France* | 56 | 1572 | 813 | 759 |
| *Germany* | 150 | 2179 | 1071 | 1108 |
| **Hungary** | 204 | 4912 | 2370 | 2542 |
| *Iceland* | 30 | 1687 | 800 | 887 |
| **Lithuania** | 142 | 2886 | 948 | 1938 |
| *Netherlands* | 79 | 1470 | 745 | 725 |
| **New Zealand** | 79 | 1763 | 852 | 911 |
| *Norway* | 131 | 2518 | 1190 | 1328 |
| **Russian Federation** | 163 | 2289 | 841 | 1448 |
| *Slovenia* | 78 | 1563 | 828 | 735 |
| **Sweden** | 145 | 3068 | 1462 | 1606 |
| **Switzerland** | 383 | 3283 | 1660 | 1623 |
| *United States* | 211 | 5807 | 2839 | 2968 |

SOURCE: IEA Third International Mathematics and Science Study (TIMSS), 1994-95.

Countries appearing in italics either did not satisfy guidelines for sample participation rates or did not meet age/grade specifications.

## Exhibit A.4 School and Student Sample Sizes by Gender - Advanced Mathematics
### Final Year of Secondary School

| Country | Number of Schools | Number of Students | Number of Males | Number of Females |
|---|---|---|---|---|
| *Australia* | 83 | 645 | 360 | 285 |
| *Austria* | 114 | 763 | 299 | 464 |
| **Canada** | 309 | 2772 | 1474 | 1298 |
| **Cyprus** | 21 | 388 | 237 | 151 |
| **Czech Republic** | 90 | 1101 | 451 | 650 |
| **France** | 61 | 1055 | 665 | 390 |
| **Germany** | 76 | 2250 | 832 | 1418 |
| **Lithuania** | 29 | 734 | 372 | 362 |
| **Russian Federation** | 113 | 1638 | 908 | 730 |
| *Slovenia* | 72 | 1521 | 746 | 775 |
| **Sweden** | 101 | 1001 | 644 | 357 |
| **Switzerland** | 197 | 1389 | 766 | 623 |
| *United States* | 199 | 2785 | 1417 | 1368 |

SOURCE: IEA Third International Mathematics and Science Study (TIMSS), 1994-95.

Countries appearing in italics either did not satisfy guidelines for sample participation rates or did not meet age/grade specifications.

## Exhibit A.5 School and Student Sample Sizes by Gender - Physics
### Final Year of Secondary School

| Country | Number of Schools | Number of Students | Number of Males | Number of Females |
|---|---|---|---|---|
| *Australia* | 85 | 661 | 417 | 244 |
| *Austria* | 114 | 763 | 306 | 457 |
| **Canada** | 306 | 2353 | 1440 | 913 |
| **Cyprus** | 21 | 367 | 230 | 137 |
| **Czech Republic** | 90 | 1087 | 426 | 661 |
| **France** | 61 | 1087 | 670 | 417 |
| **Germany** | 74 | 709 | 487 | 222 |
| **Norway** | 66 | 1048 | 781 | 267 |
| **Russian Federation** | 91 | 1259 | 724 | 535 |
| *Slovenia* | 51 | 726 | 566 | 160 |
| **Sweden** | 101 | 1012 | 651 | 361 |
| **Switzerland** | 197 | 1359 | 727 | 632 |
| *United States* | 203 | 3114 | 1617 | 1497 |

Countries appearing in italics either did not satisfy guidelines for sample participation rates or did not meet age/grade specifications.

## Exhibit A.6 Grades Tested in TIMSS

| Country | Upper Grade - Population 1 | | Upper Grade - Population 2 | |
|---|---|---|---|---|
| | Country's Name for Upper Grade | Years of Formal Schooling Including Upper Grade[1] | Country's Name for Upper Grade | Years of Formal Schooling Including Upper Grade[1] |
| [2] Australia | 4 or 5 | 4 or 5 | 8 or 9 | 8 or 9 |
| Austria | 4 | 4 | 4. Klasse | 8 |
| Belgium (Fl) | | | 2A & 2P | 8 |
| Belgium (Fr) | | | 2A & 2P | 8 |
| Bulgaria | | | 8 | 8 |
| Canada | 4 | 4 | 8 | 8 |
| Colombia | | | 8 | 8 |
| [3] Cyprus | 4 | 4 | 8 | 8 |
| Czech Republic | 4 | 4 | 8 | 8 |
| Denmark | | | 7 | 7 |
| England | Year 5 | 5 | Year 9 | 9 |
| France | | | 4ème (90%) or 4ème Technologique (10%) | 8 |
| Germany | | | 8 | 8 |
| Greece | 4 | 4 | Secondary 2 | 8 |
| Hong Kong | Primary 4 | 4 | Secondary 2 | 8 |
| Hungary | 4 | 4 | 8 | 8 |
| Iceland | 4 | 4 | 8 | 8 |
| Iran, Islamic Rep. | 4 | 4 | 8 | 8 |
| Ireland | 4th Class | 4 | 2nd Year | 8 |
| Israel | 4 | 4 | 8 | 8 |
| Japan | 4 | 4 | 2nd Grade Lower Secondary | 8 |
| Korea | 4th Grade | 4 | 2nd Grade Middle School | 8 |
| Kuwait | 5 | 5 | 9 | 9 |
| Latvia | 4 | 4 | 8 | 8 |
| Lithuania | | | 8 | 8 |
| [4] Netherlands | 6 | 4 | Secondary 2 | 8 |
| [3,5] New Zealand | Standard 3 | 4.5–5.5 | Form 3 | 8.5 - 9.5 |
| [3] Norway | 3 | 3 | 7 | 7 |
| [3] Philippines | | | 1st Year High School | 7 |
| Portugal | 4 | 4 | Grade 8 | 8 |
| Romania | | | 8 | 8 |
| [6] Russian Federation | | | 8 | 7 or 8 |
| Scotland | Year 5 | 5 | Secondary 2 | 9 |
| Singapore | Primary 4 | 4 | Secondary 2 | 8 |
| Slovak Republic | | | 8 | 8 |
| Slovenia | 4 | 4 | 8 | 8 |
| Spain | | | 8 EGB | 8 |
| [3] South Africa | | | Standard 6 | 8 |
| [3] Sweden | | | 7 | 7 |
| [3] Switzerland | | | 0 | 0 |
| (German) | | | 7 | 7 |
| (French and Italian) | | | 8 | 8 |
| Thailand | Primary 4 | 4 | Secondary 2 | 8 |
| United States | 4 | 4 | 8 | 8 |

SOURCE: IEA Third International Mathematics and Science Study (TIMSS), 1994-95.

1 Years of schooling based on the number of years children in the grade level have been in formal schooling, beginning with primary education (International Standard Classification of Education Level 1). Does not include preprimary education.

2 Australia: Each state/territory has its own policy regarding age of entry to primary school. In 4 of the 8 states/territories students were sampled from grades 3 and 4; in the other four states/territories students were sampled from grades 4 and 5. In 4 of the 8 states/territories students were sampled from grades 7 and 8; in the other four states/territories students were sampled from grades 8 and 9.

3 Indicates that there is a system-split between the lower and upper grades. In Cyprus, system-split occurs only in the large or city schools. In Switzerland there is a system-split in 14 of 26 cantons.

4 In the Netherlands kindergarten is integrated with primary education. Grade-counting starts at age 4 (formerly kindergarten 1). Formal schooling in reading, writing, and arithmetic starts in grade 3, age 6.

5 New Zealand: The majority of students begin primary school on or near their 5th birthday so the "years of formal schooling" vary.

6 Russian Federation: 70% of students in the seventh grade have had 6 years of formal schooling; 70% in the eighth grade have had 7 years of formal schooling.