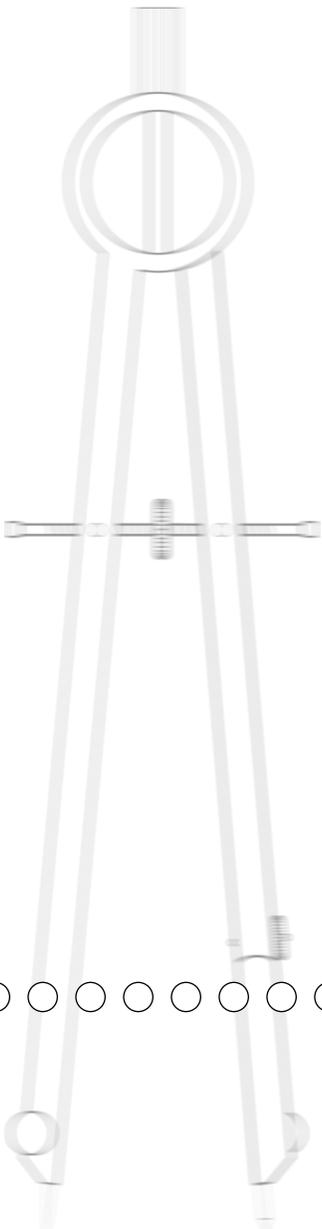




Quality Control in Data Collection for TIMSS 1999 Benchmarking

Kathleen M. O'Connor
Steven E. Stemler







9

Quality Control in Data Collection for TIMSS 1999 Benchmarking

Kathleen M. O'Connor
Steven E. Stemler

9.1 Overview

To verify that the TIMSS international data-collection procedures were applied uniformly in each of the benchmarking jurisdictions, the International Study Center instituted a program for quality assurance in data collection¹. Quality Control Monitors (QCMs) were recruited by the International Study Center to document procedures in a sample of schools in each participating Benchmarking entity. The International Study Center selected approximately five schools in each state and two or three schools in each district/consortium to take part in the quality control program.

The major responsibility of Benchmarking QCM was to observe the TIMSS test administration in selected schools. QCMs were assigned to schools and completed a Classroom Observation Record documenting test administration procedures for each session they observed.

In preparation for their task, QCMs were given an overview of the TIMSS 1999 survey operations procedures and were trained by the staff of the International Study Center in how to conduct and document their quality control task. In order to facilitate the training, the TIMSS International Study Center developed a manual to inform QCMs about the TIMSS Benchmarking project and to describe in detail their roles and responsibilities.

QCMs were provided with the following materials to conduct their task: the *Manual for Quality Control Monitors* (TIMSS, 1998), the *TIMSS 1999 Benchmarking Test Administrator's Manual* (Westat, 1999), the Classroom Observation Forms, the list of schools selected for site visits, and contact information for the region's Westat supervisor.

○○○

1. See O'Connor, K.M., and Stemler, S.E. (2000) for information about the international quality control effort, which documented the data collection in the TIMSS 1999 countries.

Before beginning their task, QCMs assembled in Boston for a training meeting. At the meeting, QCMs were trained to do the following:

- Contact the supervisors to gather information about sites being observed, including the contact information for the school coordinator
- Contact the school coordinators to explain the site visits
- Review the condition of the achievement test booklets prior to testing
- Collect and review the tracking forms being used to sample students and to document participation status
- Observe the testing sessions
- Interview the school coordinator
- Document findings.

In total, 18 Benchmarking quality control monitors were recruited and trained. They observed a total of 98 testing sessions.

9.2 Observing the TIMSS Test Administration

The Classroom Observation Record was designed to allow the QCM to keep a simple and accurate record of the major activities relating to the test administration. The record had four sections:

1. Preliminary activities of the test administrator
2. Test session activities
3. General impressions
4. Interview with the school coordinator.

9.2.1 Preliminary Activities of the Test Administrator

Section A of the Classroom Observation Record dealt with preparations for the testing session. Monitors were asked to note whether the test administrator had checked the testing materials, read the administration script, organized space for the session, and arranged for the necessary equipment (pencils, timers, etc.).

Exhibit 9.1 summarizes the results for this section. It shows that in almost all cases, the preparatory testing procedures were followed. In the rare instances where deviations occurred, appropriate corrections were made. In the few instances where QCMs reported a discrepancy between information on the Student

Tracking Form and information listed on the Student Identification Form, the errors were usually limited to one student in the group and consisted of a mismarking of the student's gender or a mismarked digit on the student ID number.

In the few cases where it was reported that there was not enough room for students, QCMs noted that this was due to unavoidable circumstances (e.g., the test was administered in a small classroom, the desks were too narrow, the students sat at round tables).

The absence of a visible wall clock was also considered an environmental restriction more than a limitation of the implementation of the testing procedures. In many of the cases the room had a clock, but not all students were able to see it.

In general, QCMs observed no procedural deviations in preparations for the testing that were severe enough to compromise the integrity of the test administration.

Exhibit 9.1 Preliminary Activities of the Test Administrator

Question	Yes	No	N/A
Had the test administrator verified adequate supplies of the test booklets?	97 ⁺	1 ^{**}	-
Had the test administrator familiarized himself or herself with the script prior to testing?	94 ⁺	4 ^{**}	-
Were all the seals intact on the test booklets prior to distribution?	41	1	55 ⁺
Did the Student Identification information on test booklet correspond with the Student Tracking Form?	85	10	3
Was there adequate seating space for the students to work without distractions?	85	13	-
Was there adequate room for the test administrator to move about the room during testing?	93	5	-
Did the test administrator have a stopwatch or timer for accurately timing testing sessions?	96	2	-
Did the test administrator have an adequate supply of pencils and other materials?	97	-	1
Was there a wall clock visible for the students to check their timing during the testing?	83	14	1

⁺ Seals were not used on the booklets in these states, districts, or consortia

^{*} Represents the number of respondents answering either Definitely Yes or Probably Yes

^{**} Represents the number of respondents answering either Definitely No or Probably No

9.2.2 Test Session Activities

Section B of the classroom observation record dealt with the test session activities themselves. These included the extent to which the test administrator followed the script, how the test booklets were distributed and collected, and the various announcements made during the testing session.

The achievement test was administered in two sessions, with a short break between. Exhibit 9.2 documents the activities associated with the first testing session and shows that at least 70% of the test administrators followed their script exactly when preparing the students and delivering instructions for Session 1. Where changes were made, they tended to be additions to the script.

Further examination of Exhibit 9.2 shows that in more than 75% of the sessions, the test administrator collected booklets one at a time from students. In the remaining sessions, students laid their booklets down on their desks during a brief 1 to 2 minute break.

Note that in 35 of 98 testing sessions (36%), the length of the testing session did not equal the time allowed. In each instance, all students had finished early.

Finally, booklets were rarely collected at the end of Session 1; rather, students were given a very short 1 to 2 minute break while the books remained on their desks.

Exhibit 9.2 Test Administrator's Activities—Testing Session 1

Question	Yes	No	N/A
Did the test administrator follow the test administrator's script exactly in...			
...preparing the students?	72	23 (minor changes) 2 (major changes)	1
...distributing the materials?	58	30 (minor changes) 10 (major changes)	-
...giving General Directions?	62	30 (minor changes) 6 (major changes)	-
...giving instructions for Part I?	79	11 (minor changes) 8 (major changes)	-
If the test administrator made changes to the script, would you describe them as...			
...additions?	39	20	39
...revisions?	24	28	46
...deletions?	21	23	54
Did the test administrator distribute test booklets one-at-a-time to students?	76	22	-
Did the test administrator distribute the test booklets according to the booklet assignment on the Student Tracking Form?	94	4	-
Did the test administrator record attendance correctly on the Student Tracking Form?	91	1	6
Did the total testing time for Session 1 equal the time allowed?	63	35	-
Did the test administrator announce "you have 10 minutes left" prior to the end of Session 1?	94	4	-
Were any other "time remaining" announcements made during Session 1?	10	88	-
At the end of Session 1, did the test administrator collect the test booklets one at a time from students?	2	96	-

Exhibit 9.3 summarizes QCMs' observations from the second testing session. The amount of time it took to restart the testing sessions ranged from 0 to 23 minutes; however, the vast majority of sessions were restarted in five minutes or less. In fact, because booklets were rarely collected during the break, testing typically resumed in 1-2 minutes.

Exhibit 9.3 shows that in about 30% of the testing sessions, the time used for Session 2 was less than the full time allowed. The test administrators reported that all of the students in those sessions had finished the exam early, had finished reviewing their work, and in many cases “were becoming unruly and impatient.”

In 47 of 98 sessions (48%), booklets were collected one at a time from students. When they were not, students were simply asked to pass them to the front of their rows.

Exhibit 9.3 also reveals that in about two-thirds of the sessions observed, no break was given between testing and the administration of Student Questionnaires. Administrators often reported that students already had time to rest due to the fact that students often finished testing early. As a result, no official break preceded administration of the Student Questionnaires at many of the observed testing sessions.

A final statistic from Exhibit 9.3 worth noting is that in about two-thirds of the testing sessions, students requested additional time to complete the Student Questionnaire. It was almost always the case that these students were given an extra 5-10 minutes to complete the questionnaires.

Exhibit 9.3 Test Administrator's Activities—Testing Session 2

Question	Yes	No	N/A
Was the time spent to restart the testing in Session 2 equal to 5 minutes?	1	97	-
Did the total testing time for Session 2 equal the time allowed?	65	33	-
Did the test administrator announce "you have 10 minutes left" prior to the end of Session 2?	94	4	-
Were any other "time remaining" announcements made during Session 2?	7	91	-
At the end of Session 2, did the test administrator collect the test booklets one at a time from the students?	47	51	-
When the test administrator read the script for the end of testing Session 2, did he or she announce a break to be followed by the <i>Student Questionnaire</i> ?	32	61	5
How accurately did the test administrator follow the script to end the testing and signal a break?	40 (no changes)	33 (minor changes) 21 (major changes)	4
If there were any changes, would you describe them as...			
...additions?	20	24	54
...some minor changes?	32	20	46
...omissions?	21	24	53
At the end of the break, did the test administrator distribute the Student Questionnaires and give directions as specified in the script?	56	29	13
Did the students ask for additional time to complete the questionnaire?	66	26	6
At the end of the session, prior to dismissing the students, did the test administrator thank the students for participating in the study?	86	8	4

Exhibit 9.4 presents the results of the remaining questions asked about the test session activities. These questions dealt with topics such as student compliance with instructions, and the alignment between scripted instructions and their implementation.

The results show that in almost all of the sessions, the students complied well or very well with the instructions to stop testing. Additionally, in nearly 70% of the sessions students were given extra time to complete the Student Questionnaire.

Exhibit 9.4 Test Session Activities

Question	Very well	Well	Fairly well	Not well	N/A
When the test administrator ended Session 1, how well did the students comply with the instructions to "stop work"?	93	2	2	-	1
When the test administrator ended Session 2, how well did the students comply with the instructions to "stop work"?	93	4	-	-	1
	Exactly	Longer	Shorter	N/A	
How does the total time allocated for the administration of the <i>Student Questionnaire</i> compare with the time specified in the script?	13	70	6	9	
	Very orderly	Somewhat orderly	Not orderly at all	N/A	
How orderly was the dismissal of students?	60	29	5	4	

9.2.3 General Impressions

Section C dealt with the quality control monitors general observations and overall impressions of the test administration. It covered topics such as how well the test administrator monitored the behavior of the students during the testing, and any unusual circumstances that may have come up during the session (e.g., cheating, emergency situations, student refusal to participate, defective instrumentation).

Examination of the results presented in Exhibit 9.5 shows that in almost all sessions, the testing took place without any problems. In roughly 7% of sessions, QCMs reported seeing evidence of students attempting to cheat on the test. When asked to expand on this, QCMs generally indicated that students were either whispering to each other after they were done or were looking around at their neighbors to see whether their test booklets were indeed different. Because the TIMSS test design involves eight different booklets distributed among the students, students usually did not have the same booklet as their neighbors, so any students who may have tried to copy a neighbor's answers would have been deterred by the test design.

Finally, a large proportion of testing sessions had one or more students leave the room for an “emergency” during testing. Typically these emergencies were bathroom breaks. In many of these instances, booklets were not collected from the student; instead, the students left the booklets on their desk.

Exhibit 9.5 Summary Observations of the QCMs

Question	Yes	No	N/A
During the testing situation did the test administrator walk around the room to be sure students were working on the correct section of the test and/or behaving properly?	93	4	1
In your opinion, did the test administrator address students' questions appropriately?	94	3	1
Did you see any evidence of students attempting to cheat on the tests (e.g., by copying from a neighbor)?	6	90	2
Were any defective booklets detected and replaced before the testing began?	-	95	3
Were any defective booklets detected and replaced after the testing began?	-	97	1
If any defective test booklets were replaced, did the test administrator replace them appropriately?	-	4	94
Did any students refuse to take the test either prior to the testing or during the testing?	9	86	3
If a student refused, did the test administrator accurately follow the instructions for excusing the student (collect the test booklet and record the incident on the Student Tracking Form)?	2	2	94
Did any students leave the room for an “emergency” during the testing?	43	52	3
If yes, did the test administrator address the situation appropriately (collect the booklet, and if the student was readmitted, return the test booklet and record time out of the room on the test booklet)?	13	25	60

Finally, Exhibit 9.6 indicates that in almost all of the testing sessions, QCMs found the behavior of students to be orderly and cooperative. Where it was less than perfect, the test administrator was almost always able to control the students and the situation. For the great majority of sessions, QCMs reported that the overall quality of the sessions was either excellent or very good.

Exhibit 9.6 Summary Observations of Student Behavior

Question	Extremely	Moderately	Somewhat	Hardly at all	N/A
To what extent would you describe the students as orderly and cooperative?	65	26	6	-	1
	No, no late students	No, not admitted	Yes, before testing began	Yes, after testing began	N/A
Were any late students admitted to the testing room?	74	2	8	13	1
	Excellent	Very good	Good	Fair	Poor
In general, how would you describe the overall quality of the testing session?	52	24	12	5	4
	Definitely Yes	Some effort was made	Hardly any effort was made	N/A	
If the students were not cooperative and orderly, did the test administrator make an effort to control the students and the situation?	23	9	5	61	

9.2.4 Interview with the School Coordinator

In Section D of the Classroom Observation Record, the QCMs recorded details of the interview with the school coordinator. Issues addressed included shipping of assessment materials, satisfaction with arrangements for the test administration, the responsiveness of Westat to queries, necessity for make-up sessions, and, as a check on within-school sampling activities, the organization of classes in the school.

The results presented in Exhibit 9.7 show that TIMSS 1999 was an administrative success in the eyes of the school coordinators. In 80% or more of the cases, school coordinators reported that Westat was responsive to their questions or concerns, and that relations were cordial and cooperative.

About half of the school coordinators reported that they were able to collect the completed Teacher Questionnaires prior to student testing. Of the rest, the vast majority reported that they were missing only one or two questionnaires and were expecting them to be turned in shortly.

It was estimated that the Teacher Questionnaires would take about 60 minutes to complete. Of the school coordinators who had administered the Teacher Questionnaire at the time of the interview, about 61% indicated that the estimate was about right, while about 11% reported that the questionnaires took longer and about 28% that they took less time to complete.

Finally, it is worth noting that in about 53% of the cases, school coordinators indicated that students were given special instructions, motivational talks, or incentives prior to testing. Students were given special instructions more often than motivational talks or special incentives, and most frequently these were contained in a letter sent home to the students' parents.

Exhibit 9.7 Interview with the School Coordinator

Question	Yes	No	N/A
Was Westat responsive to your questions or concerns?	86	4	8
Were you able to collect completed <i>Teacher Questionnaires</i> prior to the test administration?	52	39	7
It was expected that the <i>Teacher Questionnaire</i> would require about 60 minutes to complete. In your opinion, was that estimate correct?	43	8 (longer) 20 (less time)	27
Were you satisfied with the accommodations (testing room) you were able to arrange for the testing?	86	8	4
Did the students receive any special instructions, motivational talk, or incentives to prepare them for the assessment?	55	40	3
Were students given any opportunity to practice on questions like those in the tests before the testing session?	4	92	2
Is this a complete list of the mathematics classes in this grade in this school?	85	4	9
To the best of your knowledge, are there any students in this grade level who are <i>not</i> in any of these mathematics classes?	18	76	4
To the best of your knowledge, are there any students in this grade level in more than one of these mathematics classes?	6	89	3
If there were another TIMSS Benchmarking assessment, would you be willing to serve as a school coordinator?	92	4	2

Perhaps the biggest tribute to the successful planning and implementation of TIMSS 1999 was the fact that nearly 94% of respondents said that if there were to be another TIMSS Benchmarking assessment, they would be willing to serve as the school coordinator. Furthermore, the results shown in Exhibit 9.8 suggest that practically all of the school coordinators thought the testing sessions went well, and that most thought that staff members in their school felt positive about the TIMSS 1999 testing.

Exhibit 9.8 Interview with the School Coordinator (continued)

Question	Very well	Satisfactory	Unsatisfactory	N/A
Overall, how would you say the session went?	61	30	1	6
	Positive	Neutral	Negative	N/A
Overall, how would you rate the attitude of the other school staff members towards the TIMSS testing?	57	34	4	3

9.3 Summary

In summary, the observations by the quality control monitors indicate that the data collected in the TIMSS 1999 Benchmarking study met strict standards for quality, and that as a result there can be a high level of confidence in the findings.

References

TIMSS (1998). *Manual for quality control monitors* (TIMSS 1999 Doc. Ref. No. 98-0023). Prepared by the International Study Center at Boston College. Chestnut Hill, MA: Boston College.

Westat (1999). *TIMSS 1999 Benchmarking test administrator manual*. Prepared by Westat, Inc. Rockville, MD: Westat, Inc.

