CHAPTER 8

Creating the TIMSS 2019 International Database

Kamil Kowolik
Mark Cockle
Milena Taneva

Preparing the TIMSS 2019 International Database and ensuring its integrity was a complex endeavor requiring extensive collaboration among IEA Hamburg, the TIMSS & PIRLS International Study Center, Statistics Canada, and the national centers of participating countries. Once the countries had created their data files and submitted them to IEA Hamburg, an exhaustive process of checking and editing known as "data cleaning" began. Data cleaning is the process of checking data for inconsistencies and formatting the data to create a standardized output.

For each TIMSS assessment, the overriding concerns of the data cleaning process are to ensure the following:

- All information in the database conformed to the internationally defined data structure

- The content of all codebooks and documentation appropriately reflected national adaptations to questionnaires

- All variables used for international comparisons were in fact comparable across countries (after harmonization, where necessary)

- All institutions involved in this process applied quality control measures throughout in order to assure the quality and accuracy of the TIMSS 2019 data.

For TIMSS 2019, IEA Hamburg was responsible for checking the data files from each country, applying standardized data cleaning rules to verify the accuracy and consistency of the data, and documenting any deviations from the international file structure. In addition, IEA Hamburg was responsible for processing and cleaning the data collected by eAssessment Player that delivers the assessment to students, importing student achievement response data for human-scoring into

IEA's Scoring System, and implementing machine scoring rules for achievement items according to specifications from the TIMSS & PIRLS International Study Center.

For countries participating in eTIMSS, student achievement data files were created by IEA Hamburg from the raw data collected by the eAssessment Player as well as the achievement item scores assigned through the Scoring System. For paperTIMSS instruments (including "bridge" booklets forming the basis for a link between paperTIMSS and eTIMSS) and paper-based context questionnaires, data files were created at each country's national center and reviewed prior to submission to IEA Hamburg. The National Research Coordinators (NRCs) from each participating country collaborated with IEA Hamburg to resolve any queries which emerged during the data cleaning process, and the NRCs checked interim versions of the national/benchmarking participant database(s) produced by IEA Hamburg. The TIMSS & PIRLS International Study Center provided the NRCs with univariate data almanacs containing summary item statistics on each variable so that the national centers could evaluate their data from an international perspective (see Chapter 10).

The TIMSS & PIRLS International Study Center also conducted all operational psychometric analyses of the achievement and context questionnaire data, as documented in Chapter 12 (achievement scaling) and Chapter 16 (context scaling), and produced achievement scores (plausible values), and context questionnaire scores, as well as other derived variables based on the context data. Using the Within-School Sampling Software (WinW3S)[1] database and response data provided by IEA Hamburg, Statistics Canada in collaboration with IEA Hamburg calculated the sampling weights, population coverage, and school and student participation rates—as documented in Chapter 3 and Chapter 9.

Data Sources

All data collected as part of TIMSS 2019 arrived at IEA Hamburg for processing and cleaning before going to the TIMSS & PIRLS International Study Center for verification and analysis and to Statistics Canada for calculating sampling weights and outcomes. This included data collected from: 1) paperTIMSS instruments, including achievement booklets and all context questionnaires; 2) IEA's Online SurveySystem which countries could use to administer home, teacher, and school questionnaires; and 3) the eAssessment Player that delivered the TIMSS assessment and a short questionnaire to students.

Data Entry and Verification of Paper Instruments

Each national center was responsible for entering the responses collected in paperTIMSS achievement booklets and paper-based context questionnaires into data files using the IEA Data Management Expert (DME) software. The DME is a software system developed by IEA Hamburg that facilitates data entry

---

1 WinW3S is a software developed by IEA Hamburg that stores participation information at school, teacher, class, and student levels in a relational database while maintaining a hierarchical ID system. The software allows users to perform all necessary within-school sampling according to the TIMSS standards, and also provides some data validation in and across these levels.

TIMSS & PIRLS
International Study Center
Lynch School of Education
BOSTON COLLEGE

CHAPTER 8: CREATING THE INTERNATIONAL DATABASE
METHODS AND PROCEDURES: TIMSS 2019 TECHNICAL REPORT   8.2

and includes validation checks to identify inconsistencies. As a general principle, national centers were instructed to enter data for any questionnaire that contained at least one valid response, discarding unused or empty instruments. This applied to countries that administered paperTIMSS as well as eTIMSS countries, as these countries administered at least some questionnaires on paper and also administered paper "bridge" booklets of achievement items to a subsample of students.

National centers entered responses from the paper instruments into data files using a predefined international codebook. The codebook defines the structure of the data to be entered and contains information about the variable names, lengths, labels, and missing codes, as well as variable ranges for continuous measures or counts and valid values for nominal or ordinal questions.

As documented in Chapter 5, countries participating in TIMSS are expected to make national adaptations to certain questions in the international questionnaires (e.g., the questions about parents' education must be adapted to the national context). Countries making such adaptations were required to adapt the codebook structure to reflect the adaptations made to the national questionnaire versions before beginning the data entry process.

To ensure consistency across participating countries, the basic rule for data entry in the DME required national staff to enter data "as is" without any interpretation, correction, truncation, imputation, or cleaning.

The guiding principles for data entry included the following:

- Responses to closed response items were coded as "1" if the first option was used, "2" if the second option was marked, and so on

- Responses to open response questions, for example number of students in the TIMSS class, were entered "as is" even if the value was outside the originally expected range

- Responses to filter questions and filter-dependent questions were entered exactly as filled in by the respondent, even if the information provided is logically inconsistent

- Non-response, ambiguous responses, responses given outside of the expected format, or conflicting responses (e.g., selection of two options in a multiple-choice question) were coded as "omitted or invalid."

As each respondent ID number was entered it was checked by the DME software for alignment with a five-digit checksum generated by WinW3S. A mistype in either the ID or the checksum resulted in an error message prompting the person entering the data to check the entry. The data-verification module of DME also checked for a range of other issues such as inconsistencies in identification codes and out-of-range or otherwise invalid codes. When such issues were flagged by the software, the individuals entering the data were prompted to resolve the inconsistency or confirm that an issue existed before resuming data entry.

Double-Data Entry

To check data entry reliability in participating countries, national centers were required to enter a 5 percent sample of each survey instrument (achievement booklet or questionnaire) twice by two different data entry persons (punchers) operating independently. IEA Hamburg recommended that countries begin the double-data entry process as early as possible during the data capture period in order to identify possible systematic misunderstandings or mishandlings of data-entry rules and to initiate appropriate remedial actions—for example, retraining national center staff. Those entering the data were required to resolve discrepancies between the first and second data entries by consulting the original questionnaire and applying the international rules in a uniform way.

Although it was desirable that each and every discrepancy be resolved before submission of the complete dataset, the acceptable level of disagreement between the originally entered and double-entered data was established at 1 percent or less for questionnaire data and at 0.1 percent or less level for achievement data. Values above this level required resolution of the discrepancy and re-entry of data.

The level of disagreement between the originally entered and double-entered data was evaluated by IEA Hamburg, and it was found that in general the margin of error observed for processed data was well below the required threshold.

Data from Online Questionnaire Administration

As documented in Chapter 6: Survey Operations Procedures for TIMSS 2019, national centers had the option of administering the school, teacher, and home questionnaires online through IEA's Online SurveySystem instead of or in addition to using paper-based questionnaires. In addition, National Research Coordinators from participating countries completed the TIMSS 2019 Curriculum Questionnaire through this system.

To ensure confidentiality, national centers provided every respondent with a letter containing individual login information along with information on how to access the online questionnaire. This login information corresponded to the ID and checksum provided from WinW3S, meaning that the identity validation step occurring at the national centers for paper-based questionnaires occurred when the respondents' logged-in to the survey.

Online administration of questionnaires had a number of advantages. Because responses were collected in digital format and stored directly on the IEA Hamburg server, there was no need for data entry, reducing the workload for national centers. Also, the online system does not allow for inconsistent response patterns, meaning that the data collected had fewer inconsistencies when compared with data collected through the paper-based questionnaires. For example, if the directions ask the respondent to "Check one circle for each line," the system does not allow the respondent to check more than one response category on each line.

TIMSS & PIRLS
International Study Center
Lynch School of Education
BOSTON COLLEGE

CHAPTER 8: CREATING THE INTERNATIONAL DATABASE
METHODS AND PROCEDURES: TIMSS 2019 TECHNICAL REPORT   8.4

The TIMSS 2019 online questionnaires also include skip logic, which minimized response burden and improved data consistency. The TIMSS questionnaires have a number of questions that filter out respondents—meaning the subsequent questions are not applicable given the response to the filter question. For example, Question 10A of the eighth grade school questionnaire reads "Does your school have a school library? If no, go to Question 11." If a respondent chooses "No," the online survey skips directly to Question 11, omitting Questions 10B. Not only does the skip logic save the respondents' time, it also results in fewer inconsistencies in the data received by IEA Hamburg and instead produces planned missingness of the skipped responses which are coded in the final database as "not applicable."

Data Verification at the National Centers

Before sending the data to IEA Hamburg for further processing, national centers carried out mandatory validation and verification steps on all entered data and undertook corrections as necessary.

While the questionnaire data were being entered, the data manager or other staff at each national center used the information from the Teacher Tracking Forms to verify the completeness of the materials. Student participation information (e.g., whether a student participated in the assessment or was absent) was entered via WinW3S.

The validation process was supported by an option in WinW3S to generate an inconsistency report. This report listed all of the types of discrepancies between variables recorded during the within-school sampling and test administration process and made it possible to cross-check these data against data entered in the DME, the database for online respondents, and the uploaded student data on the central international server.

Data managers were requested to resolve such issues before final data submission to IEA Hamburg. If inconsistencies remained or the national center could not solve them, IEA Hamburg asked the center to provide documentation on these problems.

Upon submitting the validated data to IEA Hamburg, NRCs also provided extensive documentation including hard copies or electronic scans of all original Student and Teacher Tracking Forms, Student Listing Forms, and when applicable, a report on procedural activities collected as part of the online Survey Activities Questionnaire (see Chapter 6).

Data from eTIMSS Administration

As described in Chapter 4, the eTIMSS assessment was designed to run on PCs and tablets using USB or local server delivery. For both delivery methods, the student response data were stored in a SQLite database, the contents of which could be uploaded to the IEA Hamburg server immediately following the assessment, or later off-site. Following data upload, IEA Hamburg performed some pre-processing, verification, and cleaning steps and then student responses to constructed response items were sent to

the IEA Online Scoring System (IEA CodingExpert software), which almost immediately made student responses available to be allocated to scorers. Scoring took place directly on the IEA Hamburg server—allowing IEA Hamburg to monitor, in real time, the progress of scoring within countries.

Also available online to national centers was an upload monitor listing all the student records that had been uploaded to the IEA Hamburg server. In the rare cases that duplicate IDs were detected, the IDs were flagged and national centers indicated which record to keep. The data monitor also allows a list of IDs to be downloaded so that they can be used to update data availability status in WinW3S.

Description of eAssessment Data

The assessment Player recorded student item responses as well as other actions taken by the student and the data were stored in a SQLite database. Student actions were broken down into timestamped events that recorded process data such as navigation behavior and tool use, but also messages to the student that were created by the system (e.g. time remaining towards the end of the test). The events for process data and student responses were stored separately, with the events in an "event table" and item responses in a "response table." There were also auxiliary tables containing the student ID together with the language in which the assessment was administered and information about whether the data had already been uploaded to the IEA server. Other tables were used for error handling.

Each event or item response was stored with both general attributes and attributes specific to that response or event. The following general attributes were recorded:

- Two timestamp parts: The first recorded events and item responses in standard Unix time format and gave the elapsed time in seconds since January 1, 1970. Since a more precise time information was needed for event data, the second timestamp added the milliseconds.

- A sequential number recording the correct sequence of actions: This number reflected the exact order of events and responses and had to coincide with the sequence obtained using the timestamp information.

- A screen ID number: This number indicated the specific screen (or item) on which the response was saved, or the event occurred.

- A page identifier: Due to the rotation of item blocks within booklets, an item could be displayed in different positions in the assessment. Therefore, it was necessary to also include a "page number" as a general attribute.

- An item ID number: For recording responses, the item identifier referred to the particular item or item input (e.g., keyboard field) on the screen. This number corresponded to a given "raw variable name" specified by the TIMSS & PIRLS International Study Center.

- An event-type ID number: For recording events, using ID numbers instead of names helped to minimize data traffic during the assessment administration. A separate reference look-up table held the actual event names that corresponded to the event-type ID numbers.

- A response ID number: For recording responses, this identifier indicated if a response was changed later during the response process. It showed the sequential number (ID) under which the subsequent answer was saved. The final answer the student gave to an item was marked with a "NULL" value for this field.

Attributes specific to each event were stored as JavaScript Object Notation (JSON) objects. JSON objects in general hold for each attribute the name of the attribute (property) and the value of the property. Exhibit 8.1 shows an example extract from the event table for the "UI:IsLoaded" event type. This event indicates that the appropriate test form was loaded with the first item presented to the student. The event-specific attribute is the "index" which is set to zero for the first page of the test, stores as the JSON object {"index":0}.

Exhibit 8.1: Extract from the Event Table for Event Type "UI:IsLoaded"

Event-Type Id	Screen ID	Page Identifier	Information
26	13617	0	{"index":0}

In the response table, each response was stored in a separate record. The response table held the entire response history of each item the student worked on. All item responses were stored as one or more records with string of characters indicating the student response. This could be a single number, but also an extended string containing information about drawn lines or the dragging and dropping of objects. In addition, the student response table contained typed student responses that were later transferred to the Scoring System for human scoring, along with screenshot images of responses from the line-drawing tool. Responses that did not need human scoring were machine scored. For these responses, a set of detailed scoring rules provided by the TIMSS & PIRLS International Study Center were incorporated in a scoring algorithm and applied to each response to determine the appropriate score.

Pre-Processing and Scoring eTIMSS Data

Some pre-processing steps were required to prepare eAssessment data in a suitable format for scoring and further processing. Data uploads from the eTIMSS Players were processed at IEA Hamburg by several data servers that received and time stamped and then extracted the raw data from the uploaded SQLite databases into the "central" SQL database for all countries. This new structure contained a separate database for each country and grade, including all data from the original SQLite databases with the addition of identifiers relating to the import of data and additional fields for scoring purposes.

For eTIMSS countries, the new mode of administration allowed for a substantial portion of the digital items to be machine scored, particularly in mathematics. For eTIMSS items suitable for machine-scoring, the TIMSS & PIRLS International Study Center used scoring guides as the basis for developing machine scoring specifications that could be accurately applied without human judgment of student responses. Developing the machine scoring specifications involved testing each item in the eTIMSS Player, reviewing the output, and writing rules in terms of the output to classify all possible responses to a code in the item's scoring guide. The scoring unit at IEA Hamburg reviewed all specifications and provided feedback on an item-by-item basis, resulting in several rounds of revision until the rules for all items were clarified. The scoring unit at IEA Hamburg then applied the scoring rules for all machine-scored items and the data analysis team at the TIMSS & PIRLS International Study Center independently replicated the results to validate the scoring.

The IEA Scoring System (CodingExpert software) was used by NRCs and their scoring staff to score the eTIMSS constructed response items that were not suitable for machine scoring. Although scoring supervisors controlled the distribution of responses to scorers within countries, the responses themselves became available in the system soon after upload (with some delay due to the asynchronous handling of the import to the central database and thence to the scoring system). To avoid unnecessary scoring, therefore, it was essential that any duplicates in the central database were dealt with before import to the scoring system. In addition to measures to prevent a database from being uploaded a second time from the client side, checks were made to the results database creation date and content to ensure any possible duplicates were flagged before import. There were, however, some kinds of duplicate records that could be legitimate. Two databases with the same student ID but with different creation times could have originated in several scenarios. For example, this could be simply a case of the test administrator mistakenly using the same ID twice for two different students, or an interruption in the assessment may have led to part 1 being conducted from one USB stick and part 2 from a second. Such cases needed to be reconciled by IEA Hamburg's data processing procedures. When scoring was completed, the student response data were transferred to tables prepared for import into the data processing system (DPE) employed at IEA Hamburg for all large-scale international assessments.

Cleaning the International and National Databases

To ensure the integrity of the international database, a uniform data cleaning process was followed, involving regular consultation between IEA Hamburg and the NRCs. After each country had submitted its data, codebooks, and documentation, IEA Hamburg, in collaboration with the NRCs, conducted a four-step cleaning procedure upon the submitted data and documentation:

- A structural check
- A check of the identification (ID) variables
- Linkage cleaning
- Background cleaning

The data cleaning process included numerous iterations of the four-step cleaning procedure and were completed on each national data set in close collaboration with national centers. This repeated multi-step cleaning ensured that all data were properly cleaned and that any new errors that could have been introduced during the data cleaning were rectified. The cleaning process was repeated as many times as necessary until all data were made consistent and comparable. Any inconsistencies detected during the cleaning process were resolved in collaboration with national centers, and all corrections made during the cleaning process were documented in a cleaning report, produced for each country.

After the final cleaning iteration, each country's data were sent to Statistics Canada for the calculation of sampling weights, and then the data, including sampling weights, were sent to the TIMSS & PIRLS International Study Center so that the psychometric analyses (as described in Chapter 11 and Chapter 12) could be conducted. The NRCs were provided with interim data products to review at different points in the process.

Preparing National Data Files for Analysis

The main objectives of the data cleaning process were to ensure that the data adhered to international formats, that school, teacher, and student information could be linked across different survey files, and that the data reflected the information collected within each country in an accurate and consistent manner.

As illustrated in Exhibit 8.2, the program-based data cleaning consisted of a set of activities explained in the following subsections. IEA Hamburg carried out all of these activities in close communication with the national centers as well as with the TIMSS & PIRLS International Study Center for achievement data.
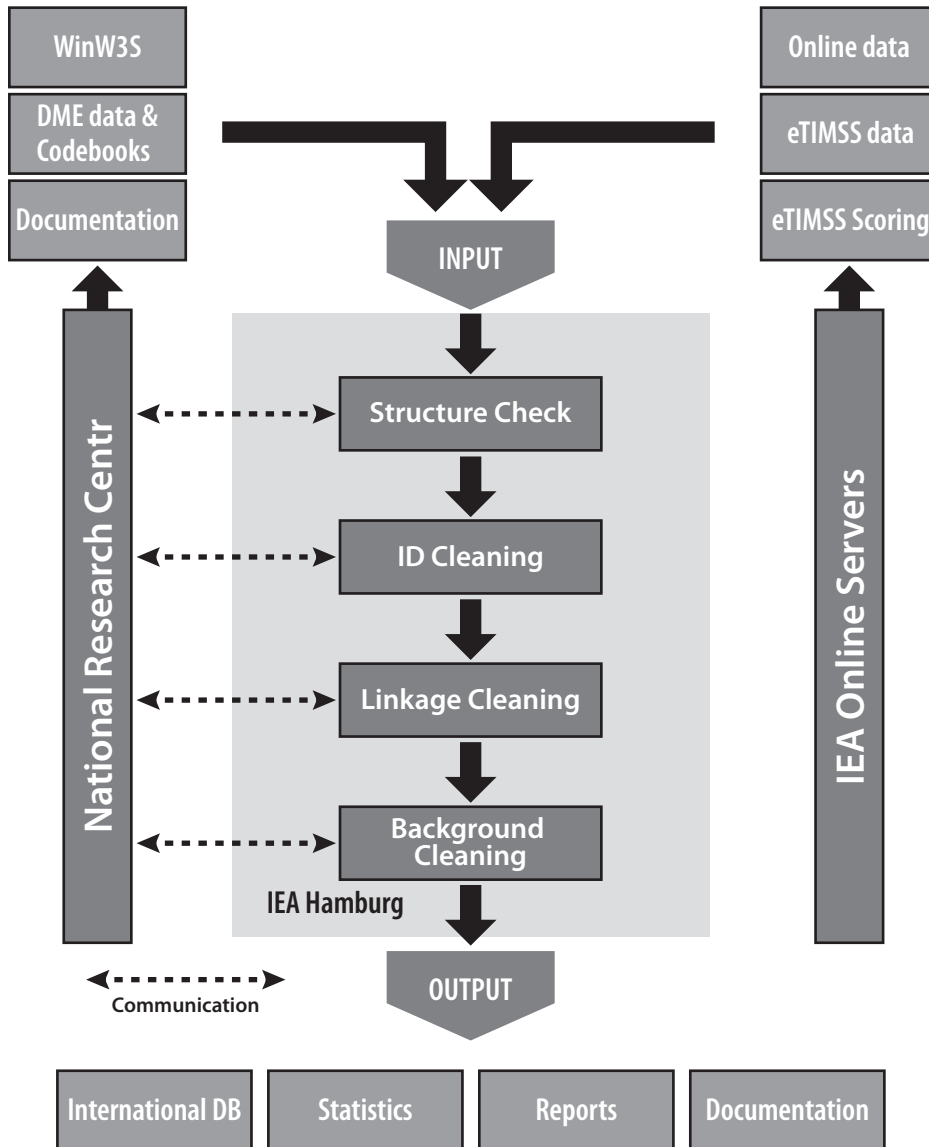
Checking Documentation, Import, and Structure

For each country, data cleaning began with a review of data file structures and its data documentation, including a review of National Adaptation Forms, Student Tracking Forms, Teacher Tracking Forms, Student-Teacher Linkage Forms, and the Survey Activities Questionnaire.

After the review, IEA Hamburg first merged the tracking information and sampling information captured in the WinW3S database with the student-level database containing the corresponding student

data from eTIMSS or paperTIMSS achievement assessments. During this step, IEA Hamburg staff also merged the data from the school and teacher questionnaires for both the online and paper modes of administration. At this stage, data from the different sources was transformed and imported into one SQL database so that this information would be available during all further data-processing stages.

The first checks identified differences between the international and the national file structures. Some countries made adaptations (such as adding national variables or omitting or modifying international variables) to their questionnaires. The extent and nature of these changes differed across countries: some countries administered the questionnaires without any modifications (apart from translations and necessary adaptations relating to cultural or language-specific terms), whereas other countries inserted response categories within existing international variables or added national variables.

To keep track of adaptations, staff at the TIMSS & PIRLS International Study Center asked the national centers to complete National Adaptation Forms. In their adaptations, countries sometimes modified the structure and values of the international codebooks, and if IEA Hamburg had to recode variables in the national data files to ensure that the resulting data remained comparable across countries. The national adaptation process is described in Chapter 5 and details about country-specific adaptations to the international instruments can be found in Supplement 2 of the *TIMSS 2019 User Guide for the International Database*.

IEA Hamburg then discarded variables created purely for verification purposes during data entry and made provision for adding new variables necessary for analysis and reporting, including reporting variables, derived variables, sampling weights, and scale scores.

Once IEA staff had ensured that each data file matched the international format, they applied a series of standard data cleaning rules for further processing. Processing during this step employed software developed by IEA Hamburg that identifies and corrects inconsistencies in the data. Each potential problem flagged at this stage was identified by a unique problem number, and then described and recorded in a database. The action taken by the cleaning program or IEA Hamburg staff with respect to each problem was also recorded.

IEA Hamburg referred problems that could not be rectified automatically to the responsible NRC so that national center staff could check the original data-collection instruments and tracking forms to trace the source of these errors. Wherever possible, staff at IEA Hamburg suggested a remedy and asked the national centers to either accept it or propose an alternative. If a national center could not solve the issue through verification of the instruments or forms, IEA Hamburg applied a general cleaning rule to the files to rectify the error. When all automatic updates had been applied, IEA Hamburg staff used SQL recoding scripts to directly apply any remaining corrections to the data files.

Checking Identification Variables

Each record in a data file needs to have a unique identification number. The existence of records with duplicate ID numbers in a file implies an error of some kind. Some countries administered the school, teacher, and home questionnaire (fourth grade only) online in addition to the paper mode. Therefore, by mistake a respondent could have completed both the paper and the online versions of the questionnaire. Similarly, it was possible for an eTIMSS login to be used (and uploaded) twice. If two records in a TIMSS 2019 database shared the same ID number and contained exactly the same data, IEA Hamburg deleted one of the records and kept the other one in the database. In the rare case that both records contained different data and IEA staff found it impossible to identify which record contained the more reliable or complete version of the data, national centers were asked which record to keep.

Although the ID cleaning covered all data from all instruments, it focused mainly on the student data file. In addition to checking the unique student ID number, it was crucial to check variables pertaining to student participation and exclusion status, as well as students' birth dates and dates of testing in order to calculate student age at the time of testing. The Student Tracking Forms provided an important tool for resolving anomalies in the database.

As mentioned previously, IEA Hamburg conducted all cleaning procedures in close cooperation with the national centers. After national center staff had cleaned the identification variables, they passed the clean databases with information about student participation and exclusion on to Statistics Canada, which used this information to calculate students' participation rates, exclusion rates, and student sampling weights.

Checking Linkages

As data on students, parents, teachers, and schools appeared in a number of different data files, a process of linkage cleaning was implemented to ensure that the data files would correctly link together. The linking of the data files followed a hierarchical system of identification codes that included school, class, and student components. These codes linked the students with their class and/or school membership. Further information on linkage codes can be found in Chapter 6: Survey Operations Procedures for TIMSS 2019.

Linkage cleaning consisted of a number of checks to verify that student entries matched across achievement files, student context questionnaire data files, scoring reliability files, and home background files. In addition, at this stage, checks were conducted to ensure that teacher and student records linked correctly to the appropriate schools. The Student Tracking Forms, Teacher Tracking Forms, and Student-Teacher Linkage Forms were crucial in resolving any anomalies. IEA Hamburg also liaised with NRCs about any problematic cases, and the national centers were provided with standardized reports listing all inconsistencies identified within the data.

Resolving Inconsistencies in Context Questionnaire Data

The amount of inconsistent and implausible responses in questionnaire data files varied considerably across countries. IEA Hamburg determined the treatment of inconsistent responses on a question-by-question basis, using all available documentation to make an informed decision. IEA Hamburg staff also checked all questionnaire data for consistency across the responses given. For example, Question 1 in the school questionnaire asked for the total school enrollment in all grades, and Question 2 asked for the enrollment in the target grade only. Logically, the number given as a response to Question 2 could not exceed the number provided by school principals in Question 1. Similarly, it is not possible that the number of years a teacher has been teaching altogether (Question 1 in the teacher questionnaires) exceeds the minimum possible age of a beginning teacher in all participating countries (Question 3 in the teacher questionnaires). IEA Hamburg flagged inconsistencies of this kind and then asked the national centers to review these issues. IEA staff recoded those cases that could not be corrected as "invalid."

Filter questions, which appeared in some questionnaires, directed respondents to a particular set of questions that only applies to a subset of respondents. IEA Hamburg applied the following cleaning rule to these filter questions and the dependent questions that followed, for instance: If a respondent answered "No" to Question 10A in the school questionnaire "Does your school have a school library?" IEA Hamburg recoded any responses to the dependent question 10B as "logically not applicable." Also, following the same example, if the filter question was omitted but at least one valid response was found in the dependent questions then IEA Hamburg recoded the filter question to "Yes." This of course is only possible for dichotomous filter questions (e.g., with response options such "Yes/No").

IEA Hamburg also applied what are known as split variable checks to questions where the answer was coded into several variables. For example, Question 5 in the student questionnaire asked students: "Do you have any of these things at your home?" Student responses were captured in a set of nine variables, each one coded as "Yes" if the corresponding "Yes" option was filled in and "No" if the "No" option was filled in. Occasionally, students checked the "Yes" boxes but left the "No" boxes unchecked. Because, in these cases, it was clear that the unchecked boxes meant "No," these responses were recoded accordingly.

In addition, student reports to items on gender and age in the student questionnaire were checked against the tracking information provided by the School Coordinator or Test Administrator during the within-school sampling and test/questionnaire administration process. When information on gender or birth year and month was missing in the student questionnaire but the student participated, this information, when available, was copied over from the tracking data to the questionnaire. If discrepancies were found between existing tracking and questionnaire gender and age data, IEA Hamburg queried the case with the national center, and the national center investigated which source of information was correct. If unresolved, tracking data was used rather than questionnaire data (or vice versa for some items at the eighth grade).

Handling of Missing Data

Two types of entries were possible during the TIMSS 2019 data capture: valid data values and missing data values. Missing data can be assigned a value of omitted/invalid, or not administered during data capture. IEA Hamburg applied additional missing codes to the data to facilitate further analyses. This process led to four distinct types of missing data in the international database:

- **Omitted or invalid:** The respondent had a chance to answer the question but did not do so, leaving the corresponding item or question blank. This code was also used if the response was uninterpretable or out-of-range.

- **Not administered:** This signified that the item or question was not administered to the respondent, which meant that the respondent could not read and answer the question. The not administered missing code was used for those student test items that were not in the set of assessment blocks administered to a student either deliberately (due to the rotation of assessment blocks) or, in rare cases, due to technical failure or incorrect translations. This missing code was also used for those records that were included in the international database but did not contain a single response to one of the assigned questionnaires. For example, this situation applied to home questionnaire data for students who participated in the student test but the parent/guardian did not answer the home questionnaire. In addition, the not administered code was used for individual questionnaire items that a national center decided not to include in the country-specific version of the questionnaire.

- **Logically not applicable:** The respondent answered a preceding filter question in a way that made the following dependent questions not relevant to him or her.

- **Not reached:** This applied only to the individual items of the student achievement test and indicated those items that students did not attempt due to a lack of time. "Not reached" codes were derived as follows: First, the last answer given by a student in a session is identified. This could be either a valid or invalid response to an item. The first omitted response after this last answer is coded as "omitted," but all following responses to these items in the session are then coded as "not reached." For example, the response pattern "1942999999" (where "9" represents "omitted") is recoded to "19429RRRRR" (where "R" represents "not reached").

Data Cleaning Quality Control

Because TIMSS 2019 was a large and highly complex study with very high standards for data quality, maintaining these standards required an extensive set of interrelated data checking and data cleaning procedures. To ensure that all procedures were conducted in the correct sequence, that no special

requirements were overlooked, and that the cleaning process was implemented independently of the persons in charge, the data quality control process included the following steps:

- Thorough testing of all data cleaning programs: Before applying the programs to real datasets, IEA Hamburg applied them to simulation datasets containing all possible problems and inconsistencies

- Registering all incoming data and documents in a dedicated database: IEA Hamburg recorded the date of arrival as well as specific issues requiring attention

- Carrying out data cleaning according to strict rules: Deviations from the cleaning sequence were not possible, and the scope for involuntary changes to the cleaning procedures was minimal

- Documenting all systematic data recoding that applied to all countries: IEA Hamburg recorded all changes to data in the comprehensive cleaning documentation provided to national centers

- Logging every "manual" correction to a country's data files in a recoding script: Logging these changes, which occurred only occasionally, allowed IEA Hamburg staff to undo changes or to redo the whole manual-cleaning process at any later stage of the data cleaning process

- Repeating, on completion of data cleaning for a country, all cleaning steps from the beginning: This step allowed IEA Hamburg to detect any problems that might have been inadvertently introduced during the data cleaning process

- Working closely with national centers at various steps of the cleaning process: IEA Hamburg provided national centers with the processed data files and accompanying documentation so that center staff could thoroughly review and correct any identified inconsistencies.

IEA Hamburg compared national adaptations recorded in the documentation for the national datasets with the structure of the submitted national data files. IEA Hamburg staff then recorded any identified deviations from the international data structure in the national adaptation database and for the supplementary materials provided with the *TIMSS 2019 User Guide for the International Database*. Whenever possible, IEA Hamburg recoded national deviations to ensure consistency with the international data structure.

Interim Data Products

Before the TIMSS 2019 International Databases were finalized, three major interim versions of the data files were sent to each country—each country receiving only its own data. In addition, countries that administered eTIMSS received files with student raw responses. These raw response files are the

trace of what students answered and are in this sense comparable to the completed paper booklets that paperTIMSS countries would have available for checking. The first version of the databases was sent as soon as the data could be considered "clean" as regards identification codes and linkage issues. Documentation, with a list of the cleaning checks and corrections made in the data, was included in the first sendout to enable the National Research Coordinators to review the cleaning process before the 7th NRC meeting Agadir, Morocco in December 2019. A second version of the data was sent to the countries when all national adaptations and the feedback resulting from the review of the first version were implemented at the end of February 2020. National Research coordinators were asked to confirm that the data is ready for the operational psychometric analysis used for achievement scaling. A third version of the data files was sent to countries when the weights and international achievement scores were available and had been merged with the data files. This version, sent to the countries in advance of the 8th NRC Meeting in June 2020 contained only those records that were used in the analysis and reports to be released in December 2020 and satisfied the sampling standards, allowed the NRCs to replicate the results presented in the international reports.

Interim data products were accompanied by detailed data processing and national adaptation documentation, codebooks, and summary statistics. The summary statistics were created by the TIMSS & PIRLS International Study Center and included weighted univariate statistics for all questionnaire variables for each country. For categorical variables, representing the majority of variables, the percentages of respondents choosing each of the response options were displayed. For continuous numeric variables, various descriptive statistics were reported, including the minimum, maximum, mean, median, mode, and percentiles. For both types of variables, the percentages of missing data were reported. Additionally, for the achievement items, the TIMSS & PIRLS International Study Center provided item analysis and reliability statistics listing information regarding the number of valid cases, percentages, percentage correct, Rasch item difficulty, scoring reliability, and so forth. These statistics were used for a more in-depth review of the data at the international and national levels in terms of plausibility, unexpected response patterns, etc. More information on item almanacs and reviewing item statistics is available in Chapter 10.

Final Product—the TIMSS 2019 International Databases

The data cleaning effort implemented at IEA Hamburg ensured that the TIMSS 2019 international databases contained high-quality data. More specifically, the process ensured that:

- Information coded in each variable was internationally comparable
- National adaptations were reflected appropriately in all variables

TIMSS & PIRLS
International Study Center
Lynch School of Education
BOSTON COLLEGE

IEA

CHAPTER 8: CREATING THE INTERNATIONAL DATABASE
METHODS AND PROCEDURES: TIMSS 2019 TECHNICAL REPORT 8.16

- All entries in the database could be successfully linked within and across levels

- Sampling weights and student achievement scores were available for international comparisons.

Supplements to the *TIMSS 2019 User Guide for the International Database* document all national adaptations made to questionnaires by individual countries and how they were handled in the data. The description of country-specific items also can be found in this supplement, as well as recoding requirements by the TIMSS & PIRLS International Study Center.

TIMSS & PIRLS
International Study Center
Lynch School of Education
BOSTON COLLEGE