IEA
TIMSS
LONGITUDINAL
2023 — 2024

TIMSS 2023 Longitudinal

Technical Report

Addendum to the TIMSS 2023 Technical Report

Editors:
Bethany Fishbein
Eugenio Gonzalez
Matthias von Davier

Co-funded by the European Union

IEA
TIMSS & PIRLS
BOSTON COLLEGE

TIMSS 2023 Longitudinal Technical Report (Addendum to the TIMSS 2023 Technical Report)

Editors: Bethany Fishbein, Eugenio Gonzalez, and Matthias von Davier

With contributions from:
Darryl Cooney
Susan Flicop
Dihao Leng
Tahira Qadri
Katherine Reynolds
Liqun Yin

Suggested citation:

Fishbein, B., Gonzalez, E., & von Davier, M. (Eds.). (2025). *TIMSS 2023 Longitudinal Technical Report (Addendum to the TIMSS 2023 Technical Report)*. Boston College, TIMSS & PIRLS International Study Center. https://doi.org/10.6017/lse.tpisc.timss.de4812

Publishers:
TIMSS & PIRLS International Study Center, Lynch School of Education and Human Development, Boston College, and International Association for the Evaluation of Educational Achievement (IEA)

For more information about TIMSS or TIMSS Longitudinal, contact:
IEA's TIMSS & PIRLS International Study Center
Lynch School of Education and Human Development
Boston College
Chestnut Hill, MA 02467
United States
tel: +1-617-552-1600
e-mail: timssandpirls@bc.edu
timssandpirls.bc.edu

Contents

Introduction

The TIMSS 2023 Longitudinal Study is an extension of TIMSS 2023 that combines results from the TIMSS 2023 data collection with a second data collection from the same schools and students in the following school year. The results from these two data collections report on students' mathematics and science achievement and related contextual factors at two time points: once in 2023 and again in 2024. National samples assessed as part of TIMSS 2023 served as the first data collection for fourth- and eighth-grade students in 2023, and the second data collection took place in 2024 when most of these students were toward the end of their fifth- and ninth-grade school years. Collecting these data enables analyses to explore how growth in achievement varies within and across education systems, among different subgroups of students, and in relation to students' attitudes, engagement, and school experiences. Nine countries participated in TIMSS 2023 Longitudinal for Grades 4 and 5, and three countries for Grades 8 and 9. Descriptive results from the study are presented and discussed in the *TIMSS 2023 Longitudinal International Results in Mathematics and Science* report.

Building upon the foundation of TIMSS 2023, the TIMSS 2023 Longitudinal Study adhered to the same quality standards and principles for the methods and procedures used to develop, implement, and report the TIMSS 2023 results. However, assessing the same students a second time in 2024 required adapting some procedures and developing new ones to ensure the quality and comparability of the results. While the *TIMSS 2023 Technical Report* (von Davier et al., 2024) provides complete documentation of the TIMSS 2023 methods and procedures, this addendum describes the methods and procedures implemented for the 2024 data collection and for the joint analysis of the results from both years conducted for the longitudinal study. For a complete picture of the methods and procedures, readers should consult the TIMSS 2023 report. References are provided to specific chapters where relevant.

The data collection in 2024 used the same mathematics and science achievement instruments as TIMSS 2023; however, booklets were assigned to students under a carefully designed procedure so that students would not be administered the same items twice, and so the difficulty of the items was aligned with students' anticipated learning gains from the previous school year. Students and their parents, teachers, and school principals completed revised versions of the TIMSS 2023 Context Questionnaires in 2024. These contained some overlapping items with the 2023 instruments to allow for examining change in select contextual constructs, as well as some new items to examine constructs of interest in relation to learning gains. Section 1 of this report describes the design for administering achievement instruments, and Section 2 provides information about the questionnaire revision process.

As in TIMSS 2023, the longitudinal 2024 data collection followed rigorous standardized survey operation procedures with close attention to quality control throughout. Participating countries were provided with targeted guidance to ensure that students from the original TIMSS 2023 samples could be tracked reliably over time and that operational procedures were consistently followed in both data collections. Section 3 of this report provides an overview of the survey operations for the 2024 data collection and documents the feedback collected from each country's National Research Coordinator (NRC) on the 2024 survey activities, and Section 4 documents observations from the International Quality Control Program on the 2024 data collection.

Section 5 describes the implementation of sampling procedures for the 2024 data collection and documents outcomes for the TIMSS 2023 Longitudinal Study samples. The success of the longitudinal study rested on participating countries' ability to find the students sampled for TIMSS 2023 again in 2024 and reassess as many as possible, including those who were absent in 2023 and any students who may have repeated or skipped grade levels. TIMSS 2023 Longitudinal adhered to the same quality standards for sampling implementation as TIMSS 2023, while considering that some students originally sampled for TIMSS 2023 may no longer be accessible in 2024.

To provide high-quality measures of students' growth in achievement over time, psychometric analyses were conducted to estimate mathematics and science achievement for both 2023 and 2024 data collections on the same scale, capitalizing on the information collected from students at both time points. Section 6 of this report describes the analyses and related quality control, and Section 7 provides information on variance estimation of the achievement estimates. Section 8 documents the creation of context questionnaire scales based on the data collected in 2024.

Acknowledgments

The TIMSS 2023 Longitudinal Study was a collaborative effort involving numerous organizations and individuals worldwide, including the TIMSS & PIRLS International Study Center, IEA Amsterdam, IEA Hamburg, RTI International, RM Results, and the National Research Coordinators and their teams in the participating countries. Staff from the TIMSS & PIRLS International Study Center, IEA Amsterdam, and RTI contributed to content within this report. A full list of organizations and individuals involved in the study is provided in Appendix D of the *TIMSS 2023 Longitudinal International Results* report.

SECTION 1

Mathematics and Science Achievement Instruments

TIMSS 2023 Longitudinal administered the same pool of digital assessment booklets at both time points, following a carefully designed protocol to ensure that individual students were not administered the same assessment material twice. Chapter 1 of the *TIMSS 2023 Technical Report* describes the development process TIMSS follows each cycle for the mathematics and science assessments to measure trends.

The digital assessments were organized in blocks of mathematics and science items assembled into booklets. These were administered to students according to a group-adaptive design, which optimizes measurement across the range of abilities found in each of the TIMSS participating countries. The group-adaptive assessment included 14 booklets at each grade level, with seven "more difficult" and seven "less difficult." Under the group-adaptive design, each country administers the entire assessment pool at each time point; however, the proportions of more and less difficult booklets vary according to the mathematics and science ability levels of students in the country.

For the 2024 data collection, students received a booklet of the same difficulty level they received in 2023, ensuring they were not administered the same assessment material twice. The assignment scheme aimed for students not to receive less difficult content in 2024 compared to 2023, in anticipation of learning gains over time. Exhibit 1 illustrates the 2024 booklet assignment based on the 2023 booklet assignments.

Exhibit 2 lists the proportion of more difficult and less difficult booklets administered in the TIMSS 2023 Longitudinal countries. The same rotations were used in each country for both administrations. Most countries administered an equal balance of more and less difficult booklets (50/50). Jordan and Kosovo administered a greater proportion of less difficult booklets (30% more difficult, 70% less difficult), and Korea administered a greater proportion of more difficult booklets (70% more difficult, 30% less difficult).

Exhibit 1: TIMSS 2023 Longitudinal Group-Adaptive Booklet Assignments

Booklet Difficulty	2023 Booklet Assignment	2024 Booklet Assignment
More Difficult Booklets	Booklet 1	Booklet 3
	Booklet 2	Booklet 4
	Booklet 3	Booklet 5
	Booklet 4	Booklet 6
	Booklet 5	Booklet 7
	Booklet 6	Booklet 1
	Booklet 7	Booklet 2
Less Difficult Booklets	Booklet 8	Booklet 10
	Booklet 9	Booklet 11
	Booklet 10	Booklet 12
	Booklet 11	Booklet 13
	Booklet 12	Booklet 14
	Booklet 13	Booklet 8
	Booklet 14	Booklet 9

Exhibit 2: TIMSS 2023 Longitudinal Group-Adaptive Booklet Rotations (More Difficult / Less Difficult)

Country	Grades 4 & 5	Grades 8 & 9
Georgia	50/50	
Italy	50/50	
Jordan	30/70	30/70
Korea, Rep. of	70/30	70/30
Kosovo	30/70	
Montenegro	50/50	
North Macedonia	50/50	
Slovenia	50/50	
Sweden	50/50	50/50

SECTION 2

Context Questionnaires

An important purpose of the TIMSS 2023 Longitudinal Study is to facilitate analyses that explore how growth in achievement varies in relation to students' attributes and their attitudes, engagement, and school experiences. Participating students, their parents, teachers, and school principals completed context questionnaires in both administrations to collect information about contexts related to student learning. In addition, NRCs from the participating countries completed curriculum questionnaires, gathering information about national curriculum policies and practices related to their educational systems and mathematics and science curricula. Chapter 2 of the *TIMSS 2023 Technical Report* describes the process for developing the context questionnaires administered in TIMSS 2023.

The context questionnaires administered in 2023 were revised by a special Questionnaire Item Review Committee (QIRC) prior to their administration in 2024. The revision process for the 2024 longitudinal questionnaires involved the following:

- Removing items that did not need to be administered twice (e.g., information that should not change, such as student participation in preprimary education)
- Revising items to focus on the grade level assessed in 2024
- Adding new items to allow examination of new topics and constructs relevant to the longitudinal study

Staff at the TIMSS & PIRLS International Study Center proposed initial revisions for review by the QIRC and NRCs. The questionnaires were revised based on feedback from these groups. After finalizing the international questionnaire versions, the same procedures for preparing national questionnaire instruments in TIMSS 2023 were applied to the revised questionnaires used in 2024 (see Chapter 5 of the *TIMSS 2023 Technical Report*).

Revisions to TIMSS 2023 Context Questionnaires for the 2024 Administration

This section summarizes the major changes made to the TIMSS 2023 Context Questionnaires for the 2024 data collection. The summary of changes below does not provide a list of item-by-item edits; instead, the questionnaires administered are available to view in their entirety at the TIMSS 2023 Longitudinal Context Questionnaires webpage.

A matching table for common and unique items is provided in the TIMSS 2023 Longitudinal International Database documentation.

Student Questionnaire

Items/Topics Retained from 2023

- Demographic information
- Home resources
- Level of education completed by parents/guardians
- Expectations for future education
- Experiences at school
- Digital self-efficacy
- Attitudes toward mathematics and science
- Perceptions and experiences of instruction
- Environmentally responsible behaviors and attitudes

Items/Topics Added or Revised for 2024

- Student well-being
- Mathematics and science homework

Items/Topics Removed for 2024

- Student and parent country of birth

Home Questionnaire

Items/Topics Retained from 2023

- Home resources
- Language(s) spoken at home
- Expectations for child's future education
- Perception of child's school
- Encouragement of environmentally responsible behaviors

Items/Topics Added or Revised for 2024

- Activities and engagement with child's learning and school

Items/Topics Removed for 2024

- Student early literacy and numeracy development
- Student country of birth
- Language(s) spoken by student before starting school
- Student participation in pre-primary education

Teacher Questionnaire

Items/Topics Retained from 2023

- Teacher background and education
- School emphasis on academic success
- Job satisfaction
- Students limiting instruction
- Instructional emphases, strategies, and tools
- Class enrollment
- Instruction time
- Topics taught to students
- Assessment practices
- Participation in professional development

Items/Topics Added or Revised for 2024

- Confidence in teaching

Items/Topics Removed from 2023

- School safety
- Challenges in teaching
- Challenges in using digital devices

School Questionnaire

Items/Topics Retained from 2023

- Enrollment of students in target grade
- Socioeconomic composition of student body
- Percentage of students with language of test as native language
- Instruction time
- Instructional resources
- School emphasis on academic success
- School emphasis on mathematics and science
- School emphasis on environmental sustainability
- Principal experience and education

Items/Topics Added or Revised for 2024

- School innovation
- Computer access
- Enrichment and remediation programs

Items/Topics Removed from 2023

- Total enrollment of students in school
- School geographic location

- Student access to library or media center at school
- School discipline and safety
- Student readiness for instruction
- COVID-19 pandemic

Curriculum Questionnaire

Items/Topics Retained from 2023

- Instructional time in mathematics and science curricula
- Types of materials included in mathematics and science curricula
- Policies on use of digital devices in mathematics and science curricula
- TIMSS mathematics and science topics covered by mathematics and science curricula

Items/Topics Added or Revised for 2024

- Organization of students for transitioning between Grades 4/5 and Grades 8/9

Items/Topics Removed from 2023

- Structure of education system
- Language(s) of instruction
- Early childhood education
- National examinations
- National initiatives related to environmental sustainability
- National initiatives related to social-emotional learning
- Teacher preparation
- Principal preparation
- Policy related to COVID-19 pandemic
- Evaluation procedures for mathematics and science curricula

SECTION 3

Survey Operations Procedures

The TIMSS 2023 Survey Operations Procedures, described in Chapter 4 of the *TIMSS 2023 Technical Report*, were the basis for the procedures used for the 2024 data collection. NRCs were provided with a Survey Operations Procedures Unit for TIMSS 2023 Longitudinal, information on sampling preparations and student tracking, and guidelines on preparing and administering the revised context questionnaires for 2024. Additional training sessions were provided for countries on using tracking forms and software, scoring constructed-response items, and data management.

For the test administration in 2024, participating countries were instructed to use the same Test Administrator Manual and School Coordinator Manual from TIMSS 2023, but updated to include a reference to students having already been assessed in 2023. As in TIMSS 2023, national and international quality control programs were conducted to document the quality of the test administration in 2024. Scoring and submitting databases and documentation followed the same procedures as TIMSS 2023.

Tracking Students Over Time

Conducting data collection in 2024 required participating countries to recontact the students sampled for TIMSS 2023 in 2024 and reassess as many as possible, including those who were absent in 2023 and any students who may have repeated or skipped grade levels. Countries were provided guidelines and targeted support to locate and document the participation of all students sampled for TIMSS 2023. Most of these students were found in the same schools, often in the same classrooms, but the situation varied across countries. In some cases, countries had to contact students who were spread across classrooms or even across grades. This required extra effort, but it was the exception rather than the norm across the participating countries.

For the 2024 data collection, countries used tracking forms together with IEA's Within-School Sampling Software (WinW3S) to help track the participation of students. The tracking forms completed by School Coordinators also documented whether students moved between grades, classes, or schools, as well as their mathematics and science teachers in each school year. The tracking information was used to quantify the level of attrition in each country and evaluate the quality of the samples. Section 5 of this report provides information about the implementation of the student tracking procedures and participation.

Linking Students and Teachers

As in TIMSS 2023, the 2024 administration collected questionnaire data from all mathematics and science teachers of the participating students during the school year in which the 2024 data collection occurred. In some cases, students had the same teachers in both years, but this was not always the case. Because of the variation across countries in how teachers are linked to students, the TIMSS 2023 Longitudinal International Database includes the 2023 teachers and 2024 teachers as separate records. These are meant to be analyzed separately.

Survey Activities Questionnaire

Similar to the TIMSS 2023 data collection, NRCs were asked to complete a Survey Activities Questionnaire seeking feedback on all aspects of their experience conducting the 2024 data collection. The feedback solicited in the questionnaire included evaluating the administration of the assessment materials and the effectiveness of the operations procedures and documentation.

The Survey Activities Questionnaire for the 2024 data collection was adapted from the one used in TIMSS 2023. It included seven sections focusing on the following aspects of the TIMSS survey operations procedures:

- Instructions and training
- Obtaining participation and tracking students
- Preparing the assessment instruments
- Administering the assessments
- Implementing the National Quality Control Program
- Preparing for and scoring the constructed-response items
- Creating and submitting the databases and documentation

Most items in the Survey Activities Questionnaire included accompanying comment fields, in which NRCs were encouraged to explain their responses, provide additional information, or suggest improvements for the process.

This section summarizes select results from the Survey Activities Questionnaire, based on the 2024 data collection, as reported in Appendix A. Responses were provided by nine NRCs from the nine participating countries, three of which participated at both grade levels. Results from the 2023 data collection and more information about TIMSS Survey Operations Procedures referenced below are presented in Chapter 4 of the *TIMSS 2023 Technical Report*.

↓ Appendix A: TIMSS 2023 Longitudinal Survey Activities Questionnaire Results for the 2024 Data Collection

Instructions and Training

The first section of the questionnaire asked NRCs about the Survey Operations and Procedures Unit, which was created specifically for TIMSS 2023 Longitudinal. Seven NRCs considered the Longitudinal Survey Operations and Procedures Unit to be clear and sufficient, and all nine felt that the training for tracking students and using the WinW3S software was clear and sufficient. Eight countries considered the virtual scoring training to be sufficient, and seven felt the data management training was also clear and sufficient. Most countries used the TIMSS 2023 manuals, as the procedures were the same, except for tracking the students sampled for TIMSS 2023 over the next school year for the 2024 data collection.

Obtaining Participation and Tracking Students

The second section asked NRCs about obtaining cooperation from the schools as well as tracking students who participated in TIMSS 2023. Countries were encouraged to inform schools selected for TIMSS 2023 about both data collections as early as possible. Three countries utilized the sample invitations provided by the TIMSS & PIRLS International Study Center to secure school cooperation for the 2024 data collection. Only one NRC reported difficulties in adapting the School Coordinator manual and adjusting the WinW3S software for the longitudinal study. Almost all countries reported that 90–100% of students assessed in 2023 were in the same schools in 2024. Students who were no longer in the same schools were usually not accessible and therefore did not participate in the 2024 data collection. Section 5 of this report provides more information about the implementation of the student tracking procedures and participation.

Preparing the Assessment Instruments

This section focused on preparing assessment instruments for the 2024 data collection, which included preparing national context questionnaires and preparing digital devices for test administration. The same procedures for preparing national questionnaire instruments in TIMSS 2023 were applied to the revised questionnaires for 2024 (see Chapter 5 of the *TIMSS 2023 Technical Report*).

Countries that administered the home questionnaire on paper reported no major problems with using the InDesign software needed to prepare the materials, and all NRCs corrected their materials as requested by the verifiers throughout the process. The only country that did not implement quality control checks for the paper home questionnaire used an external printing company that conducted its own checks, with which the country reportedly worked for 25 years. All other countries utilizing the paper home questionnaire checked a sample of the materials.

All countries that administered the achievement test using the USB method implemented quality control checks on the USBs before distributing them to schools. Some tested all the USBs, while others verified the file sizes or tested only a subset of them.

Countries are always asked to maintain the security of the test materials and report any security breaches. The only security breach noted concerned the teacher questionnaire and the ability of a teacher to enter a different subject code than the one assigned. Finally, one country reported difficulty fixing some layout issues using the IEA's Online Survey System (OSS) for the online questionnaires.

Administering the Assessment

This section addressed the administration of the assessment in 2024, including the extent to which NRCs were notified about errors in the testing materials sent to schools. Very few errors in the materials were reported, and issues were resolved in almost all cases. One country noted that when the USB sticks did not work, schools used the spare ones that were provided.

Eight countries appointed School Coordinators from within the participating schools, and all but two countries used the same School Coordinators as they did for TIMSS 2023. Five of the countries used the same Test Administrators from the TIMSS 2023 test administration, and almost all provided additional training by the national center staff. Six of the nine NRCs reported that at least one additional person helped the Test Administrators during the testing sessions. This was sometimes the classroom teacher, School Coordinator, or an information technology (IT) consultant/expert who assisted with computer setup and troubleshot any technical issues that occurred during the testing session.

Three NRCs reported that technical issues required stopping or postponing testing sessions for a whole class. NRCs noted electrical outages that affected entire classes or schools, as well as other technical issues (such as problems with logging in).

Five of the nine reporting countries had enough computers or tablets for each school to test all students in the selected class(es) simultaneously. Where this was not the case, schools held multiple sessions with two to seven sessions per school.

National Quality Control Program

This section described the National Quality Control Program that each country implemented during the 2024 data collection. The national program was in addition to the International Quality Control Program visits, documented in Section 4 of this report.

As part of national quality assurance activities, NRCs were required to send National Quality Control Monitors to 10% of the participating schools to observe the test

administration and to document compliance with the prescribed procedures. They were told to follow the same program procedures used in TIMSS 2023. Three countries did not conduct this program for several reasons, including having Test Administrators from the national centers who provided daily feedback and other quality assurance reviews. The quality monitors in participating countries visited an average of more than 14% of participating schools and reported following the same program for the longitudinal study as they did for the 2023 data collection. No NRCs reported that there were any major problems or special circumstances that occurred frequently.

Preparing for and Scoring the Constructed-Response Items

This section covered the preparation for and scoring of the constructed-response items in 2024, which followed procedures from TIMSS 2023 (see Chapter 7 of the *TIMSS 2023 Technical Report*). All but one country reported using the same materials for training the scorers in 2024 that were used for training the TIMSS 2023 scorers. The NRC reporting changes explained that some clarifications were provided based on experience from 2023. No country reported experiencing any significant problems when scoring the constructed-response items. Two countries used only experienced scorers, one used only new scorers, and the rest used a combination of experienced and new scorers. Six of the countries took two weeks to complete the scoring, with the others using three or four weeks.

Creating and Submitting the Databases and Documentation

The last section of the TIMSS 2023 Longitudinal Survey Activities Questionnaire addressed the use of IEA's WinW3S and Data Management Expert (DME) software as well as the data quality checks when submitting the databases and documentation. Two NRCs reported issues using WinW3S, which were resolved with the help of staff at IEA Hamburg. No one reported major problems when using the DME software. All NRCs applied the required data quality checks, with all issues resolved by the country or with support from IEA Hamburg. Chapter 8 of the *TIMSS 2023 Technical Report* provides information about data processing and verification in TIMSS.

SECTION 4

International Quality Assurance Observations

As for the TIMSS 2023 data collection, International Quality Control Monitors (IQCMs) conducted quality observations of a sample of the 2024 longitudinal test administration sessions for the TIMSS International Quality Assurance Program. In each of the nine participating countries, one IQCM was recruited and trained who had served in the same role for 2023. Some hired assistants to support them in their activities. In total across the nine participating counties, 179 test sessions were observed for the 2024 longitudinal data collection, including 134 fifth-grade sessions and 45 ninth-grade sessions.

Depending on when the 2024 data collection took place in the country, IQCMs participated in one of two training sessions offered in February and September 2024. As all IQCMs had also served in the same role in TIMSS 2023, the training focused on reviewing established procedures and introducing updates to procedures for the 2024 data collection. During the data collection period, to the extent possible, IQCMs conducted site visits to the same schools they had visited during the TIMSS 2023 data collection to document the administration of the test and questionnaires and to determine whether any aspects of the test administration changed from the previous year. As part of these visits, IQCMs interviewed the School Coordinator and Test Administrator to gather feedback on the TIMSS 2023 Longitudinal Study and the 2024 testing procedures. IQCMs completed a Classroom Observation Record for each visit, detailing their findings and submitting their observations.

The content of the Classroom Observation Records was essentially identical to that of the TIMSS 2023 program, with a few additions. New questions were added addressing how students from the TIMSS 2023 data collection were tracked for the 2024 data collection the following year, as well as the overall participation of originally sampled students from TIMSS 2023. These additional questions were introduced to assess the sampling and administrative procedures associated with the continuity of TIMSS 2023 to 2024.

The Classroom Observation Record was similar in structure to that of TIMSS 2023, organized into the following sections for each school visited:

- Preliminary Activities

- Administration Process
- Summary Observations
- Interview with the School Coordinator and Test Administrator

This section summarizes select results of the TIMSS 2023 Longitudinal International Quality Assurance program for the 2024 data collection, highlighting unique aspects compared to TIMSS 2023. Frequencies of IQCM responses for all activities are reported in Appendix B. The "Not applicable" category includes instances in which a question was not displayed due to logical filtering.

Results from the 2023 data collection are presented in Chapter 6 of the *TIMSS 2023 Technical Report*.

↓ Appendix B: TIMSS 2023 Longitudinal International Quality Assurance Program Results for the 2024 Data Collection

Preliminary Activities

The first section addressed activities that took place before the testing sessions. Overall, IQCMs reported that the 2024 data collections were conducted in accordance with the international procedures. Results were similar to those for the TIMSS 2023 data collection.

Administration Process

This section reported on the activities conducted during the testing sessions, including administration of the achievement test and the student questionnaire. IQCMs reported that in more than 75% of observations, the Test Administrators followed the Test Administration Script exactly, with minor deviations including repeating instructions in simpler words, simplifying instructions, answering and clarifying questions about instructions, and providing more explanations of the examples. For the Student Questionnaire following the achievement test, Test Administrators followed the administration script exactly in more than 80% of sessions observed, with minor deviations to simplify and rephrase instructions in simpler words or to provide examples to clarify instructions.

Summary Observations

This section summarized observations on overall student cooperation and compliance with instructions during the assessment administration, student participation compared to TIMSS 2023, and general observations of the test sessions.

Student Cooperation

According to observations on student cooperation, it was rare for a student to refuse to take the test. This occurred in less than 2% of observations at both grades. In all cases except one, the proper procedure was followed to excuse the student from the testing session.

In over 90% of the observed test sessions, students were described as extremely or moderately orderly and cooperative across both grade levels. In the classrooms where the students were not listed as extremely orderly and cooperative, IQCMs reported that the Test Administrator made at least some effort to control the situation in 95% of the fifth-grade sessions and in all the ninth-grade sessions.

Summary Reports on Student Participation

For the schools visited, IQCMs reported on student participation in 2024 in comparison with the original sample of students in the TIMSS 2023 data collection. If a student did not participate, reasons for missing the assessment included refusal to participate, student absences, school transfers, school system exits, or grade changes. There were also some reports of new students participating in the 2024 data collection who did not take the test in the previous year as part of TIMSS 2023. In most such cases, proper procedures were followed to mark them as new students in the relevant forms to record this information. More detailed information about student participation for the full national samples in both years is provided in Section 5 of this report.

General Observations of Test Administration

Although rare, IQCMs reported that a few students experienced technical issues during the computer-based assessment. Issues were similar to those reported by IQCMs in TIMSS 2023, such as login difficulties, software malfunctions (such as players crashing, white screens, and user interface problems like unresponsive scrollbars or rulers), hardware failures (e.g., nonfunctioning USB sticks, computers shutting down, and incompatible ports), and network-related problems. However, in most cases, Test Administrators were able to resolve the issues, such as by restarting software, swapping hardware, or consulting IT staff. Many classrooms had IT personnel available to support the administration (81% of fifth-grade sessions and 93% of ninth-grade sessions), which helped ensure that students could complete their assessments despite the disruptions.

Interview with the School Coordinator and Test Administrator

School Coordinators

Almost all School Coordinators reported that the 2024 data collection in their school went "very well" or "satisfactorily" overall (99% of fifth-grade sessions and 100% of ninth-grade sessions). In addition, the School Coordinators noted that the School Coordinator Manual worked well for them, and most other school staff members had positive attitudes toward TIMSS testing.

Many of the School Coordinators interviewed in the observed schools served as a School Coordinator in the TIMSS 2023 data collection (76% of fifth-grade and 80% of ninth-grade), indicating their familiarity with the TIMSS data collection procedures and their preparedness to serve as a School Coordinator again for the administration of the TIMSS assessment in their school.

School Coordinators in the observed schools were asked to indicate how they identified students for inclusion in the 2024 data collection. They reported that this was done primarily by using student lists (names or other identifiers) provided by the NRC, which were based on participants from the 2023 data collection. These lists were reviewed and updated by school staff to reflect changes, such as students who had transferred, repeated the grade, or were newly enrolled. In most cases, the sampled classes remained the same from one year to the next, making the tracking process straightforward. Many visited schools maintained detailed and legally mandated pedagogical records, allowing School Coordinators to easily verify student status. In some countries, formal procedures and systematic record-keeping further facilitated accurate identification. In some instances, class teachers and school principals assisted in verifying and updating student information, ensuring that the final lists reflected the correct composition for the longitudinal follow-up. Overall, almost 85% of School Coordinators interviewed did not find it difficult to track sampled students for the 2024 data collection.

More than 92% of the School Coordinators interviewed across the participating countries reported that the national centers were responsive to the school's questions and concerns. As a reflection of the successful planning and implementation of the 2024 data collection, approximately 85% of the interviewed School Coordinators said that they would be willing to serve as a School Coordinator in future international assessments.

IQCMs also provided details from the School Coordinator interviews on whether participants, including students, parents, teachers, and principals, were informed about TIMSS before the administration and if any incentives were offered. As in the TIMSS 2023 data collection, IQCMs reported that participants were clearly informed about how their personal data would be used. In some cases, students and teachers also received small

incentives for their participation, such as token money for snacks, certificates of participation, chocolates, and food.

Test Administrators

Over 80% of the interviewed Test Administrators did not report experiencing any problems during the data collection. Those who did experience issues reported that most were related to technology, network connectivity, or software issues.

As a reflection of the successful planning and implementation of TIMSS 2023 Longitudinal, over 78% of the interviewed Test Administrators said that they would be willing to serve as Test Administrators in future international assessments.

SECTION 5

Sampling Implementation

The target population for the 2024 data collection consisted of students who were selected and eligible to participate in the TIMSS 2023 data collection, regardless of their participation status in 2023. Chapter 3 and Chapter 9 of the *TIMSS 2023 Technical Report* provide information about the TIMSS 2023 sample design and implementation, which were the basis for the TIMSS 2023 Longitudinal samples. This section documents the outcomes of the national samples for TIMSS 2023 Longitudinal, including details of sample sizes, population coverage and exclusions, and participation rates for both 2023 and 2024 data collections, as well as attrition rates between the two administrations.

The TIMSS 2023 Longitudinal samples included all selected and eligible students who were assessed at one or both time points. Students participating in the longitudinal study may be missing data for one time point due to absence, parental refusal, lost instruments, or attrition. Students who were not found in the selected schools in 2024 were considered to be inaccessible and were not expected to participate in the 2024 data collection. This non-participation was categorized as attrition. As long as such inaccessible students participated in 2023, they were still considered to be participating in the longitudinal study overall. New students enrolled in selected classes in 2024 were not considered part of the target population.

Exhibit 3 presents the grades identified as the target grades for sampling according to the international target populations for the fourth and eighth year of schooling, as well as the grades identified as the next year higher, and the average ages of the students in the longitudinal samples at the time of each data collection. Students were part of the target sample in 2024 even if they repeated or skipped a grade level. However, in almost all cases, the same students assessed in 2023 at the fourth and eighth grades were reassessed at the fifth and ninth grades, respectively.

Exhibit 3: TIMSS 2023 Longitudinal National Grade Definitions and Ages of Students Assessed

Grades 4–5

Country	Country's Name for Fourth Year of Schooling*	Average Age at Time of Testing in 2023	Country's Name for Fifth Year of Schooling*	Average Age at Time of Testing in 2024
Georgia	Grade 4	10.2	Grade 5	11.2
Italy	Primary Grade 4	9.8	Primary Grade 5	10.7
Jordan	Grade 4	9.9	Grade 5	10.9
Korea, Rep. of	Elementary School Grade 4	10.4	Elementary School Grade 5	11.4
Kosovo	Grade 4	10.0	Grade 5	11.1
Montenegro	Grade 4	9.9	Grade 5	10.9
North Macedonia	Grade 4	9.8	Grade 5	10.9
Slovenia	Grade 4	10.0	Grade 5	10.9
Sweden	Grade 4	10.8	Grade 5	11.7

Grades 8–9

Country	Country's Name for Eighth Year of Schooling*	Average Age at Time of Testing in 2023	Country's Name for Ninth Year of Schooling*	Average Age at Time of Testing in 2024
Jordan	Grade 8	14.0	Grade 9	14.9
Korea, Rep. of	Middle School Grade 2	14.4	Middle School Grade 3	15.4
Sweden	Grade 8	14.8	Grade 9	15.8

* Countries' names for the years of formal schooling were reported by National Research Coordinators.
The target population in 2024 included all students enrolled in selected and eligible schools, regardless of grade enrolled in 2024.

TIMSS 2023 Longitudinal Quality Standards

TIMSS 2023 Longitudinal adhered to the same quality standards for sampling implementation as TIMSS 2023 (see Chapter 9 of the *TIMSS 2023 Technical Report*), but with consideration that some students originally sampled for TIMSS 2023 were no longer accessible in 2024. Based on the combined 2023 and 2024 samples, the *TIMSS 2023 Longitudinal International Results* may be annotated for population coverage, exclusions, and participation. Participation rates were calculated for each data collection and for the combined data. However, in determining whether countries met the TIMSS standards for sampling participation for the TIMSS 2023 Longitudinal Study, participation rates were calculated solely based on students present in the same schools in 2024 as in 2023, excluding inaccessible students (attrition cases).

Since the 2023 data collection set the foundation for the 2024 administration, it was imperative that countries planning to participate in TIMSS 2023 Longitudinal maintained the highest levels of participation in 2023. It was expected that the 2024 school and

classroom participation rates would be similar to those in 2023, but possibly lower if any 2023 participating schools chose not to participate in 2024. Countries were encouraged to be diligent in conducting make-up sessions in both 2023 and 2024 to ensure the highest possible participation, and to carefully track reasons for student non-participation in both years so that attrition cases could be properly identified.

Two unique situations are worth mentioning. First, for the combined longitudinal dataset, a classroom was considered to have participated in both 2023 and 2024 data collections when at least 50% of the students participated in any of the data collections. This could result in a class being classified as not participating in 2023 because fewer than 50% of the students participated, but being considered as participating in the combined longitudinal dataset if participation in 2024 surpassed 50%. Second, if an entire school refused to participate in 2024, all students in the school were classified as non-participating (and not inaccessible), regardless of their possible attrition status. This happened in a few schools that initially agreed to participate and noted some students were no longer at the school and inaccessible, but the school as a whole did not participate in the data collection.

Population Coverage and Exclusions

Exhibit 4 summarizes population coverage and exclusions for the combined TIMSS 2023 Longitudinal samples. National coverage of the international target population and exclusion rates were the same as documented for TIMSS 2023. Any student identified for exclusion in TIMSS 2023 was considered excluded for the 2024 data collection. Some students may have been identified as excluded in 2024 but not in 2023; therefore, their 2023 data were retained, and they were treated as eligible for the combined sample.

Exhibit 4: Coverage of the TIMSS 2023 Longitudinal Target Populations

Grades 4–5

Country	International Target Population		Exclusions from National Target Population		
	Coverage	Notes on Coverage	School-Level Exclusions	Within-Sample Exclusions	Overall Exclusions
[1] Georgia	93%	Students taught in Georgian	1.2%	2.3%	3.5%
[2] Italy	100%		0.9%	6.1%	7.0%
Jordan	100%		0.0%	0.3%	0.3%
Korea, Rep. of	100%		1.2%	2.2%	3.4%
[2] Kosovo	100%		4.7%	3.4%	8.1%
[2] Montenegro	100%		2.1%	6.6%	8.7%
North Macedonia	100%		1.5%	2.0%	3.5%
Slovenia	100%		2.3%	2.8%	5.1%
[2] Sweden	100%		1.6%	4.7%	6.2%

Grades 8–9

Country	International Target Population		Exclusions from National Target Population		
	Coverage	Notes on Coverage	School-Level Exclusions	Within-Sample Exclusions	Overall Exclusions
Jordan	100%		0.0%	0.3%	0.3%
Korea, Rep. of	100%		1.7%	1.4%	3.1%
[2] Sweden	100%		1.6%	4.8%	6.4%

1 National Target Population does not include all of the International Target Population.
2 National Defined Population covers 90% to 95% of National Target Population.

Target Population and Sample Sizes

Exhibit 5 shows the number of schools and students in each country's target population (after school-level exclusions), as well as an estimate of the student population size based on the combined longitudinal sample data, providing a validity check on the sampling procedure. These figures may differ from those reported in TIMSS 2023 if more students participated in the overall longitudinal study compared to the 2023 data collection alone. In general, the population sizes estimated from the samples closely matched the corresponding population sizes from the sampling frames.

Exhibits 6 and 7 report sample sizes achieved in terms of schools and students, respectively, providing more detailed numbers for each of the two data collection years.

Exhibit 5: TIMSS 2023 Longitudinal Target Population and Sample Sizes

Grades 4–5

Country	Target Population		Overall Longitudinal Sample*		
	Schools	Students	Total Number of Schools that Participated	Total Number of Students Assessed	Student Population Size Estimated from Sample
Georgia	1,895	49,493	156	4,491	48,046
Italy	6,636	514,619	152	4,610	508,046
Jordan	3,866	184,981	224	6,172	170,810
Korea, Rep. of	5,555	454,905	156	4,507	407,382
Kosovo	621	25,363	152	4,747	23,806
Montenegro	141	7,332	141	4,559	7,375
North Macedonia	315	19,322	150	4,866	18,428
Slovenia	449	21,658	149	5,015	18,063
Sweden	3,253	126,222	160	5,358	121,217

Grades 8–9

Country	Target Population		Overall Longitudinal Sample*		
	Schools	Students	Total Number of Schools that Participated	Total Number of Students Assessed	Student Population Size Estimated from Sample
Jordan	2,976	184,107	225	7,051	172,511
Korea, Rep. of	3,033	468,202	176	4,410	411,984
Sweden	1,619	119,797	154	5,303	121,796

* The overall TIMSS 2023 Longitudinal sample includes schools and students assessed at one or both time points.

Exhibit 6: TIMSS 2023 Longitudinal School Sample Sizes

Grades 4–5

Country	Schools in Original Sample	Eligible Schools in Original Sample	Original Schools that Participated		Replacement Schools that Participated		Total Schools that Participated	
			2023	2024	2023	2024	2023	2024
Georgia	156	156	151	146	5	5	156	151
Italy	152	152	146	146	6	6	152	152
Jordan	224	224	224	224	0	0	224	224
Korea, Rep. of	156	156	146	142	10	7	156	149
Kosovo	152	152	152	151	0	0	152	151
Montenegro	141	141	141	140	0	0	141	140
North Macedonia	150	150	149	149	0	0	149	149
Slovenia	150	150	144	141	5	5	149	146
Sweden	164	161	159	157	1	1	160	158

Grades 8–9

Country	Schools in Original Sample	Eligible Schools in Original Sample	Original Schools that Participated		Replacement Schools that Participated		Total Schools that Participated	
			2023	2024	2023	2024	2023	2024
Jordan	225	225	225	224	0	0	225	224
Korea, Rep. of	176	176	140	138	36	36	176	174
Sweden	157	155	154	153	0	0	154	153

Exhibit 7: TIMSS 2023 Longitudinal Student Sample Sizes

Grades 4–5

Country	Students Sampled in Participating Schools	Students Withdrawn from Class/ School	Students Excluded	Students Eligible	Students Absent		Students Inaccessible in 2024 (Attrition)	Students Assessed	
					2023	2024		2023	2024
Georgia	4,688	37	115	4,536	125	425	250	4,411	3,861
Italy	4,970	14	307	4,649	196	222	86	4,453	4,341
Jordan	6,500	267	22	6,211	146	297	668	6,065	5,246
Korea, Rep. of	4,719	85	88	4,546	191	618	219	4,355	3,709
Kosovo	5,028	71	126	4,831	222	341	166	4,609	4,324
Montenegro	4,839	18	170	4,651	234	288	71	4,417	4,292
North Macedonia	5,261	165	97	4,999	366	508	162	4,633	4,361
Slovenia	5,312	16	144	5,152	406	696	83	4,746	4,373
Sweden	5,743	60	243	5,440	301	490	302	5,139	4,648

Grades 8–9

Country	Students Sampled in Participating Schools	Students Withdrawn from Class/ School	Students Excluded	Students Eligible	Students Absent		Students Inaccessible in 2024 (Attrition)	Students Assessed	
					2023	2024		2023	2024
Jordan	7,571	442	28	7,101	139	470	621	6,962	6,010
Korea, Rep. of	4,586	50	64	4,472	215	332	144	4,257	3,996
Sweden	5,900	48	246	5,606	524	881	148	5,082	4,577

Students attending a sampled class at the time the sample was chosen but leaving the class before the 2023 assessment was administered were classified as "withdrawn."
Students with a disability or language barrier that prevented them from participating in the assessment were classified as "excluded."
Students not present when the assessment was administered, and not subsequently assessed in a make-up session, were classified as "absent."

Participation Rates

The TIMSS Standards for sampling participation are described in Chapter 9 of the *TIMSS 2023 Technical Report*, and Chapter 3 provides information on how participation rates were calculated.

In determining whether countries met the TIMSS standards for sampling participation based on their TIMSS 2023 Longitudinal samples, 2024 participation rates were calculated solely based on students present in the same schools in 2024 (as in 2023), excluding inaccessible students (attrition cases). The 2024 student participation rate was calculated as the ratio of the number of selected students who participated in 2024 to the total number of selected students who should have been assessed in participating schools and classes, excluding students who were inaccessible (attrition).

School, classroom, student, and overall participation rates are reported separately for the 2023 and 2024 data collections in Exhibit 8 (weighted) and Exhibit 9 (unweighted). Modest decreases in participation are observed between 2023 and 2024 at the student level. North Macedonia had one school treated as non-participating in 2023 due to low class participation, which is now participating in 2024. Korea did not meet the standards for participation in 2024 at the Grade 9 level, with only 80% weighted school participation and 74% weighted overall participation before replacement. In 2023, Korea had just met the standards with 76% weighted overall participation before replacement.

Exhibit 8: TIMSS 2023 Longitudinal Participation Rates (Weighted)

Grades 4–5

Country	School Participation		Class Participation	Student Participation	Overall Participation	
	Before Replacement	After Replacement			Before Replacement	After Replacement
Georgia						
2023	97%	100%	100%	98%	95%	98%
2024	97%	100%	100%	90%	87%	90%
Italy						
2023	96%	100%	100%	96%	92%	96%
2024	96%	100%	100%	95%	92%	95%
Jordan						
2023	100%	100%	100%	97%	97%	97%
2024	100%	100%	100%	95%	95%	95%
Korea, Rep. of						
2023	93%	100%	100%	96%	89%	96%
2024	93%	100%	100%	89%	83%	89%
Kosovo						
2023	100%	100%	100%	95%	95%	95%
2024	100%	100%	100%	93%	93%	93%
Montenegro						
2023	100%	100%	100%	95%	95%	95%
2024	100%	100%	100%	94%	94%	94%
North Macedonia						
2023	99%	99%	100%	93%	92%	92%
2024	100%	100%	100%	90%	90%	90%
Slovenia						
2023	96%	99%	99%	92%	88%	91%
2024	96%	99%	99%	87%	83%	86%
Sweden						
2023	100%	100%	100%	95%	94%	95%
2024	100%	100%	100%	90%	90%	90%

Exhibit 8: TIMSS 2023 Longitudinal Participation Rates (Weighted) (Continued)

Grades 8–9

Country	School Participation		Class Participation	Student Participation	Overall Participation	
	Before Replacement	After Replacement			Before Replacement	After Replacement
Jordan						
2023	100%	100%	100%	98%	98%	98%
2024	100%	100%	100%	93%	93%	93%
† Korea, Rep. of						
2023	80%	100%	100%	95%	76%	95%
2024	80%	100%	100%	92%	74%	92%
Sweden						
2023	99%	99%	100%	91%	90%	90%
2024	99%	99%	100%	85%	84%	84%

Participation rates for 2024 were calculated excluding students not assessed due to inaccessibility (attrition).

TIMSS guidelines for sampling participation: The minimum acceptable participation rates were 85% of schools, 95% of classes, and 85% of students, or a combined rate (the product of school, class, and student participation) of 75%. Participants not meeting these guidelines for either time point were annotated as follows:

† Met guidelines for sample participation rates only after replacement schools were included

‡ Nearly satisfied guidelines for sample participation rates after replacement schools were included

≡ Did not satisfy guidelines for sample participation rates

Exhibit 9: TIMSS 2023 Longitudinal Participation Rates (Unweighted)

Grades 4–5

Country	School Participation		Class Participation	Student Participation	Overall Participation	
	Before Replacement	After Replacement			Before Replacement	After Replacement
Georgia						
2023	97%	100%	100%	97%	94%	97%
2024	97%	100%	100%	90%	87%	90%
Italy						
2023	96%	100%	100%	96%	92%	96%
2024	96%	100%	100%	95%	91%	95%
Jordan						
2023	100%	100%	100%	98%	98%	98%
2024	100%	100%	100%	95%	95%	95%
Korea, Rep. of						
2023	94%	100%	100%	96%	90%	96%
2024	94%	100%	100%	86%	80%	86%
Kosovo						
2023	100%	100%	100%	95%	95%	95%
2024	100%	100%	100%	93%	93%	93%
Montenegro						
2023	100%	100%	100%	95%	95%	95%
2024	100%	100%	100%	94%	94%	94%
North Macedonia						
2023	99%	99%	100%	93%	92%	92%
2024	100%	100%	100%	90%	90%	90%
Slovenia						
2023	96%	99%	99%	92%	88%	91%
2024	96%	99%	99%	86%	82%	85%
Sweden						
2023	99%	99%	100%	94%	93%	94%
2024	99%	99%	100%	91%	89%	90%

Exhibit 9: TIMSS 2023 Longitudinal Participation Rates (Unweighted) (Continued)

Grades 8–9

Country	School Participation		Class Participation	Student Participation	Overall Participation	
	Before Replacement	After Replacement			Before Replacement	After Replacement
Jordan						
2023	100%	100%	100%	98%	98%	98%
2024	100%	100%	100%	93%	93%	93%
† Korea, Rep. of						
2023	80%	100%	100%	95%	76%	95%
2024	80%	100%	100%	92%	73%	92%
Sweden						
2023	99%	99%	100%	91%	90%	90%
2024	99%	99%	100%	84%	83%	83%

Participation rates for 2024 were calculated excluding students not assessed due to inaccessibility (attrition).

TIMSS guidelines for sampling participation: The minimum acceptable participation rates were 85% of schools, 95% of classes, and 85% of students, or a combined rate (the product of school, class, and student participation) of 75%. Participants not meeting these guidelines for either time point were annotated as follows:

† Met guidelines for sample participation rates only after replacement schools were included

‡ Nearly satisfied guidelines for sample participation rates after replacement schools were included

≡ Did not satisfy guidelines for sample participation rates

The overall TIMSS 2023 Longitudinal participation rates and attrition rates are reported in Exhibit 10 (weighted) and Exhibit 11 (unweighted). The overall longitudinal participation rate is measured with replacement schools, based on the number of students assessed at one or both time points. Participation rates are also reported for students who participated at both time points (excluding attrition cases), only in 2023, and only in 2024 (excluding attrition cases). The exhibits also report an attrition rate, measured as the ratio of the number of students not assessed in 2024 because they were inaccessible to the total number of students who should have been assessed in participating schools and classes.

Most countries had weighted attrition rates of less than 6%, except for Jordan, with 12% at Grade 5 and 11% at Grade 9. Cited reasons for attrition included inaccessibility to students in refugee camps and students no longer in the selected school due to transfer from public to private schools, financial hardships, and joining the workforce at Grade 9.

Exhibit 10: TIMSS 2023 Longitudinal Overall Participation Rates (Weighted)

Grades 4–5

Country	Overall Longitudinal Participation (2023 or 2024)*	Participation in 2023 and 2024	Participation in 2023	Participation in 2024	Percent of Students Inaccessible in 2024 (Attrition Rate)
Georgia	99%	88%	98%	90%	5%
Italy	99%	92%	96%	95%	2%
Jordan	99%	92%	97%	95%	12%
Korea, Rep. of	99%	85%	96%	89%	5%
Kosovo	99%	89%	95%	93%	3%
Montenegro	98%	90%	95%	94%	2%
North Macedonia	97%	85%	92%	90%	4%
Slovenia	96%	81%	91%	86%	2%
Sweden	98%	86%	95%	90%	5%

Grades 8–9

Country	Overall Longitudinal Participation (2023 or 2024)*	Participation in 2023 and 2024	Participation in 2023	Participation in 2024	Percent of Students Inaccessible in 2024 (Attrition Rate)
Jordan	99%	91%	98%	93%	11%
† Korea, Rep. of	99%	88%	95%	92%	3%
Sweden	94%	80%	90%	84%	3%

* Students were considered participating in TIMSS 2023 Longitudinal if they were assessed at one or both time points. Overall participation rates were calculated with replacement schools excluding students not assessed in 2024 due to inaccessibility (attrition).

† Met guidelines for sample participation rates only after replacement schools were included. See Exhibits 8 and 9.

Exhibit 11: TIMSS 2023 Longitudinal Overall Participation Rates (Unweighted)

Grades 4–5

Country	Overall Longitudinal Participation (2023 or 2024)*	Participation in 2023 and 2024	Participation in 2023	Participation in 2024	Percent of Students Inaccessible in 2024 (Attrition Rate)
Georgia	99%	88%	97%	90%	6%
Italy	99%	92%	96%	95%	2%
Jordan	99%	93%	98%	95%	11%
Korea, Rep. of	99%	82%	96%	86%	5%
Kosovo	98%	90%	95%	93%	3%
Montenegro	98%	91%	95%	94%	2%
North Macedonia	97%	85%	92%	90%	3%
Slovenia	96%	80%	91%	85%	2%
Sweden	98%	86%	94%	90%	6%

Grades 8–9

Country	Overall Longitudinal Participation (2023 or 2024)*	Participation in 2023 and 2024	Participation in 2023	Participation in 2024	Percent of Students Inaccessible in 2024 (Attrition Rate)
Jordan	99%	91%	98%	93%	9%
† Korea, Rep. of	99%	89%	95%	92%	3%
Sweden	94%	79%	90%	83%	3%

* Students were considered participating in TIMSS 2023 Longitudinal if they were assessed at one or both time points. Overall participation rates were calculated with replacement schools excluding students not assessed in 2024 due to inaccessibility (attrition).

† Met guidelines for sample participation rates only after replacement schools were included. See Exhibits 8 and 9.

SECTION 6

Student Achievement Estimation

The main goal of the TIMSS 2023 Longitudinal Study was to estimate students' growth in mathematics and science achievement over one year of schooling. To fulfill this goal, the psychometric analyses involved joint estimation of achievement at the two time points and reporting the difference between them as average growth. The results are reported using the existing TIMSS metric established in 1995. Producing these estimates relied on the item response theory (IRT) item parameters used in TIMSS 2023. It involved adaptations of procedures that followed the same rigorous quality standards as those observed in TIMSS 2023 (see Chapter 11 and Chapter 12 of the *TIMSS 2023 Technical Report*).

The TIMSS 2023 Longitudinal dataset had several unique aspects to account for in the analyses, such as the need to estimate growth based on the joint achievement data over two time points, imputing achievement estimates for students who participated only at one time point, and handling an increased number of potential conditioning variables for the latent regression model. This section describes the procedures implemented to estimate achievement for mathematics and science jointly across time points on the TIMSS reporting scale. Additional reviews and analyses were conducted to ensure the quality and validity of the results.

Data Review

As in TIMSS 2023, the data collected in 2024 underwent extensive item-by-item and country-by-country reviews to identify any statistical and psychometric issues that may jeopardize the quality of the TIMSS 2023 Longitudinal growth measures. Details of the types of review conducted are described in Chapter 10 of the *TIMSS 2023 Technical Report*.

With the same achievement items administered in both 2023 and 2024, an essential part of the data review involved comparing item statistics for each country between the two time points. The procedures were similar to those used for reviewing item statistics in TIMSS, focusing on examining differences between years in item percent correct, item discrimination, item nonresponse rates, and human-scoring reliability. Additionally, IRT-based root mean square deviation (RMSD) statistics were calculated for each item by country using the 2024 data and then compared to the same statistics computed with the 2023 data.

The primary aim of the TIMSS 2023 Longitudinal data review was to detect any unusual changes in item statistics between years, which might indicate a problem in using the item to measure growth. Item statistics were compared in tabular form for each item, country by country, and plots were produced for each country to compare changes in the item statistics between years visually. Between 2023 and 2024, countries were expected to show overall patterns of increased item percent correct and similar or decreased nonresponse rates. Item discrimination and scoring reliability rates were expected to remain similar.

The review flagged a small number of instances where constructed-response items showed unusual changes in difficulty within a country from one year to the next. Further exploration of the data and communication with the NRCs revealed that the implementation of scoring for specific items was not consistently followed across both time points, and data for the items were removed from either the 2023 or the 2024 data collection results. In addition, any item deletions or recodes that were implemented in the 2023 data based on the TIMSS 2023 item review were also implemented in the 2024 data. Appendix C includes a list of deleted items and recodes in the TIMSS 2023 Longitudinal data. Items are listed according to their assigned item block, sequence, and unique item ID used to identify the variables in the data files.

↓ Appendix C: Modifications to the TIMSS 2023 Longitudinal Achievement Data

In 2024, a small number of students were assigned the wrong test booklet, resulting in the same block of items being administered to the same student in both 2023 and 2024. For these cases, the 2024 response data for the repeated items were deleted.

Sample Sizes

As described earlier in Section 5, The TIMSS 2023 Longitudinal samples include all students who participated in at least one of the two data collections. The psychometric analysis implementation was used to impute plausible values (PVs) of achievement at each time point for both mathematics and science, regardless of whether students had data for both years or for only one of them. The approach imputed results for students missing data for one of the years, maximizing the information provided by their available data and the observed relationships between variables across the distribution.

This procedure relied on the majority of students in each country having achievement item responses and context questionnaire item responses at both time points, with just small percentages of students having data at only one time point. Exhibit 12 reports the sample sizes for the TIMSS 2023 Longitudinal achievement estimation, corresponding to the total student sample sizes, along with the percentages of students with data in both years, 2023 only, or 2024 only.

Exhibit 12: TIMSS 2023 Longitudinal Sample Sizes for Achievement Estimation

Grades 4–5

Country	Total Number of Students	Percent of Students by Participation Status		
		Participated in 2023 & 2024	Participated in 2023 Only	Participated in 2024 Only
Georgia	4,491	84.2%	14.3%	1.5%
Italy	4,610	91.1%	5.6%	3.3%
Jordan	6,172	81.9%	16.3%	1.9%
Korea, Rep. of	4,507	81.3%	15.2%	3.5%
Kosovo	4,747	88.2%	8.6%	3.1%
Montenegro	4,559	90.8%	5.8%	3.4%
North Macedonia	4,866	85.2%	10.5%	4.3%
Slovenia	5,015	82.4%	12.1%	5.5%
Sweden	5,358	82.8%	13.4%	3.8%

Grades 8–9

Country	Total Number of Students	Percent of Students by Participation Status		
		Participated in 2023 & 2024	Participated in 2023 Only	Participated in 2024 Only
Jordan	7,051	81.6%	17.0%	1.4%
Korea, Rep. of	4,410	86.7%	9.9%	3.4%
Sweden	5,303	82.4%	13.4%	4.2%

Because of rounding some results may appear inconsistent.

Braided Conditioning Models

As described in Chapter 12 of the *TIMSS 2023 Technical Report*, TIMSS imputes plausible values of student achievement data in both mathematics and science through conditioning. This involves estimating and applying a latent regression model that includes a substantial amount of all available contextual information.

Given student participation in one or both data collection years, a "braided" conditioning approach was implemented, which estimated three different latent regression models, each optimizing the available context information of the group for which the corresponding set of achievement results is estimated. The following three models were estimated for each country:

- Model 1: Used both 2023 and 2024 context variables
- Model 2: Used only 2023 context variables
- Model 3: Used only 2024 context variables

This braided approach maximizes the available context data for estimating achievement for the three groups of students. All TIMSS 2023 Longitudinal students were included in estimating each model, but the final set of plausible values was selected for each student based on whether they participated and produced context data in both years (Model 1), in 2023 only (Model 2), or in 2024 only (Model 3).

The same primary conditioning variables were used in estimating all three latent regression models. In addition to students' gender and the language of the test, four primary conditioning variables were included, representing the class average mathematics achievement and average science achievement, separately for each year. This differed from the scaling procedures in TIMSS 2023, which used a single class average for both subjects. However, using separate class averages for the two subjects in both years accounted for roughly 70% of the between-class, between-subject, and between-year variance. By using the four separate class averages, the regression model accounts for this variance and improves the simultaneous estimation of the two subjects and years. The interim achievement averages were based on expected a posteriori (EAP) estimates using the TIMSS 2023 item parameters with the MIRT package (Chalmers, 2012) in the R statistical programming language (R Core Team, 2024).

Principal Component Analysis

TIMSS uses principal component analysis (PCA) to eliminate multicollinearity and to reduce the number of variables included in the latent regression model (see Chapter 12 of the *TIMSS 2023 Technical Report*). Responses to the context questionnaire items in both years were preprocessed and used in the calculation of the principal components. This included responses to the student questionnaire and the home questionnaire. Principal component analysis was carried out separately for each country, using the variables corresponding to each of the three models.

Two new procedures were used in TIMSS 2023 Longitudinal to remove redundant sources of variance from the PCA and maximize the amount of unique variance used for the latent regression model. The first one was introduced before running the PCA, and checked each variable for collinearity with other variables in the model. In this step, collinear variables were removed from the PCA. A second step partialled out the variance already accounted for by the primary conditioning variables in conducting the PCA. In other words, the PCA was run based on the residual variance of the contextual variables after partialling out the variance from the primary conditioning variables.

The new procedures for partialling out variance as well as using separate class EAPs were introduced carefully in a stepwise review to ensure they were implemented properly and had no unexpected impact on the results. First, results were produced using the same procedures used for TIMSS 2023. Second, results were produced using the separate four class EAP means. Third and fourth iterations added the steps to remove collinear variables

and partial out the variance of the primary conditioning variables from the variables used for the principal components. The results of all four iterations were compared, looking at the variance explained and the average achievement and standard deviations resulting from each model, ensuring no substantial change in the overall results while improving the amount of variance explained by the model.

Exhibit 13 provides a summary of the number of conditioning variables used for achievement estimation in each of the three models. In general, the rule is to select as many principal components to account for 90% of the context variance, but not exceeding 5% of the total sample size.

Exhibit 13: Conditioning Variables Used for TIMSS 2023 Longitudinal Proficiency Estimation

Grades 4–5

Country	Number of Primary Variables	Model 1 (2023 & 2024 Context Data)			Model 2 (2023 Context Data Only)			Model 3 (2024 Context Data Only)		
		Number of PCs Available	Number of PCs Retained	Percent of Variance Explained	Number of PCs Available	Number of PCs Retained	Percent of Variance Explained	Number of PCs Available	Number of PCs Retained	Percent of Variance Explained
Georgia	5	1,254	224	61	681	224	75	576	224	80
Italy	5	1,209	230	59	668	230	75	558	230	80
Jordan	5	1,251	308	70	679	308	86	575	304	90
Korea, Rep. of	5	1,106	225	63	600	225	78	523	225	83
Kosovo	5	1,241	237	62	669	237	77	574	237	82
Montenegro	5	1,240	227	61	666	227	76	576	227	80
North Macedonia	6	1,260	243	63	682	243	78	579	243	83
Slovenia	5	1,184	250	65	642	250	81	552	250	85
Sweden	5	1,157	267	67	628	267	83	538	267	88

Grades 8–9

Country	Number of Primary Variables	Model 1 (2023 & 2024 Context Data)			Model 2 (2023 Context Data Only)			Model 3 (2024 Context Data Only)		
		Number of PCs Available	Number of PCs Retained	Percent of Variance Explained	Number of PCs Available	Number of PCs Retained	Percent of Variance Explained	Number of PCs Available	Number of PCs Retained	Percent of Variance Explained
Jordan	5	1,666	352	72	612	329	90	1,055	352	84
Korea, Rep. of	5	1,037	220	68	532	220	83	514	220	86
Sweden	5	1,747	265	69	884	265	80	1,306	265	77

PCs = Principal Components

Latent Regression Models

The DGROUP version of Educational Testing Service's MGROUP programs (Rogers et al., 2006; Sheehan, 1985) was used to estimate the latent regression models and to impute plausible values (PVs). These programs take as input the students' responses to the items

in both years, nonresponse indicator variables for each year by content domain, the item parameters estimated for TIMSS 2023, and the conditioning variables. DGROUP was used to estimate the set of three four-dimensional models for each country, each imputing five PVs of achievement simultaneously for overall mathematics and science at each time point (mathematics in 2023, mathematics in 2024, science in 2023, and science in 2024).

Scale Transformation

The plausible values imputed for the 2023 and 2024 data were transformed onto the existing TIMSS reporting metric using linear transformations.

The TIMSS Longitudinal analysis produced a joint set of plausible values for 2023 and 2024 that is based on all information available about each student participating in the longitudinal study. The 2023 component includes plausible values that are (and must be, since they are random imputations and make use of more information) slightly different from the plausible values estimated in the TIMSS 2023 study without the longitudinal components. This is because the longitudinal study contains almost twice the amount of information about each student who participated in both the 2023 assessment and the 2024 reassessment. The TIMSS 2023 and 2024 joint estimates are necessary for analyses that utilize data from both years, because achievement data and context data are correlated across years.

The transformation constants for TIMSS 2023 Longitudinal were derived by finding the linear transformation equation to convert the newly estimated 2023 PVs in the logit metric produced for TIMSS 2023 Longitudinal countries onto their reported results in TIMSS 2023. This was done by matching the mean and standard deviation across all five PVs of 2023 achievement for each subject, using the pooled 2023 data from only longitudinal countries.

The constants were calculated using data from all students across longitudinal countries that participated in 2023, taking into account their 2023 PVs published for TIMSS 2023, as well as the 2023 component of the joint 2023 and 2024 PVs from the TIMSS 2023 Longitudinal Study. Students with data from both years were assigned the 2023 PV component from longitudinal Model 1, and students with data from only 2023 were assigned the 2023 PV component from longitudinal Model 2.

For each grade, the linear transformations are given by

$$PV_k^* = A_k + B_k \times PV_k$$

where PV_k are the five plausible values of scale k (mathematics or science) before transformation, PV_k^* are the plausible values of scale k after transformation, and A_k and B_k are the linear transformation constants.

Transformation constants were obtained by first computing the means (μ_k^*) and standard deviations (σ_k^*) across the five PVs for each scale using the published plausible values from TIMSS 2023. These were calculated using data only from countries that participated in TIMSS 2023 Longitudinal. Next, the means (μ_k) and standard deviations (σ_k) were calculated using the newly estimated PVs for 2023 based on the longitudinal analyses. From these calculations, the linear transformation constants are defined as:

$$B_k = \frac{\sigma_k^*}{\sigma_k}$$

and

$$A_k = \mu_k^* - B_k \times \mu_k.$$

Exhibit 14 shows the transformation constants for the TIMSS 2023 Longitudinal data. The constants were applied to both 2023 and 2024 PVs produced in the TIMSS 2023 Longitudinal achievement estimation.

Exhibit 14: Scale Transformation Constants for the TIMSS 2023 Longitudinal Data

Scale	TIMSS 2023 Published Results		Newly Estimated 2023 Longitudinal Results		A_k	B_k
	Mean	Standard Deviation	Mean	Standard Deviation		
Grades 4–5 Mathematics	497.56129	93.12510	-0.36605	1.10534	528.40150	84.25029
Grades 4–5 Science	481.89562	97.43317	-0.65593	0.95191	549.03337	102.35535
Grades 8–9 Mathematics	500.51301	123.94386	0.28592	0.99917	465.04593	124.04637
Grades 8–9 Science	493.14741	110.75920	-0.05098	0.93418	499.19185	118.56272

Validating the Results

The psychometric analysis of the TIMSS 2023 Longitudinal achievement data included extensive steps to ensure the quality of the results, many of which were parallel to those documented for TIMSS 2023 (see Chapter 12 of the *TIMSS 2023 Technical Report*). Unique to TIMSS 2023 Longitudinal is the joint psychometric analysis and achievement estimation of the 2023 and 2024 data, and the imputation of achievement for a subset of students missing data at one of the two data collection points.

To evaluate the validity of the 2024 achievement estimates generated through the braided imputation procedure, a quality control analysis was conducted for students who participated in the 2023 data collection but not in the 2024 data collection. The goal of this analysis was to assess whether the imputed values appeared in line with the results of students who participated in both time points but had similar background characteristics to

the non-participating students. This check was only done for students missing 2024 data, as the number of students missing 2023 data was relatively small.

To this end, a *synthetic reference distribution* was constructed using data from students who participated in both 2023 and 2024, reweighting the observed cases to match the background characteristics of the students who did not participate in 2024. The resulting synthetic distribution was compared to the imputed score distribution of students who did not participate in 2024.

The key assumption of this approach was that, under a well-specified imputation procedure and a plausible missing data mechanism, the distribution of imputed values for missing cases should resemble that of observed values from individuals with similar covariate profiles (Abayomi et al., 2008; Bondarenko & Raghunathan, 2016; van Buuren, 2018). Differences in background characteristics between the participating and non-participating groups may confound direct comparisons between observed and imputed values. Reweighting the observed sample to match the covariate distribution of the missing group can adjust for this to provide a counterfactual benchmark. Substantial divergence between the imputed and synthetic distributions may signal that the imputation procedure introduces bias or fails to adequately capture the underlying relationships in the data.

The analysis steps were conducted as follows:

1. Modeling 2024 Non-Participation via Propensity Scores

A logistic regression model was used to predict the probability of not participating in 2024, based on background information from 2023. Predictors included the principal component variables used in conditioning Model 2, the primary conditioning variables, and the student-level average TIMSS 2023 mathematics and science EAP scores.

2. Constructing Weights

To approximate what the group might have scored in 2024, propensity weights were constructed for 2024 participants (i.e., students who were assessed in 2024) and hence have achievement estimates that can be compared to those who have fully imputed estimates due to non-participation in 2024.

Specifically, for each student i who participated in the 2024 assessment, the following propensity weights were calculated:

$$w_i = \frac{E(X_i)}{1 - E(X_i)}$$

where $E(X_i)$ is the predicted probability of not participating in 2024, obtained from the logistic model described in Step 1. This equation derives from propensity score methods in the causal inference literature (Rosenbaum & Rubin, 1983), where it is commonly used

to estimate average treatment effects. Although the goal of the quality control measures presented here is not to make causal inferences, the idea is similar: The observed 2024 participants are reweighted to resemble the data of non-participating 2024 students based on their characteristics in 2023.

The computed propensity weights were then multiplied by students' sampling weights for students who participated in both years. This reweighted sample serves as a proxy for what the missing group might have looked like had they participated in 2024. For students not participating in 2024, only their sampling weights were used.

3. Comparing Synthetic and Imputed Distributions

The weights computed in Step 2 were used to construct reweighted distributions of 2024 achievement scores. Specifically, the synthetic achievement distribution of 2024 participants was compared to the imputed achievement score distribution for the students who did not participate in 2024. Stacked PVs, where each student's five PVs were treated as five independent observations, were used to represent achievement.

Findings

Appendix D presents overlaid distribution plots of the imputed and synthetic distributions for each country. The reweighted synthetic distributions generally aligned well with the imputed distributions, providing evidence that the conditioning approach was effective in imputing 2024 achievement for non-participating students. In some countries, certain conditioning variables were found to be associated with non-participation in 2024. Slight deviations between the distributions were observed in some countries, potentially due to variance in imputation or the relatively small number of non-participating students in 2024 with fully imputed scores.

↓ Appendix D: Imputed and Synthetic Distribution Plots for Validating TIMSS 2023 Longitudinal Achievement Estimates

SECTION 7

Standard Error Estimation

To quantify uncertainty due to sampling variance and imputation variance, each statistic reported in *TIMSS 2023 Longitudinal International Results in Mathematics and Science* is accompanied by an estimate of its standard error. These are estimated using the jackknife repeat replication (JRR) technique. In addition to providing a measure of uncertainty of results, standard errors are an essential component of confidence intervals and are necessary to conduct null hypothesis significance testing. Details about the sources of variance in TIMSS results and estimating JRR standard errors are described in Chapter 13 of the *TIMSS 2023 Technical Report*.

Appendix E provides details on the standard errors for the TIMSS 2023 Longitudinal countries' mean proficiency estimates in mathematics and science for both 2023 and 2024, reporting JRR sampling variance, imputation variance, total variance, and the overall standard error.

↓ Appendix E: TIMSS 2023 Longitudinal Summary Statistics and Standard Errors for Proficiency

SECTION 8

Context Questionnaire Scales

The TIMSS Context Questionnaires comprise sets of items, reported as scales, that measure an underlying latent construct. Scale variables created from the TIMSS 2023 Longitudinal context data, collected at both time points, can facilitate the exploration of factors related to growth in mathematics and science achievement within and across countries.

This section outlines the procedures for creating scale variables from data collected in 2024, including scales administered in both 2023 and 2024, as well as new scales introduced in 2024. Some item sets included in the TIMSS 2023 Longitudinal database were administered only in 2023. The methods and procedures for constructing, interpreting, and validating TIMSS Context Questionnaire Scales, as well as details about the scales created for the 2023 data collection, are described in Chapter 15 of the *TIMSS 2023 Technical Report*.

Exhibit 15 lists all scales administered for the 2024 data collection that have variables included in the TIMSS 2023 Longitudinal International Database. Columns indicate the assessment year the scale was first established and whether scale results are included in *TIMSS 2023 Longitudinal International Results in Mathematics and Science*.

Exhibit 15: List of TIMSS 2023 Longitudinal Context Questionnaire Scales Reporting 2024 Data

Grades 4–5

Scale Name	Respondent	Scale Score Variable Name	Scale Index Variable Name	Year Scale Metric Established	Included in TIMSS 2023 Longitudinal Results
Digital Self-Efficacy	Students	ASBLGSEC	ASDLGSEC	2023	
Students Value Environmental Preservation	Students	ASBLGVEP	ASDLGVEP	2023	
Sense of School Belonging	Students	ASBLGSSB	ASDLGSSB	2023	✓
Students' Well-Being at School*	Students	ASBLGSWB	ASDLGSWB	2024	✓
Student Bullying	Students	ASBLGSB	ASDLGSB	2023	✓
Students Like Learning Mathematics	Students	ASBLGSLM	ASDLGSLM	2023	✓
Instructional Clarity in Mathematics Lessons	Students	ASBLGICM	ASDLGICM	2023	✓
Disorderly Behavior during Mathematics Lessons	Students	ASBLGDML	ASDLGDML	2023	✓
Students Confident in Mathematics	Students	ASBLGSCM	ASDLGSCM	2023	✓
Students Like Learning Science	Students	ASBLGSLS	ASDLGSLS	2023	✓
Instructional Clarity in Science Lessons	Students	ASBLGICS	ASDLGICS	2023	✓
Disorderly Behavior during Science Lessons	Students	ASBLGDSL	ASDLGDSL	2023	✓
Students Confident in Science	Students	ASBLGSCS	ASDLGSCS	2023	✓
Parents' Perceptions of Their Child's School	Parents	ASBLHPSP	ASDLHPSP	2015	
Parents' Activities with Their Child*	Parents	ASBLHPAC	ASDLHPAC	2024	
Home Socioeconomic Status	Parents	ASBLHSES	ASDLHSES	2019	✓
Home Resources for Learning	Parents/ Students	ASBLGHRL	ASDLGHRL	2011	
Instruction Affected by Mathematics Resource Shortages	Principals	ACBLGMRS	ACDLGMRS	2011	
Instruction Affected by Science Resource Shortages	Principals	ACBLGSRS	ACDLGSRS	2011	
School Emphasis on Academic Success - Principals' Reports	Principals	ACBLGEAS	ACDLGEAS	2015	✓
School Innovation*	Principals	ACBLGSIN	ACDLGSIN	2024	
School Emphasis on Academic Success - Teachers' Reports	Teachers	ATBLGEAS	ATDLGEAS	2015	
Teachers' Job Satisfaction	Teachers	ATBLGTJS	ATDLGTJS	2015	
Teaching Limited by Students Not Ready for Instruction	Teachers	ATBLGLSN	ATDLGLSN	2015	✓
Teachers Confident in Teaching Mathematics*	Teachers	ATBLMTCM	ATDLGTCM	2024	
Teachers Confident in Teaching Science*	Teachers	ATBLSTCS	ATDLSTCS	2024	

Exhibit 15: List of TIMSS 2023 Longitudinal Context Questionnaire Scales Reporting 2024 Data (Continued)

Grades 8–9

Scale Name	Respondent	Scale Score Variable Name	Scale Index Variable Name	Year Scale Metric Established	Included in TIMSS 2023 Longitudinal Results
Home Educational Resources*	Students	BSBLGHER	BSDLGHER	2011	✓
Digital Self-Efficacy	Students	BSBLGSEC	BSDLGSEC	2023	
Students Value Environmental Preservation	Students	BSBLGVEP	BSDLGVEP	2023	
Sense of School Belonging	Students	BSBLGSSB	BSDLGSSB	2023	✓
Students' Well-Being at School*	Students	BSBLGSWB	BSDLGSWB	2024	✓
Student Bullying	Students	BSBLGSB	BSDLGSB	2023	✓
Students Like Learning Mathematics	Students	BSBLGSLM	BSDLGSLM	2023	✓
Instructional Clarity in Mathematics Lessons	Students	BSBLGICM	BSDLGICM	2023	✓
Disorderly Behavior during Mathematics Lessons	Students	BSBLGDML	BSDLGDML	2023	✓
Students Confident in Mathematics	Students	BSBLGSCM	BSDLGSCM	2023	✓
Students Value Mathematics	Students	BSBLGSVM	BSDLGSVM	2011	✓
Students' Perceptions of Mathematics Homework*	Students	BSBLGPMH	BSDLGPMH	2024	
Students Like Learning Science	Students	BSBLGSLS	BSDLGSLS	2023	
Instructional Clarity in Science Lessons	Students	BSBLGICS	BSDLGICS	2023	
Disorderly Behavior during Science Lessons	Students	BSBLGDSL	BSDLGDSL	2023	
Students Confident in Science	Students	BSBLGSCS	BSDLGSCS	2023	
Students Value Science	Students	BSBLGSVS	BSDLGSVS	2011	✓
Students' Perceptions of Science Homework*	Students	BSBLGPSH	BSDLGPSH	2024	
Students Like Learning Biology	Students	BSBLGSLB	BSDLGSLB	2023	
Instructional Clarity in Biology Lessons	Students	BSBLGICB	BSDLGICB	2023	
Disorderly Behavior during Biology Lessons	Students	BSBLGDBL	BSDLGDBL	2023	
Students Confident in Biology	Students	BSBLGSCB	BSDLGSCB	2023	
Students' Perceptions of Biology Homework*	Students	BSBLGPBH	BSDLGPBH	2024	
Students Like Learning Chemistry	Students	BSBLGSLC	BSDLGSLC	2023	
Instructional Clarity in Chemistry Lessons	Students	BSBLGICC	BSDLGICC	2023	
Disorderly Behavior during Chemistry Lessons	Students	BSBLGDCL	BSDLGDCL	2023	

Exhibit 15: List of TIMSS 2023 Longitudinal Context Questionnaire Scales Reporting 2024 Data (Continued)

Grades 8–9

Scale Name	Respondent	Scale Score Variable Name	Scale Index Variable Name	Year Scale Metric Established	Included in TIMSS 2023 Longitudinal Results
Students Confident in Chemistry	Students	BSBLGSCC	BSDLGSCC	2023	
Students' Perceptions of Chemistry Homework*	Students	BSBLGPCH	BSDLGPCH	2024	
Students Like Learning Physics	Students	BSBLGSLP	BSDLGSLP	2023	
Instructional Clarity in Physics Lessons	Students	BSBLGICP	BSDLGICP	2023	
Disorderly Behavior during Physics Lessons	Students	BSBLGDPL	BSDLGDPL	2023	
Students Confident in Physics	Students	BSBLGSCP	BSDLGSCP	2023	
Students' Perceptions of Physics Homework*	Students	BSBLGPPH	BSDLGPPH	2024	
Students Like Learning Earth Science	Students	BSBLGSLE	BSDLGSLE	2023	
Instructional Clarity in Earth Science Lessons	Students	BSBLGICE	BSDLGICE	2023	
Disorderly Behavior during Earth Science Lessons	Students	BSBLGDEL	BSDLGDEL	2023	
Students Confident in Earth Science	Students	BSBLGSCE	BSDLGSCE	2023	
Students' Perceptions of Earth Science Homework*	Students	BSBLGPEH	BSDLGPEH	2024	
Instruction Affected by Mathematics Resource Shortages	Principals	BCBLGMRS	BCDLGMRS	2011	
Instruction Affected by Science Resource Shortages	Principals	BCBLGSRS	BCDLGSRS	2011	
School Emphasis on Academic Success - Principals' Reports	Principals	BCBLGEAS	BCDLGEAS	2015	✓
School Innovation*	Principals	BCBLGSIN	BCDLGSIN	2024	
School Emphasis on Academic Success - Teachers' Reports	Teachers	BTBLGEAS	BTDLGEAS	2015	
Teachers' Job Satisfaction	Teachers	BTBLGTJS	BTDLGTJS	2015	
Teaching Limited by Students Not Ready for Instruction	Teachers	BTBLGLSN	BTDLGLSN	2015	✓
Teachers Confident in Teaching Mathematics*	Teachers	BTBLMTCM	BTDLMTCM	2024	
Teachers Confident in Teaching Science*	Teachers	BTBLSTCS	BTDLSTCS	2024	

* Indicates the LCA-CS method was used to specify the raw cut points for the scale regions.

Context Scales Common in 2023 and 2024

Most of the scales reported in 2024 were also reported in 2023, and were calculated using the same item parameters, transformation constants, and scale region cut points estimated

for TIMSS 2023. These estimated results for the 2024 data collection on the same scale as the 2023 data collection were saved under separate variables for analysis.

Although some changes in the way in which students respond to context scales may be expected from year to year, quality checks were conducted to ensure that any changes were of a reasonable magnitude and that the TIMSS 2023 item parameters had an equally good fit to the 2024 data as they did in 2023. Scatterplots were created to examine changes in response patterns across the scale categories between years for each country. In addition, IRT-based RMSD statistics were computed for each item by country using the 2024 data and then compared to the same statistics calculated for the 2023 data. Visual inspection of scatterplots comparing statistics allowed for evaluating if there were unusually large changes for any country on an item relative to other countries. Cronbach's alpha coefficients and results of the principal component analysis conducted to evaluate the reliability of the scale in each country, as well as the scale's relationship with achievement, were compared between 2023 and 2024. Based on these analyses, the scales common to 2023 and 2024 generally behaved similarly between years for all countries, providing evidence that the scale results could be compared over time. The procedures identified two scales in one country with altered item translations for the 2024 instruments, rendering them noncomparable between years. Data collected in 2024 were removed for the scale items in this country.

Context Scales New in 2024

A small number of new scales were introduced into the context questionnaires for the 2024 data collection: *Students' Well-Being at School*, *Students' Perceptions of Mathematics Homework* (Grade 9 only), *Students' Perceptions of Science Homework* (Grade 9 only, including separate science versions), *Parents' Activities with Their Child* (Grade 5 only), *School Innovation*, *Teachers Confident in Teaching Mathematics*, and *Teachers Confident in Teaching Science*.

Creating and validating the scales new to 2024 followed the same procedures as described in Chapter 15 of the *TIMSS 2023 Technical Report* for new scales created in TIMSS 2023. Item parameters were estimated using the Rasch partial credit model (PCM; Masters, 1982), and scale estimates were produced using weighted maximum likelihood estimation (Warm, 1989). Item parameters and scale transformation constants were estimated based on the pooled distribution of longitudinal countries at each grade, with each country weighted equally. Cut points to categorize respondents into reporting categories based on their scale scores were identified using the TIMSS LCA-based cut score method (LCA-CS method). Scales were evaluated for reliability and validity, noting some limited generalizability, especially at Grade 9, with only three countries contributing to the scale parameters. In particular, the Grade 9 *Students' Perceptions of Science Homework* scales were established based on data from only one country for the integrated

science scale, only two countries for the separate biology, chemistry, and physics scales, and only one country for the separate earth science scale.

Detailed information on the TIMSS 2023 Longitudinal scales that were created new for the 2024 data collection can be found in the following appendices:

- Appendix F provides a list of items comprising each scale, along with their corresponding response categories.

  ↓ Appendix F: TIMSS 2023 Longitudinal Context Scale Descriptions

- Appendix G provides international item parameters and item fit statistics. For each item, the delta parameter δ_i shows the estimated overall location of the item on the scale, and the tau parameters τ_{ij} show the location of the steps, expressed as deviations from delta ($b_{ij} = \delta_i - \tau_{ij}$).

  ↓ Appendix G: TIMSS 2023 Longitudinal Context Scale Item Parameters and Item Fit Statistics

- Appendix H reports the scale transformation constants applied to the international distribution of logit scores to put the 2024 estimates on the TIMSS (10,2) reporting metric.

  ↓ Appendix H: TIMSS 2023 Longitudinal Context Scale Transformation Constants

- Appendix I provides the equivalence tables of raw and transformed scale scores with the cut points used to create the scale regions.

  ↓ Appendix I: TIMSS 2023 Longitudinal Context Scale Equivalence Tables of Raw and Transformed Scale Scores

- Appendix J reports country-level Cronbach's alpha reliability coefficients and principal component analysis results.

  ↓ Appendix J: TIMSS 2023 Longitudinal Context Scale Reliability and Principal Component Analysis

- Appendix K reports country-level estimates of the relationship between the results of each scale and 2024 mathematics and science achievement.

  ↓ Appendix K: TIMSS 2023 Longitudinal Context Scale Relationships with Achievement

References

Abayomi, K., Gelman, A., & Levy, M. (2008). Diagnostics for multivariate imputations. *Journal of the Royal Statistical Society Series C: Applied Statistics, 57*(3), 273–291. https://doi.org/10.2139/ssrn.1010415

Bondarenko, I., & Raghunathan, T. (2016). Graphical and numerical diagnostic tools to assess suitability of multiple imputations and imputation models. *Statistics in Medicine, 35*(17), 3007–3020. https://doi.org/10.1002/sim.6926

Chalmers, R. P. (2012). mirt: A multidimensional item response theory package for the R environment. *Journal of Statistical Software, 48*(6), 1–29. https://doi.org/10.18637/jss.v048.i06

Masters, G. N. (1982). A Rasch model for partial credit scoring. *Psychometrika, 47*(2), 149–174.

R Core Team (2024). R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria. https://www.R-project.org/

Rogers, A., Tang, C., Lin, M.-J., & Kandathil, M. (2006). DGROUP [computer software]. Princeton, NJ: Educational Testing Service.

Rosenbaum, P. R., & Rubin, D. B. (1983). The central role of the propensity score in observational studies for causal effects. *Biometrika, 70*(1), 41–55. https://doi.org/10.2307/2335942

Sheehan, K. M. (1985). M-Group: Estimation of group effects in multivariate models [computer software, Version 3.2]. Princeton, NJ: Educational Testing Service.

van Buuren, S. (2018). *Flexible Imputation of Missing Data*. 2nd Edition, CRC Press, Boca Raton. https://doi.org/10.1201/9780429492259

von Davier, M., Fishbein, B., & Kennedy, A. (Eds.). (2024). *TIMSS 2023 Technical Report (Methods and Procedures)*. Boston College, TIMSS & PIRLS International Study Center. https://timss2023.org/methods

Warm, T. A. (1989). Weighted likelihood estimation of ability in item response theory. *Psychometrika, 54*(3), 427–450.