

Chapter 6

Quality Assurance in the TIMSS Advanced 2008 Data Collection

Ieva Johansone

6.1 Introduction

To ensure the quality of the TIMSS Advanced 2008 data, considerable effort was made to develop standardized materials and survey operations procedures (see Chapter 5 for more information). In addition, the TIMSS & PIRLS International Study Center developed an ambitious international quality control program to document data collection activities in the participating countries. To implement this program, the IEA Secretariat, in cooperation with national centers, nominated an international Quality Control Monitor (QCM) in each of the participating countries.

The TIMSS & PIRLS International Study Center conducted a QCM training program that involved introducing the QCMs to the TIMSS Advanced 2008 survey operations procedures, the design of the test booklets, and the background questionnaires. During the training, each QCM received the materials necessary to complete their monitoring tasks. The materials included a copy of



the *TIMSS Advanced 2008 International Quality Control Monitor Manual* (TIMSS & PIRLS International Study Center, 2007a), *Classroom Observation Record*, *TIMSS Advanced 2008 Survey Operations Procedures Units 2 and 3 Manuals* (TIMSS & PIRLS International Study Center, 2007d, 2007e), *TIMSS Advanced 2008 School Coordinator Manual* (TIMSS & PIRLS International Study Center, 2007c), and *TIMSS Advanced 2008 Test Administrator Manual* (TIMSS & PIRLS International Study Center, 2007f).

The major task of the international QCMs was to conduct site visits to a random sample of 24 participating schools (12 per subject, which was about 10 percent of the sampled schools). During the site-visits, the QCMs observed TIMSS Advanced classroom testing sessions and discussed the TIMSS Advanced test administration with the Test Administrator or School Coordinator. Where necessary, the QCMs were permitted to recruit one or more assistants in order to efficiently cover the territory and testing timetable. A total of 28 international QCMs and their assistants were trained across the 10 countries that participated in TIMSS Advanced 2008. Altogether, these monitors observed 200 testing sessions. The results of the QCM observations are reported in Section 6.2.

In addition to the international quality control program, the National Research Coordinators (NRCs) were asked to complete the *Survey Activities Questionnaire* about their experiences with the TIMSS Advanced 2008 survey operations procedures and the quality of the assessment materials.

6.2 Quality Control Observations of the TIMSS Advanced 2008 Test Administration

For each testing session observed, QCMs completed the TIMSS Advanced 2008 Classroom Observation Record. The observation record was organized into four sections, listed below.

Section A: Preliminary Activities of the Test Administrator

Section B: Test Administration Activities

Section C: Summary Observations

Section D: Interview with the School Coordinator and/or Test Administrator

6.2.1 Preliminary Activities of the Test Administrator

Section A of the Classroom Observation Record addressed the extent to which the Test Administrator had prepared for the testing session. QCMs were asked to note the following activities of the Test Administrator: checking the testing materials, reading the administration script, organizing space for the session, and arranging for the necessary equipment.

Exhibit 6.1 summarizes the results for Section A. In nearly all testing sessions, Test Administrators observed the proper preparatory procedures. For those few deviations that occurred, QCMs provided reasonable explanations. For example, QCMs noted that Test Administrators who did not have a watch with a second hand had a cell phone clock or a classroom clock available to monitor the time remaining in the test session. In general, no procedural deviations were observed that were severe enough to jeopardize the integrity of the test administration.

Exhibit 6.1 Percentages of QCM Responses for Preliminary Activities of the Test Administrator

Question	Yes (%)	No (%)	Not Answered (%)
Had the Test Administrator verified adequate supplies of the test booklets were available?	95	5	0
Had the Test Administrator familiarized himself or herself with the test administration script prior to the testing?	94	6	0
Did the student identification information on the test booklets and student questionnaires correspond with the Student Tracking Form?	99	1	0
Was there adequate seating space for the students to work without distractions?	92	8	0
Was there adequate room for the Test Administrator to move around during the testing to ensure that student were following directions correctly?	93	7	0
Did the Test Administrator have a watch with a seconds hand (or a stopwatch) for timing the testing sessions accurately?	94	6	0

6.2.2 Assessment Session Activities

Section B of the Classroom Observation Record addressed the activities that took place during the assessment session and the administration of the student questionnaire. The activities, such as following the Test Administrator script, distributing and collecting test booklets, and making announcements during the testing session were reported by the QCMs and are presented in Exhibits 6.2 through 6.4.

Activities carried out during the achievement testing session are presented in Exhibit 6.2. One of the most important ways to standardize assessment administrations was following the Test Administrator's script. QCMs reported that, in almost all of their observations, the Test Administrators followed the script exactly when preparing students, distributing test materials, and giving directions. Of the few changes made, most were considered minor. Typically, the script changes were additions rather than revisions or deletions. In

the 5 percent of the sessions where the testing time did not equal the time allowed, it was because students had completed the test before the allotted time had elapsed. When the allotted time was over, the Test Administrator instructed students to complete the calculator survey. In some cases, students had completed the test and the calculator survey before the testing time was over. In 98 percent of the cases, the Test Administrator made sure that students stopped working immediately after announcing the end of the testing session.

Exhibit 6.2 Percentages of QCM Responses for the Assessment Session

Question	Yes (%)	No (%)	Not Answered (%)
Did the Test Administrator follow the Test Administrator's script exactly in each of the following tasks?			
Preparing the students	85	14 (Minor changes) 1 (Major changes)	0
Distributing the materials	91	7 (Minor changes) 1 (Major changes)	1
Giving directions	81	15 (Minor changes) 2 (Major changes)	2
<i>If the Test Administrator made changes to the script, how would you describe them?</i>			
Additions	16	7	1 (Not Answered) 76 (Not Applicable)
Revisions	8	15	1 (Not Answered) 76 (Not Applicable)
Deletions	5	17	2 (Not Answered) 76 (Not Applicable)
Did the Test Administrator distribute the test booklets according to the booklet assignment on the Student Tracking Form?	99	0	1
Did the Test Administrator record attendance correctly on the Student Tracking Form?	93	0	7
Was the total time for testing correct as indicated in the script?	94	5	1
Did the Test Administrator announce, "You have 10 minutes left" prior to the end of the testing session?	94	6	0
Were there any other time-remaining announcements made during testing session?	10	90	0
Did the Test Administrator read the script to announce the end the testing session and to ask students to complete the Calculator Survey?	90	10	0
At the end of the testing session, did the Test Administrator make sure all students had closed booklets and stopped working?	98	2	0
Were all booklets collected and secured after the testing session?	99	1	0

After the calculator survey was completed, the Test Administrators announced a break to be followed by the student questionnaire, unless the questionnaire was to be administered on a different date. In such cases, QCMs were not required to observe the questionnaire administration. In 85 percent of all observations, the student questionnaire was administered after a short break following the assessment session.

In most cases, Test Administrators kept to the testing script for signaling a break before administering the student questionnaire. Of those who did make changes, only 3 percent were reported as major changes. There were no observations of students requesting additional time to complete the student questionnaire.

Exhibit 6.3 Percentages of QCM Responses for the Student Questionnaire Administration

Question	Yes (%)	No (%)	Not Answered (%)
When the Test Administrator read the script to end the testing session and the calculator survey, did he/she announce a break to be followed by the student questionnaire?	85	13	2
Did the Test Administrator accurately read the script to end the testing and signal a break?	78	13 (Minor changes) 3 (Major changes)	6
<i>If there were changes, how would you describe them?</i>			
<i>Additions</i>	7	8	8 (Not Answered) 77 (Not Applicable)
<i>Revisions</i>	8	7	8 (Not Answered) 77 (Not Applicable)
<i>Omissions</i>	5	9	9 (Not Answered) 77 (Not Applicable)
Did the Test Administrator distribute the student questionnaires and give directions as specified in the script?	85	5	10
Did the students ask for additional time to complete the questionnaire?	0	93	7
At the end of the testing session, prior to dismissing the students, did the Test Administrator thank the students for participating in the study?	87	11	2

Exhibit 6.4 provides observations on student compliance with instructions and the alignment of the scripted instructions with their implementation. The results show that in all of the sessions, students complied well or very well with the instructions to stop working. Almost always, the dismissal of students had been very orderly or somewhat orderly.

Exhibit 6.4 Percentages of QCM Responses for Student Cooperation

Question	Very Well (%)	Well (%)	Fairly Well (%)	Not Well at All (%)	Not Answered (%)
When the Test Administrator ended the testing session, how well did the student comply with the instruction to stop working?	76	23	0	0	1

Question	Very Orderly (%)	Somewhat Orderly (%)	Not Orderly at All (%)	Not Answered (%)
How orderly was the dismissal of the students?	82	16	1	1

6.2.3 General Observations

Section C of the Classroom Observation Record referred to the general observations by QCMs during the testing sessions, including their overall impressions of the test administration, how well the Test Administrator monitored students, and any unusual circumstances that arose during the testing session (e.g., student refusal to participate, defective instrumentation, emergency situations, and cheating).

The results presented in Exhibits 6.5 and 6.6 show that, for most testing sessions, no problems were observed. In 99 percent of all cases, Test Administrators addressed students' questions adequately and as instructed in the *Test Administrator Manual*. In the very few sessions where a defective test instrument was detected, the Test Administrator replaced the instrument appropriately. There were no cases where a student refused to take the test. In 22 percent of the observed testing

sessions, a student left the room for an “emergency” during the testing session. In such cases, Test Administrators were instructed to collect the student’s test booklet, and give it back after he or she returned. However, in a few cases, students were instructed to close their booklets and leave them on their tables while out of the classroom.

QCMs reported no cases where students were not orderly and cooperative during the testing sessions. There were very few cases where students’ orderliness or cooperation was less than perfect or very good. In all such cases, Test Administrators managed to control the situation. QCMs reported that the overall quality of all testing sessions was fair, good, very good, or, in 56 percent of the cases, excellent.

Exhibit 6.5 Percentages of QCM Responses for General Observations

Question	Yes (%)	No (%)	Not Answered (%)
Did the Test Administrator address students’ questions appropriately?	99	1	0
Did you see any evidence of students attempting to cheat on the tests (e.g., by copying from a neighbor)?	3	97	0
Were any defective test books detected and replaced before the testing began?	1	99	0
Were any defective test books detected and replaced after the testing began?	1	99	0
<i>If any defective test books were replaced, did the Test Administrator replace them appropriately?</i>	1	1	0 (Not Answered) 98 (Not Applicable)
Did any students refuse to take the test either prior to the testing or during the testing?	0	100	0
<i>If a student refused, did the Test Administrator accurately follow the instructions for excusing the student (collect the test book and record the incident on the Student Tracking Form)?</i>	0	0	1 (Not Answered) 99 (Not Applicable)
Did any students leave the room for an “emergency” during the testing?	22	78	0
<i>If a student left the room for an emergency during the testing, did the Test Administrator address the situation appropriately (collect the test booklet, and if re-admitted, return the test booklet)?</i>	18	4	0 (Not Answered) 78 (Not Applicable)

Exhibit 6.6 Percentages of QCM Responses for Observations of Student Behavior

Question	Extremely (%)	Moderately (%)	Somewhat (%)	Hardly (%)	Not Answered (%)
To what extent would you describe the students as orderly and cooperative?	67	28	5	0	0

	No, There Were No Late Students (%)	No, They Were Not Admitted (%)	Yes, Before Testing Began (%)	Yes, After Testing Began (%)	Not Answered (%)
Were any late students admitted to the testing room?	84	3	8	4	1

	Excellent (%)	Very Good (%)	Good (%)	Fair (%)	Poor (%)	Not Answered (%)
In general, how would you describe the overall quality of the testing session?	56	29	10	5	0	0

6.2.4 Interview with the Test Administrator or School Coordinator

As the final step of each observation, the QCMs conducted an interview with the Test Administrator or School Coordinator. Details of the interview were recorded in Section D of the Classroom Observation Record. The interview addressed activities such as shipment of assessment materials, arrangements for test administration, responsiveness of the NRC to queries, necessity for make-up sessions, and, as a validation of within-school sampling procedures, the organization of classes in the school.

The results, presented in Exhibits 6.7 and 6.8, show that, overall, School Coordinators considered the TIMSS Advanced 2008 administration in their schools a success. Mistakes that did occur tended to be minor and easily remedied. There were only a few cases where shipments of test materials had something missing; and, in all such cases, they were resolved before the testing date. Note that the relatively high percentages of School Coordinators or Test

Administrators not responding to questions about receiving some of the items in Exhibit 6.7 occurred because some countries did not use the particular form or shipment method.

Exhibit 6.7 Receipt of Materials and Test Administration, Percentages of Responses from QCM Interviews with the Test Administrator and/or School Coordinator

Question	Yes (%)	No (%)	Not Answered (%)
Prior to the assessment day did you have time to check your shipment of materials from your TIMSS Advanced National Research Coordinator?	96	3	1
Did you receive the correct shipment of the following items?			
School Coordinator Manual	87	12	1
Test Administrator Manual	99	0	1
Student Tracking Forms	99	0	1
Test booklets	97	1	2
Student Questionnaires	99	0	1
Teacher Questionnaires	99	0	1
School Questionnaire	99	0	1
Test Administration Form	99	0	1
Teacher Tracking Form	80	1	19
Envelopes or boxes addressed to the National Center for the purpose of returning the materials after the assessment	82	4	14
Was the National Research Coordinator responsive to your questions or concerns?	96	1	3
Was the estimated time of 45 minutes to complete the teacher questionnaire a correct estimate?	54	5 (Took longer) 11 (Took less time)	30
Were you satisfied with the accommodations (testing room) you were able to arrange for the testing?	86	4	10
Do you anticipate that a makeup session will be required at your school?	25	62	13
<i>If you anticipate a makeup session, do you intend to conduct one?</i>	16	9	13 (Not Answered) 62 (Not Applicable)
Did the students receive any special instructions, a motivational talk, or incentives to prepare them for the assessment?	65	21	14
Is this a complete list of the classes in this grade in this school?	79	4	17
To the best of your knowledge, are there any students in this grade level who are not in any of these classes?	2	91	7
To the best of your knowledge, are there any students in this grade level in more than one of these classes?	1	92	7
If there was another international assessment, would you be willing to serve as a school coordinator?	84	7	9

In order to better estimate the time needed to complete the teacher questionnaires, QCMs asked if the current estimate of 45 minutes was appropriate. From all cases where teacher questionnaires already were completed, 54 percent of the School Coordinators reported that the estimate of 45 minutes was about right. Five percent reported that the questionnaires took longer, and 11 percent said that they took less time to complete. In more than half the cases, School Coordinators indicated that students were given special instructions, motivational talks, or incentives by a school official or the classroom teacher prior to testing. In 25 percent of the observed classes, the School Coordinator anticipated that a make-up session was needed, and most of them were sure that they would be conducting one.

Because sampling classes required a complete list of all classes in the school at the target grade, QCMs were asked to verify that the class list did indeed include all classes. In spite of complicated course structures in some countries, almost all School Coordinators reported that the complete list of classes had been documented, and that all students appeared in one and only one of these classes. Additional comments from School Coordinators showed that some were very confused by the question: commenting, for example, that the selected class was the only eligible one in the school. Thus, a relatively high percentage of QCMs did not answer the question, marking it as not applicable.

A tribute to the planning and implementation of TIMSS Advanced 2008 was the fact that 84 percent of respondents said they would be willing to serve as a School Coordinator in future international assessments. Furthermore, the results in Exhibit 6.8 suggest that the majority of School Coordinators believed that the testing sessions went very well and that school staff members had mostly positive attitudes towards the TIMSS Advanced 2008 testing.

Exhibit 6.8 Overall Impressions, Percentages of Responses from QCM Interviews with the Test Administrator and/or School Coordinator

Question	Very Well, No Problems (%)	Satisfactorily, Few Problems (%)	Unsatisfactorily, Many Problems (%)	Not Answered (%)
Overall, how would you say the session went?	86	12	1	1

	Positive (%)	Neutral (%)	Negative (%)	Not Answered (%)
Overall, how would you rate the attitude of the other school staff members towards the survey?	77	17	2	4

	Worked Well (%)	Needs Improvement (%)	Not Answered (%)
Overall, do you feel the School Coordinator Manual worked well or does it need improvement?	79	3	18

6.3 Survey Activities Questionnaire

The Survey Activities Questionnaire was designed to elicit information about NRCs experiences in preparing for and conducting the TIMSS Advanced 2008 data collection, with a focus on identifying and selecting samples, translating test instruments, assembling and printing test materials, packing and shipping the test materials, scoring constructed-response items, entering and verifying data, and implementing the national quality assurance program.

This section reports information gathered from the Survey Activities Questionnaire, reflecting the quality of the TIMSS Advanced 2008 survey materials and procedures in the participating countries. All participating countries, except one, completed the Survey Activities Questionnaire.

6.3.1 Sampling

The first part of the Survey Activities Questionnaire asked about sampling schools and classes within the sampled schools. None of the

participating countries reported problems in selecting their samples using the manuals provided by the TIMSS & PIRLS International Study Center. Two countries did not use the Windows Within-school Sampling Software (WinW3S) provided by the IEA DPC to select classes. One country chose to use their own software, because they felt their experience using this software would make the process more efficient. The other country did not have to sample classes because the survey was administered to all classes within the sampled schools.

Two NRCs encountered organizational constraints in their systems that necessitated a deviation from the sample design. In each case, the Statistics Canada sampling expert was consulted to ensure that the altered design remained compatible with TIMSS Advanced 2008 standards.

6.3.2 Preparing the Survey Instruments

In translating the survey instruments, NRCs reported mostly using a combination of their own staff and outside experts. All NRCs reported that they had gone through the process of external translation verification (organized by the IEA Secretariat) of the achievement materials and background questionnaires.

The NRCs were asked to answer some questions about assembling and printing the survey instruments, as well as issues related to checking the materials and securely storing them. All NRCs answered that they were able to assemble the achievement booklets according to the instructions provided, and only one country did not go through the process of an external review of instrument layout by the TIMSS & PIRLS International Study Center. Nearly all countries conducted the recommended quality control checks during the printing process, with missing pages being the most common in detected errors. NRCs were able to fix all of the systematic errors before sending the tests for administration.

All countries reported that they followed procedures to protect the security of the tests during assembly and printing. None of the NRCs were concerned that there might have been a breach of security.

Some questions in the questionnaire addressed the extent to which NRCs detected errors in the testing materials as they were packed for shipping to School Coordinators. A few errors were found in the materials. All errors that were discovered before distribution were remedied. In cases where errors were found after distribution, they were mainly minor and remedied by School Coordinators. In more severe cases, the provided replacement materials were used.

6.3.5 Scoring Constructed-response Items

The Survey Activities Questionnaire collected information from NRCs about preparation for scoring and scoring the constructed-response items. The scoring process was an ambitious effort, requiring recruiting and training scoring staff to score student responses including independent double scoring to verify scoring reliability. All NRCs reported that they understood the procedures of within-country reliability scoring, and only one country had difficulty understanding procedures of cross-country reliability scoring, as explained in the manuals provided by the TIMSS & PIRLS International Study Center.

Most countries used their own staff, as well as teachers and university students to score the constructed-response items. All countries used the TIMSS Advanced 2008 scoring training materials provided by the TIMSS & PIRLS International Study Center to train their scorers. Countries used anywhere between 6 and 31 scorers, and all scoring activities took up to 30 days to complete.

6.3.6 Data Entry and Verification

Within the section on entering and submitting the TIMSS Advanced 2008 data of the Survey Activities Questionnaire, NRCs reported that

they mainly used their own staff or a combination of their own staff and an external data entry company. One country used university students to enter its data.

All countries, except one where the survey instruments were scanned, used the Windows Data Entry Manager Software (IEA, 2008) to enter the TIMSS Advanced 2008 data. Most countries entered the data from a percentage of all the survey instruments twice as a verification procedure. The estimated proportion of survey instruments to be entered twice ranged from 5 to 30 percent. All NRCs reported establishing a secure storage area for the returned tests after data entry.

6.3.7 National Quality Assurance Program

As part of the national quality assurance activities, NRCs were required to send National Quality Control Observers to 10 percent of the participating schools in order to observe the test administration and document compliance with prescribed procedures. The last section of the Survey Activities Questionnaire addressed preparation for and implementation of the national quality assurance program.

In carrying out the national quality assurance program, only one national center did not use the *National Quality Control Monitor Manual* provided by the TIMSS & PIRLS International Study Center as a guide. The on-site quality control observations were conducted either by an external agency, members of the national center, a combination of the two, or, in some cases, other professionals, such as inspectors, retired teachers, mathematics and science supervisors, or ministry representatives.

References

IEA (2008). *Windows data entry manager software (WinDEM)* [Computer software and manual]. Hamburg: IEA Data Processing and Research Center.

TIMSS & PIRLS International Study Center. (2007a). *TIMSS Advanced 2008 international quality control monitor manual*. Chestnut Hill, MA: TIMSS & PIRLS International Study Center, Boston College.

TIMSS & PIRLS International Study Center. (2007b). *TIMSS Advanced 2008 national quality control monitor manual*. Chestnut Hill, MA: TIMSS & PIRLS International Study Center, Boston College.

TIMSS & PIRLS International Study Center. (2007c). *TIMSS Advanced 2008 school coordinator manual*. Chestnut Hill, MA: TIMSS & PIRLS International Study Center, Boston College.

TIMSS & PIRLS International Study Center. (2007d). *TIMSS Advanced 2008 survey operations procedures unit 2: Contacting schools and sampling classes for the TIMSS Advanced 2008 assessment*. Chestnut Hill, MA: TIMSS & PIRLS International Study Center, Boston College.

TIMSS & PIRLS International Study Center. (2007e). *TIMSS Advanced 2008 survey operations procedures unit 3: Preparing materials for the TIMSS Advanced 2008 data collection and administering the assessment*. Chestnut Hill, MA: TIMSS & PIRLS International Study Center, Boston College.

TIMSS & PIRLS International Study Center. (2007f). *TIMSS Advanced 2008 test administrator manual*. Chestnut Hill, MA: TIMSS & PIRLS International Study Center, Boston College.



