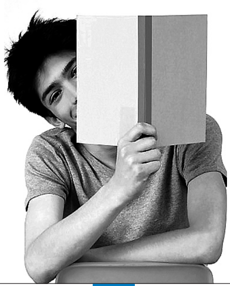


6

Field operations

PISA 2006 TECHNICAL REPORT – ISBN 978-92-64-04808-9 – © OECD 2009

OVERVIEW OF ROLES AND RESPONSIBILITIES

PISA was implemented in each country by a National Project Manager (NPM) who implemented the procedures prepared by the consortium. Each NPM typically had several assistants, working from a base location that is referred to throughout this report as a national centre (NC). For the school level operations the NPM coordinated activities with school level staff, referred to as school co-ordinators (SCs). Trained test administrators (TAs) administered the PISA assessment in schools.

National project managers

NPMs were responsible for implementing the project within their own country. They:

- Attended NPM meetings and received training in all aspects of PISA operational procedures;
- Negotiated nationally specific aspects of the implementation of PISA with the consortium, such as national and international options, oversampling for regional comparisons, additional analyses and reporting, *e.g.* by language group;
- Established procedures for the security of materials during all phases of the implementation;
- Prepared a series of sampling forms documenting sampling related aspects of the national educational structure;
- Prepared the school sampling frame and submitted this to the consortium for the selection of the school sample;
- Organised for the preparation of national versions of the test instruments, questionnaires, manuals and coding guides;
- Identified school co-ordinators from each of the sampled schools and worked with them on school preparation activities;
- Selected the student sample from a list of eligible students provided by the school co-ordinators;
- Recruited and trained test administrators to administer the tests within schools;
- Nominated suitable persons to work on behalf of the consortium as external quality monitors to observe the test administration in a selection of schools;
- Recruited and trained coders to code the open-ended items;
- Arranged for the data entry of the test and questionnaire responses, and submitted the national database of responses to the consortium;
- Submitted a written review of PISA implementation activities following the assessment.

A *National Project Manager's Manual* provided detailed information about the duties and responsibilities of the NPM. Supplementary manuals, with detailed information about particular aspects of the project, were also provided. These included:

- A *School Sampling Preparation Manual,* which provided instructions to the NPM for documenting school sampling related issues such as the definition of the target population, school level exclusions, the proportion of small schools in the sample and so on. Instructions for the preparation of the sampling frame, *i.e.* the list of all schools containing PISA eligible students, were detailed in this manual;
- A *Data Management Manual*, which described all aspects of the use of *KeyQuest*, the data entry software prepared by the consortium for the data entry of responses from the tracking instruments, test booklets and questionnaires.

School co-ordinators

School co-ordinators (SCs) co-ordinated school-related activities with the national centre and the test administrators.

The SC:

- Established the testing date and time in consultation with the NPM;
- Prepared the student listing form with the names of all eligible students in the school and sent it to the NPM so that the NPM could select the student sample;
- Received the list of sampled students on the student tracking form from the NPM and updated it if necessary, including identifying students with disabilities or limited test language proficiency who could not take the test according to criteria established by the consortium;
- Received, distributed and collected the school questionnaire;
- Received and distributed the parent questionnaire in the countries that implemented this international option;
- Informed school staff, students and parents of the nature of the test and the test date, and secured parental permission if required by the school or education system;
- Informed the NPM and test administrator of any test date or time changes;
- Assisted the test administrator with room arrangements for the test day.

On the test day, the SC was expected to ensure that the sampled students attended the test session(s). If necessary, the SC also made arrangements for a follow-up session and ensured that absent students attended the follow-up session.

A *School Co-ordinator's Manual* was prepared by the consortium that described in detail the activities and responsibilities of the SC.

Test administrators

The test administrators were primarily responsible for administering the PISA test fairly, impartially and uniformly, in accordance with international standards and PISA procedures. To maintain fairness, a TA could not be the reading, mathematics or science teacher of the students being assessed and it was preferred that they not be a staff member at any participating school. Prior to the test date, TAs were trained by national centres. Training included a thorough review of the *Test Administrator's Manual,* prepared by the consortium, and the script to be followed during the administration of the test and questionnaire. Additional responsibilities included:

- Ensuring receipt of the testing materials from the NPM and maintaining their security;
- Co-operating with the SC;
- Contacting the SC one to two weeks prior to the test to confirm plans;
- Completing final arrangements on the test day;
- Conducting a follow-up session, if needed, in consultation with the SC;
- Completing the student tracking form and the assessment session report form (a form designed to summarise session times, student attendance, any disturbance to the session, etc.);
- Ensuring that the number of tests and questionnaires collected from students tallied with the number sent to the school;
- Obtaining the school questionnaire from the SC; and
- Sending the school questionnaire, the student questionnaires and all test materials (both completed and not completed) to the NPM after the testing was carried out.

School Associates

In some countries, one person undertook the roles of both school co-ordinator and test administrator. In these cases, the person was referred to as the school associate (SA). A *School Associate's Manual* was prepared by the consortium, combining the source material provided in the individual SC and TA manuals to describe in detail the activities and responsibilities of the SA.

THE SELECTION OF THE SCHOOL SAMPLE

NPMs used the detailed instructions in the *School Sampling Preparation Manual* to document their school sampling plan and to prepare their school sampling frame.

The national target population was defined, school and student level exclusions were identified, and aspects such as the extent of small schools and the homogeneity of students within schools were considered in the preparation of the school sampling plan.

For all but a small number of countries, the sampling frame was submitted to the consortium who selected the school sample. Having the consortium select the school sample minimised the potential for errors in the sampling process, and ensured uniformity in the outputs for more efficient data processing later. It also relieved the burden of this task from national centres. NPMs worked very closely with the consortium throughout the process of preparing the sampling documentation, ensuring that all nationally specific considerations related to sampling were thoroughly documented and incorporated into the school sampling plan.

While all countries were required to thoroughly document their school sampling plan, a small number of countries were permitted to select the school sample themselves. In these cases, the national centre was required to explain in detail the sampling methods used, to ensure that they were consistent with those used by the consortium. In these cases, the standard procedure the consortium used to check that the national school sampling had been implemented correctly was to draw a parallel sample using its international procedures and compare the two samples. Further details about sampling for the main study are provided in Chapter 4.

PREPARATION OF TEST BOOKLETS, QUESTIONNAIRES AND MANUALS

As described in Chapter 2, thirteen different test booklets had to be assembled with clusters of test items arranged according to the test booklet design specified by the consortium. Test items were presented in units (stimulus material and items relating to the stimulus) and each cluster contained several units. Test units and questionnaire items were initially sent to NPMs several months before the testing dates, allowing adequate time for items to be translated. Units allocated to clusters and clusters allocated to booklets were provided a few weeks later, together with detailed instructions to NPMs about how to assemble their translated or adapted clusters into booklets.

For reference, source versions of all booklets were provided to NPMs in both English and French and were also available through a secure website. NPMs were encouraged to use the cover design provided by the OECD. In formatting translated or adapted test booklets, they had to follow as far as possible the layout in the source versions, including allocation of items to pages. A slightly smaller or larger font than in the source version was permitted if it was necessary to ensure the same page set-up as that of the source version.

NPMs were required to submit their cognitive material in units, along with a form documenting any proposed national adaptations for verification by the consortium. NPMs incorporated feedback from the verifier into their material and assembled the test booklets. These were submitted once more to the consortium, who performed a final optical check (FOC) of the materials. This was a verification of the layout, instructions to

the student, the rendering of graphic material, etc. Once feedback from the final optical check had been received and incorporated into the test booklets, the NPM was ready to send the materials to print.

The student questionnaire contained one or two modules, according to whether the Information and Computer Technology Familiarity international option questionnaire component was being added to the core component. Forty countries chose to administer this component. The core component had to be presented first in the questionnaire booklet.

Sixteen countries also administered an optional parent questionnaire.

As with the test material, source versions of the questionnaire instruments in both French and English were provided to NPMs to be used to assist in the translation of this material.

NPMs were permitted to add questions of national interest as national options to the questionnaires. Proposals and text for these were submitted to the consortium for approval as part of the process of reviewing adaptations to the questionnaires. It was recommended that the additional material should be placed at the end of the international modules. The student questionnaire was modified more often than the school questionnaire.

NPMs were required to submit a form documenting all proposed national adaptations to questionnaire items to the consortium for approval. Following approval of adaptations, the material was verified by the consortium. NPMs implemented feedback from verification in the assembly of their questionnaires, which were submitted once more in order to conduct a final optical check of the layout etc. Following feedback from the final optical check, NPMs made final changes to their questionnaires prior to printing.

The school co-ordinator (SC) and test administrator (TA) manuals (or SA manual for those countries that combined the roles of the SC and TA) were also required to be translated into the national languages. French and English source versions of each manual were provided by the consortium. NPMs were required to submit a form documenting all proposed national adaptations to the manuals to the consortium for approval. Following approval of the adaptations, the manuals were prepared and submitted to the consortium. A verification of key elements of the manuals – those related to the coding of the tracking instruments and the administration of the test – was conducted. NPMs implemented feedback from the verifier into their manuals prior to printing. A final optical check was not required for the manuals.

In countries with multiple languages, the test instruments and manuals needed to be translated into each test language. For a small number of countries, where test administrators were bilingual in the test language and the national language, it was not required for the whole of the manuals to be translated into both languages. However in these cases it was a requirement that the test script, included within the TA manual was translated into the language of the test.

THE SELECTION OF THE STUDENT SAMPLE

Following the selection of the school sample by the consortium, the list of sampled schools was returned to national centres. NPMs then contacted these schools and requested a list of all PISA-eligible students from each school. This was provided on the *student listing form*, and was used by NPMs to select the student sample.

NPMs were required in most cases to select the student sample using *KeyQuest*, the PISA student sampling and data entry software prepared by the consortium. *KeyQuest* generated the list of sampled students for each school, known as the *student tracking form* that served as the central administration document for the study and linked students, test booklets and student questionnaires.

Only in exceptional circumstances were NPMs permitted to select their student sample without using *KeyQuest*. Alternative sampling procedures required the approval of the consortium prior to implementation.

PACKAGING AND SHIPPING MATERIALS

Regardless of how materials were packaged and shipped, the following needed to be sent either to the TA or to the school:

- Test booklets and student questionnaires for the number of students sampled;
- Student tracking form;
- Two copies of the Assessment Session Report Form;
- Packing form;
- Return shipment form;
- Additional materials, *e.g.* rulers and calculators, as per local circumstances;
- Additional school and student questionnaires and a bundle of extra test booklets.

Of the thirteen separate test booklets, one was pre-allocated to each student by the *KeyQuest* software from a random starting point in each school. *KeyQuest* was then used to generate the school's student tracking form, which contained the number of the allocated booklet alongside each sampled student's name.

It was recommended that labels be printed, each with a student identification number and test booklet number allocated to that identification, as well as the student's name if this was an acceptable procedure within the country. Two or three copies of each student's label could be printed, and used to identify the test booklet, the questionnaire, and a packing envelope if used.

NPMs were allowed some flexibility in how the materials were packaged and distributed, depending on national circumstances. It was specified however that the test booklets for a school be packaged so that they remained secure, possibly by wrapping them in clear plastic and then heat-sealing the package, or by sealing each booklet in a labelled envelope. Three scenarios, summarised here, were described as illustrative of acceptable approaches to packaging and shipping the assessment materials:

- Country A: All assessment materials shipped directly to the schools; school staff (not teachers of the students in the assessment) to conduct the testing sessions; materials assigned to students before packaging; materials labelled and then sealed in envelopes also labelled with the students' names and identification numbers.
- Country B: Materials shipped directly to the schools; external test administrators employed by the National Centre to administer the tests; the order of the booklets in each bundle matches the order on the student tracking form; after the assessment has been completed, booklets are inserted into envelopes labelled with the students' names and identification numbers and sealed.
- Country C: Materials shipped to test administrators employed by the National Centre; bundles of 35 booklets sealed in plastic, so that the number of booklets can be checked without opening the packages; TAs open the bundle immediately prior to the session and label the booklets with the students' names and ID numbers from the student tracking form.

RECEIPT OF MATERIALS AT THE NATIONAL CENTRE AFTER TESTING

It was recommended that the national centre establish a database of schools before testing began to record the shipment of materials to and from schools, tallies of materials sent and returned, and to monitor the progress of the materials throughout the various steps in processing booklets after the testing.

It was recommended that upon receipt of materials back from schools, the counts of completed and unused booklets also be checked against the participation status information recorded on the student tracking form by the TA.

CODING OF THE TESTS AND QUESTIONNAIRES

This section describes PISA's coding procedures, including multiple coding, and makes brief reference to pre-coding of responses to a few items in the student questionnaire. Overall, 45% of the cognitive items across the science, reading and mathematics domains required manual coding by trained coders.

This was a complex operation, as booklets had to be randomly assigned to coders and, for the minimum recommended sample size per country of 4500 students, more than 116 000 responses had to be evaluated. An average of 26 items from each of the thirteen booklets required evaluation.

It is crucial for comparability of results in a study such as PISA that students' responses are scored uniformly from coder to coder and from country to country. Comprehensive criteria for coding, including many examples of acceptable and unacceptable responses, were prepared by the consortium and provided to NPMs in coding guides for each of science, reading and mathematics.

Preparing for coding

In setting up the coding of students' responses to open-ended items, NPMs had to carry out or oversee several steps:

- Adapt or translate the coding guides as needed and submit these to the consortium for verification;
- Recruit and train coders;
- Locate suitable local examples of responses to use in training and practice;
- Organise booklets as they were returned from schools;
- Select booklets for multiple coding;
- Single code booklets according to the international design;
- Multiple code a selected sub-sample of booklets once the single coding was completed;
- Submit a sub-sample of booklets for the International Coding Review (see Chapter 13).

Detailed instructions for each step were provided in the *Main Study NPM's Manual.* Key aspects of the process are included here.

International training

Representatives from each national centre were required to attend two international coder training sessions – one immediately prior to the field trial and one immediately prior to the main study. At the training sessions consortium staff familiarised national centre staff with the coding guides and their interpretation.

Staffing

NPMs were responsible for recruiting appropriately qualified people to carry out the single and multiple coding of the test booklets. In some countries, pools of experienced coders from other projects could be called on. It was not necessary for coders to have high-level academic qualifications, but they needed to have a good understanding of either mid-secondary level mathematics and science or the language of the test, and to be familiar with ways in which secondary-level students express themselves. Teachers on leave,

recently retired teachers and senior teacher trainees were all considered to be potentially suitable coders. An important factor in recruiting coders was that they could commit their time to the project for the duration of the coding, which was expected to take up to two months.

The consortium provided a coder recruitment kit to assist NPMs in screening applicants. These materials were similar in nature to the coding guides, but were much briefer. They were designed so that applicants who were considered to be potentially suitable could be given a brief training session, after which they coded some student responses. Guidelines for assessing the results of this exercise were supplied. The materials also provided applicants with the opportunity to assess their own suitability for the task. The number of coders required was governed by the design for multiple coding (described in a later section). For the main study, it was recommended to have 16 coders coding across the domains of science and mathematics, and an additional four coders to code reading. These numbers of coders were considered to be adequate for countries testing between 4 500 (the minimum number required) and 6 000 students to meet the timeline of submitting their data within three months of testing.

For larger numbers of students or in cases where coders would code across different combinations of domains, NPMs could prepare their own design and submit it to the consortium for approval. A minimum of four coders were required in each domain to satisfy the requirements of the multiple coding design. Given that several weeks were required to complete the coding, it was recommended that at least two back-up coders of science and mathematics and one back-up reading coder be trained and included in at least some of the coding sessions.

The coding process was complex enough to require a full-time overall supervisor of activities who was familiar with the logistical aspects of the coding design, the procedures for checking coder reliability, the coding schedules and the content of the tests and coding guides.

NPMs were also required to designate persons with subject-matter expertise, familiarity with the PISA tests and, if possible, experience in coding student responses to open-ended items to act as 'table leaders' during the coding. Table leaders were expected to participate in the actual coding and spend extra time monitoring consistency. Good table leaders were essential to the quality of the coding, as their main role was to monitor coders' consistency in applying the coding criteria. They also assisted with the flow of booklets, and fielded and resolved queries about the coding guide and about particular student responses in relation to the guide, consulting the supervisor as necessary when queries could not be resolved. The supervisor was then responsible for checking such queries with the consortium.

People were also needed to unpack, check and assemble booklets into labelled bundles so that coders could respect the specified design for randomly allocating sets of booklets to coders.

Consortium coding query service

A coding query service was provided by the consortium in case questions arose about particular items that could not be resolved at the national centre. Responses to coding queries were placed on the website, accessible to the NPMs from all participating countries.

Confidentiality forms

Before seeing or receiving any copies of PISA test materials, prospective coders were required to sign a confidentiality form, obligating them not to disclose the content of the PISA tests beyond the groups of coders and trainers with whom they would be working.

National training

Anyone who coded the PISA main survey test booklets had to participate in specific training sessions, regardless of whether they had had related experience or had been involved in the PISA field trial coding. To assist NPMs in carrying out the training, the consortium prepared training materials in addition to the detailed coding guides. Training within a country could be carried out by the NPM or by one or more knowledgeable persons appointed by the NPM. Subject matter knowledge was important for the trainer as was an understanding of the procedures, which usually meant that more than one person was involved in leading the training.

The recommended allocation of booklets to coders assumed coding by cluster. This involved completing the coding of each item separately within a cluster within all of the booklets allocated to the coder before moving to the next item, and completing one cluster before moving to the next.

Coders were trained by cluster for the seven science clusters, the four mathematics clusters and the two clusters of reading. During a training session, the trainer reviewed the coding guide for a cluster of units with the coders, and then had the coders assign codes to some sample items for which the appropriate codes had been supplied by the consortium. The trainer reviewed the results with the group, allowing time for discussion, querying and clarification of reasons for the pre-assigned codes. Trainees then proceeded to code independently some local examples that had been carefully selected by the supervisor of coding in conjunction with national centre staff. It was recommended that prospective coders be informed at the beginning of training that they would be expected to apply the coding guides with a high level of consistency, and that reliability checks would be made frequently by table leaders and the overall supervisor as part of the coding process.

Ideally, table leaders were trained before the larger groups of coders since they needed to be thoroughly familiar with both the test items and the coding guides. The coding supervisor explained these to the point where the table leaders could code and reach a consensus on the selected local examples to be used later with the larger group of trainees. They also participated in the training sessions with the rest of the coders, partly to strengthen their own knowledge of the coding guides and partly to assist the supervisor in discussions with the trainees of their pre-agreed codes to the sample items. Table leaders received additional training in the procedures for monitoring the consistency with which coders applied the criteria.

Length of coding sessions

Coding responses to open-ended items is mentally demanding, requiring a level of concentration that cannot be maintained for long periods of time. It was therefore recommended that coders work for no more than six hours per day on actual coding, and take two or three breaks for coffee and lunch. Table leaders needed to work longer on most days so that they had adequate time for their monitoring activities.

Logistics prior to coding

Sorting booklets

When booklets arrived back at the national centre, they were first tallied and checked against the session participation codes on the student tracking form. Unused and used booklets were separated; used booklets were sorted by student identification number if they had not been sent back in that order and then were separated by booklet number; and school bundles were kept in school identification order, filling in sequence gaps as packages arrived. student tracking forms were copied, and the copies filed in school identification order. If the school identification number order did not correspond with the alphabetical order of school names, it was recommended that an index of school name against school identification be prepared and kept with the binders.

Because of the time frame within which countries had to have all their coding done and data submitted to the consortium, it was usually impossible to wait for all materials to reach the national centre before beginning to code. In order to manage the design for allocating booklets to coders, however, it was recommended to start coding only when at least half of the booklets had been returned.

Selection of booklets for multiple coding

Each country was required to set aside 100 each of booklets 1, 3, 5, 6, 8 and 10 for multiple coding. The first two clusters from each of these booklets were multiple coded, except booklet 5 where the first three clusters were multiple coded. This arrangement ensured that all clusters were included in the multiple coding.

The main principle in setting aside the booklets for multiple coding was that the selection needed to ensure a wide spread of schools and students across the whole sample and to be random as far as possible. The simplest method for carrying out the selection was to use a ratio approach based on the expected total number of completed booklets.

In most countries, approximately 400 of each booklet was expected to be completed, so the selection of booklets to be set aside for multiple coding required that approximately one in four booklets was selected. Depending on the actual numbers of completed booklets received, the selection ratios needed to be adjusted so that the correct numbers of each booklet were selected from the full range of participating schools.

In a country where booklets were provided in more than one language, if the language represented 20% or more of the target population, the 600 booklets to be set aside for multiple coding were allocated in proportion to the language group. Multiple coding was not required for languages representing less than 20% of the target population.

Booklets for single coding

Single coding was required for the booklets remaining after those for multiple coding had been set aside, as well as for the clusters in the latter part of the book from those set aside for multiple coding. Some items requiring coding did not need to be included in the multiple coding. These were closed constructed response items that required a coder to assign a right or wrong code, but did not require any coder judgement. The last coder in the multiple-coding process coded these items in the booklets set aside for multiple coding, as well as the items requiring single coding from the third and fourth clusters. Other items such as multiple-choice response items required no coding and were directly data-entered.

How codes were shown

A string of small code numbers corresponding to the possible codes for the item as delineated in the relevant coding guide appeared in the upper right-hand side of each item in the test booklets. For booklets being processed by a single coder, the code assigned was indicated directly in the booklet by circling the appropriate code number alongside the item. Tailored coding record sheets were prepared for each booklet for the multiple coding and used by all but the last coder so that each coder undertaking multiple coding did not know which codes other coders had assigned.

For the reading clusters, item codes were often just 0, 1 and 9, indicating incorrect, correct and missing, respectively. Provision was made for some of the open-ended items to be coded as partially correct, usually with "2" as fully correct and "1" as partially correct, but occasionally with three degrees of correctness indicated by codes of "1", "2" and "3".

For the mathematics and science clusters, a two-digit coding scheme was adopted for the items requiring constructed responses. The first digit represented the degree of correctness code, as in reading; the second indicated the content of the response or the type of solution method used by the student. Two-digit codes were originally proposed by Norway for the TIMSS and were adopted in PISA because of their potential for use in studies of student learning and thinking.

Coder identification numbers

Coder identification numbers were assigned according to a standard three-digit format specified by the consortium. The first digit showed the combination of domains that the coder would be working across, and the second and third digits had to uniquely identify the coders within their set. For example, sixteen coders coding across the domains of science and mathematics were given identification numbers 501 to 516. Four coders who coded just reading were given identification numbers 201 to 204. Coder identification numbers were used for two purposes: implementing the design for allocating booklets to coders and monitoring coder consistency in the multiple-coding exercises.

Single coding design

Single coding of science and mathematics

In order to code by cluster, each coder needed to handle four of the thirteen booklet types at a time. For example, science cluster 1 occurred in booklets 1, 9, 10 and 12. Each of these appearances had to be coded before another cluster was started. Moreover, since coding was done item by item, the item was coded across these different booklet types before the next item was coded.

A design to ensure the random allocation of booklets to coders was prepared based on the recommended number of 16 coders and the minimum sample size of 4 500 students from 150 schools. With 150 schools and 16 coders, each coder had to code a cluster within a booklet from eight or nine schools (150 / 16 ≈ 9). Figure 6.1 shows how booklets needed to be assigned to coders for the single coding. Further explanation of the information in this table is presented below.

According to this design, cluster S1 in school subset 1 (schools 1 to 9) was to be coded by coder 501. cluster S1 in subset 2 (schools 10 to 18) was to be coded by coder 502, and so on. For cluster S2, coder 501 was to code all from subset 2 (schools 10 to 18) and coder 502 was to code all from subset 3 (schools 19 to 27). Subset 1 of cluster M2 (schools 1 to 9) was to be coded by coder 509.

Figure 6.1

Design for the single coding of science and mathematics

		Batches															
Cluster	Booklets	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
S1	1, 9, 10, 12	501	502	503	504	505	506	507	508	509	510	511	512	513	514	515	516
S2	1, 2, 8, 11	516	501	502	503	504	505	506	507	508	509	510	511	512	513	514	515
S3	2, 3, 5, 9	515	516	501	502	503	504	505	506	507	508	509	510	511	512	513	514
S4	1, 3, 4, 6	514	515	516	501	502	503	504	505	506	507	508	509	510	511	512	513
S5	4, 5, 11, 12	513	514	515	516	501	502	503	504	505	506	507	508	509	510	511	512
S6	5, 6, 8, 10	512	513	514	515	516	501	502	503	504	505	506	507	508	509	510	511
S7	1, 5, 7, 13	511	512	513	514	515	516	501	502	503	504	505	506	507	508	509	510
M1	3, 8, 12, 13	510	511	512	513	514	515	516	501	502	503	504	505	506	507	508	509
M2	4, 7, 8, 9	509	510	511	512	513	514	515	516	501	502	503	504	505	506	507	508
M3	2, 4, 10, 13	508	509	510	511	512	513	514	515	516	501	502	503	504	505	506	507
M4	3, 7, 10, 11	507	508	509	510	511	512	513	514	515	516	501	502	503	504	505	506

If booklets from all participating schools were available before the coding began, the following steps would be involved in implementing the design:

Step 1: Set aside booklets for multiple coding and then divide the remaining booklets into school subsets as above (subset 1: schools 1 to 9; subset 2: schools 10 to 18, etc. to achieve 16 subsets of schools).

Step 2: Assuming that coding begins with cluster S1: coder 501 takes booklets 1, 9, 10 and 12 for school subset 1; coder 502 takes booklets 1, 9, 10 and 12 for school subset 2; etc.; until coder 516 takes booklets 1, 9, 10 and 12 for school subset 16.

Step 3: Coders code all of the first cluster S1 item requiring coding in the booklets that they have.

Step 4: The second cluster S1 item is coded in all four booklet types, followed by the third cluster S1 item, etc., until all cluster S1 items are coded.

Step 5: For cluster S2, as per the row of the table in Figure 6.1 corresponding to S2 in the left-most column, each coder is allocated a subset of schools different from their subset for cluster S1. Coding proceeds item by item within the cluster.

Step 6: For the remaining clusters, the rows corresponding to S3, S4, etc. in the table are followed in succession.

Single coding of reading

A similar design was prepared for the single coding of reading (Figure 6.2). As the recommended number of coders for reading (4) was one quarter that recommended for coding science and mathematics, each coder was allocated ‘four subsets worth’ of schools. Also, as there were just two different clusters of reading, each of which appeared in four booklet types, each coder coded just one of the four appearances of a cluster. This ensured that a wider range of coders was used for each school subset. For the coding of cluster R1, for example, coder 201 coded this cluster in booklet 1 from school subsets 1-4 (*i.e.* schools 1-36), coder 202 coded this cluster from booklet 1 for school subsets 5-8, and so on. For the next appearance of cluster R1 (in booklet 6), coder 204 coded these from school subsets 1-4, coder 201 from school subsets 5-8, and so on.

As a result of this procedure, the booklets from each subset of schools were processed by fifteen different coders, one for each distinct cluster of science and mathematics, and four for each cluster of reading. Each student’s booklet was coded by four different coders, one for each of the four clusters in the student’s booklet. Spreading booklets among coders in this way minimised the effects of any systematic leniency or harshness in coding.

Figure 6.2
Design for the single coding of reading

		Batches			
Cluster	**Booklet**	**1-4**	**5-8**	**9-12**	**13-16**
R1	**2**	201	202	203	204
R1	**6**	204	201	202	203
R1	**7**	203	204	201	202
R1	**12**	202	203	204	201
R2	**13**	201	202	203	204
R2	**11**	204	201	202	203
R2	**9**	203	204	201	202
R2	**6**	202	203	204	201

In practice, most countries would not have had completed test booklets back from all their sampled schools before coding needed to begin. NPMs were encouraged to organise the coding in two waves, so that it could begin after materials were received back from one-half of their schools. Schools would not have been able to be assigned to school sets for coding exactly in their school identification order, but rather by identification order combined with when their materials were received and processed at the national centre.

Booklet UH

Countries using the shorter, special purpose booklet UH were advised to process this separately from the remaining booklets. Small numbers of students used this booklet, only a few items required coding, and they were not arranged in clusters. NPMs were cautioned that booklets needed to be allocated to several coders to ensure uniform application of the coding criteria for booklet UH, as for the main coding.

Multiple coding

For PISA 2006, four coders independently coded all short response and open-constructed response items from a selection of clusters from a sample of booklets. 100 of each of Booklets 1, 3, 5, 6, 8 and 10 (a total of 600 booklets) were selected for this multiple coding activity. Multiple coding was done at or towards the end of the coding period, after coders had familiarised themselves with and were experienced in using the coding guides. As noted earlier, the first three coders of the selected booklets circled codes on separate record sheets, tailored to booklet type and domain (science, reading or mathematics), using one page per student. The coding supervisor checked that coders correctly entered student identification numbers and their own identification number on the sheets, which was crucial to data quality. The UH booklet was not included in the multiple coding.

While coders would have been thoroughly familiar with the coding guides by the time of multiple coding, they may have most recently coded a different booklet from those allocated to them for multiple coding. For this reason, they needed to have time to re-read the relevant coding guide before beginning the coding. It was recommended that time be allocated for coders to refresh their familiarity with the guides and to look again at the additional practice material before proceeding with the multiple coding. As in the single coding, coding was to be done item by item. For manageability, items from the four clusters within a booklet type were coded before moving to another booklet type, rather than coding by cluster across several booklet types. It was considered that coders would be experienced enough in applying the coding criteria by this time that coding by booklet would be unlikely to detract from the quality of the data.

Multiple coding of science and mathematics

The specified multiple coding design for science and mathematics, shown in Table 6.1, assumed 16 coders with identification numbers 501 to 516. The importance of following the design exactly as specified was stressed, as it provided for links between clusters and coders. Table 6.1 shows 16 coders grouped into four groups of four, with Group 1 comprising the first four coders (501-504), Group 2 the next four (505-508), etc. The design involved two steps, with the booklets divided into two sets - booklets 1, 3, 8 and 10 made up one set, and booklet 5 the second set. The coders assigned to the second step consisted of one coder from each of the groups formed at the first step. The four codings were to be carried out by rotating the booklets to the four coders assigned to each group.

In this scenario, with all 16 coders working, booklets 1, 3, 8 and 10 were to be coded at the same time in the first step. The 100 booklet 1's, for example, were to be divided into four bundles of 25 and rotated among coders 501, 502, 503 and 504, so that each coder eventually would have coded clusters S1 and S2 from all of the 100 booklets. At the fourth rotation, after each coder had finished the multiple coding of clusters S1

and S2 from the 25 booklets in their pile, they would then single code any science or maths clusters from the second half of the booklet. The same pattern was to be followed for booklets 3, 8 and 10.

After booklets 1, 3, 8 and 10 had been put through the multiple-coding process, one coder from each of the four coding groups was selected to complete the multiple-coding of booklet 5. That is, coders 501, 506, 511 and 516 were assigned to code booklet 5,

Allocating booklets to coders for multiple coding was quite complex and the coding supervisor had to monitor the flow of booklets throughout the process.

Table 6.1
Design for the multiple coding of science and mathematics

Booklet	Coder IDs	Clusters for multiple coding	Clusters for single coding
1	501. 502. 503. 504	S1. S2	S4. S7
3	505. 506. 507. 508	S3. S4	M4. M1
8	509. 510. 511. 512	M1. M2	S2. S6
10	513. 514. 515. 516	M3. M4	S6. S1
5	501. 506. 511. 516	S5. S6. S7	S3
6	Any coders available from 501 – 516		S4. S6

Multiple coding of reading

The multiple-coding design for reading shown in Table 6.2 assumed four coders, with identification numbers 201 to 204.

If different coders were used for science or mathematics, a different multiple-coding design was necessary. The NPM would negotiate a suitable proposal with the consortium. The minimum allowable number of coders coding a domain was four; in this case each booklet had to be coded by each coder.

Table 6.2
Design for the multiple coding of reading

Booklet	Coder IDs	Clusters for multiple coding	Clusters for single coding
6	201. 202. 203. 204	R1. R2	none

Managing the coding process

Booklet flow

To facilitate the flow of booklets, it was important to have ample table surfaces on which to place and arrange them by type and school subset. The bundles needed to be clearly labelled. For this purpose, it was recommended that each bundle of booklets be identified by a batch header for each booklet type (booklets 1 to 13), with spaces for the number of booklets and school identification numbers in the bundle to be written in. In addition, each header sheet was to be pre-printed with a list of the clusters in the booklet, with columns alongside which the date and time, coder's name and identification number, and table leader's initials could be entered as the bundle was coded and checked.

Separating the coding of science, mathematics and reading

While consideration of the possibility that coders from different domains would require the same booklets at the same time was factored into the design of the single coding scheme, there was still the potential for this clash to occur. To minimise the risk of different coders requiring the same booklets, so that an efficient flow of booklets through the coding process could be maintained, it was recommended that the coding of

PISA 2006 TECHNICAL REPORT – ISBN 978-92-64-04808-9 – © OECD 2009

reading and the coding of science and mathematics be done at least partly at different times (for example, reading coding could start a week or two ahead).

Familiarising coders with the coding design

The relevant design for allocating booklets to coders was explained either during the coder training session or at the beginning of the first coding session (or both). The coding supervisor was responsible for ensuring that coders adhered to the design and used clerical assistants if needed. Coders could better understand the process if each was provided with a card indicating the bundles of booklets to be taken and in which order.

Consulting table leaders

During the initial training, practice and review, it was expected that coding issues would be discussed openly until coders understood the rationale for the coding criteria (or reached consensus where the coding guide was incomplete). Coders were not permitted to consult other coders or table leaders during the additional practice exercises (see next subsection) undertaken following the training to gauge whether all or some coders needed more training and practice

Following the training, coders were advised to work quietly, referring queries to their table leader rather than to their neighbours. If a particular query arose often, the table leader was advised to discuss it with the rest of the group.

For the multiple coding, coders were required to work independently without consulting other coders.

Monitoring single coding

The steps described here represented the minimum level of monitoring activities required. Countries wishing to implement more extensive monitoring procedures during single coding were encouraged to do so.

The supervisor, assisted by table leaders, was advised to collect coders' practice papers after each cluster practice session and to tabulate the codes assigned. These were then to be compared with the pre-agreed codes: each matching code was considered a hit and each discrepant code was considered a miss. To reflect an adequate standard of reliability, the ratio of hits to the total of hits plus misses needed to be 0.85 or more. In science and mathematics, this reliability was to be assessed on the first digit of the two-digit codes. A ratio of less than 0.85, especially if lower than 0.80, was to be taken as indicating that more practice was needed, and possibly more training.

Table leaders played a key role during each coding session and at the end of each day, by spot-checking a sample of booklets or items that had already been coded to identify problems for discussion with individual coders or with the wider group, as appropriate. All booklets that had not been set aside for multiple coding were candidates for this spot-checking. It was recommended that, if there were indications from the practice sessions that one or more particular coders might be consistently experiencing problems in using the coding guide, then more of those coders' booklets should be included in the checking. Table leaders were advised to review the results of the spot-checking with the coders at the beginning of the next day's coding. This was regarded primarily as a mentoring activity, but NPMs were advised to keep in contact with table leaders and the coding supervisor if there were individual coders who did not meet criteria of adequate reliability and would need to be removed from the pool.

Table leaders were to initial and date the header sheet of each batch of booklets for which they had carried out spot-checking. Some items/booklets from each batch and each coder had to be checked.

Cross-national coding

Cross-national comparability in assigning codes was explored through an inter-country coder reliability study (see Chapter 13).

Questionnaire coding

The main coding required for the student questionnaire internationally was the mother's and father's occupation and student's occupational expectation. Four-digit International Standard Classification of Occupations (ISCO88) codes (International Labour Organisation, 1988) were assigned to these three variables. In several countries, this could be done in a number of ways. NPMs could use a national coding scheme with more than 100 occupational title categories, provided that this national classification could be recoded to ISCO. A national classification was preferred because relationships between occupational status and achievement could then be compared within a country using both international and national measures of occupational status.

The PISA website gave a clear summary of ISCO codes and occupational titles for countries to translate if they had neither a national occupational classification scheme nor access to a full translation of ISCO.

In their national options, countries may also have needed to pre-code responses to some items before data from the questionnaire were entered into the software.

DATA ENTRY, DATA CHECKING AND FILE SUBMISSION

Data entry

The consortium provided participating countries with the data entry software *KeyQuest, which* contained the database structures for all of the booklets, questionnaires and tracking forms used in the main survey. Variables could be added or deleted as needed for national options. Approved adaptations to response categories could also be accommodated. Student response data were entered directly from the test booklets and questionnaires. Information regarding the participation of students, recorded by the SC and TA on the student tracking form, was entered directly into *KeyQuest*. Several questions from the session report form, such as the timing of the session, were also entered into *KeyQuest*.

KeyQuest performed validation checks as data were entered. Importing facilities were also available if data had already been entered into text files, but it was strongly recommended that data be entered directly into *KeyQuest* to take advantage of its PISA-specific features. A *KeyQuest* Manual provided generic technical details of the functionality of the *KeyQuest* software. A separate *Data Entry Manual* provided complete instructions specific to the main study regarding data entry, data management and validity checks.

Data Checking

NPMs were responsible for ensuring that many checks of the quality of their country's data were made before the data files were submitted to the consortium. These checks were explained in detail in the Data Entry Manual, and could be simply applied using the *KeyQuest* software. The checking procedures required that the list of sampled schools and the student tracking form for each school were already accurately completed and entered into *KeyQuest*. Any errors had to be corrected before the data were submitted. Copies of the cleaning reports were to be submitted together with the data files. More details on the cleaning steps are provided in Chapter 10.

PISA 2006 TECHNICAL REPORT – ISBN 978-92-64-04808-9 – © OECD 2009

Data submission

Files to be submitted included:

- Data for the test booklets and context questionnaires;
- Data for the international option instrument(s), if used;
- Data for the multiple-coding study;
- Session report data;
- Data cleaning reports;
- The list of sampled schools;
- Student tracking forms.

Hard or electronic copies of the last two items were also required.

After data were submitted

NPMs were required to designate a data manager who would work actively with the consortium's data processing centre at ACER during the international data cleaning process. Responses to requests for information by the processing centre were required within three working days of the request.

THE MAIN STUDY REVIEW

NPMs were required to complete a structured review of their main study operations. The review was an opportunity to provide feedback to the consortium on the various aspects of the implementation of PISA, and to provide suggestions for areas that could be improved. It also provided an opportunity for the NPM to formally document aspects such as the operational structure of the national centre, the security measures that were implemented, and the use of contractors for particular activities and so on.

The main study review was submitted to the consortium four weeks after the submission of the national database.

Reader's Guide

Country codes – the following country codes are used in this report:

OECD countries

AUS	Australia
AUT	Austria
BEL	Belgium
BEF	Belgium (French Community)
BEN	Belgium (Flemish Community)
CAN	Canada
CAE	Canada (English Community)
CAF	Canada (French Community)
CZE	Czech Republic
DNK	Denmark
FIN	Finland
FRA	France
DEU	Germany
GRC	Greece
HUN	Hungary
ISL	Iceland
IRL	Ireland
ITA	Italy
JPN	Japan
KOR	Korea
LUX	Luxembourg
LXF	Luxembourg (French Community)
LXG	Luxembourg (German Community)
MEX	Mexico
NLD	Netherlands
NZL	New Zealand
NOR	Norway
POL	Poland
PRT	Portugal
SVK	Slovak Republic
ESP	Spain
ESB	Spain (Basque Community)
ESC	Spain (Catalonian Community)
ESS	Spain (Castillian Community)
SWE	Sweden
CHE	Switzerland
CHF	Switzerland (French Community)
CHG	Switzerland (German Community)
CHI	Switzerland (Italian Community)
TUR	Turkey
GBR	United Kingdom
IRL	Ireland
SCO	Scotland
USA	United States

Partner countries and economies

ARG	Argentina
AZE	Azerbaijan
BGR	Bulgaria
BRA	Brazil
CHL	Chile
COL	Colombia
EST	Estonia
HKG	Hong Kong-China
HRV	Croatia
IDN	Indonesia
JOR	Jordan
KGZ	Kyrgyztan
LIE	Liechtenstein
LTU	Lithuania
LVA	Latvia
LVL	Latvia (Latvian Community)
LVR	Latvia (Russian Community)
MAC	Macao-China
MNE	Montenegro
QAT	Qatar
ROU	Romania
RUS	Russian Federation
SRB	Serbia
SVN	Slovenia
TAP	Chinese Taipei
THA	Thailand
TUN	Tunisia
URY	Uruguay

References

Adams, R.J., Wilson, M. & **Wang, W.C.** (1997), The multidimensional random coefficients multinomial logit model. *Applied Psychological Measurement*, No. 21, pp. 1-23.

Adams, R.J., Wilson, M. R. & **Wu, M.L.** (1997), Multilevel item response models: An approach to errors in variables regression, *Journal of Educational and Behavioural Statistics,* No. 22 (1), pp. 46-75.

Adams, R.J. & **Wu, M.L.** (2002), *PISA 2000 Technical Report,* OECD, Paris.

Bollen, K.A. & **Long, S.J.** (1993) (eds.), *Testing Structural Equation Models*, Newbury Park: London.

Beaton, A.E. (1987), Implementing the new design: The NAEP 1983-84 technical report (Rep. No. 15-TR-20), Princeton, NJ: Educational Testing Service.

Buchmann, C. (2000), Family structure, parental perceptions and child labor in Kenya: What factors determine who is enrolled in school? *Soc. Forces,* No. 78, pp. 1349-79.

Buchmann, C. (2002), Measuring Family Background in International Studies of Education: Conceptual Issues and Methodological Challenges, in Porter, A.C. and Gamoran, A. (eds.). *Methodological Advances in Cross-National Surveys of Educational Achievement* (pp. 150-97), Washington, DC: National Academy Press.

Creemers, B.P.M. (1994), *The Effective Classroom*, London: Cassell.

Cochran, W.G. (1977), *Sampling techniques,* third edition, New York, NY: John Wiley and Sons.

Ganzeboom, H.B.G., de Graaf, P.M. & **Treiman, D.J.** (1992), A standard international socio-economic index of occupational status, *Social Science Research*, No. 21, pp. 1-56.

Ganzeboom H.B. & **Treiman, D.J.** (1996), Internationally comparable measures of occupational status for the 1988 international standard classification of occupations, *Social Science Research*, No. 25, pp. 201-239.

Grisay, A. (2003), Translation procedures in OECD/PISA 2000 international assessment, *Language Testing,* No. 20 (2), pp. 225-240.

Hambleton, R.K., Swaminathan, H. & **Rogers, H.J.** (1991), *Fundamentals of item response theory*, Newbury Park, London, New Delhi: SAGE Publications.

Hambleton, R.K., Merenda, P.F. & **Spielberger, C.D.** (2005), *Adapting Educational and Psychological Tests for Cross-Cultural Assessment,* IEA Lawrence Erlbaum Associates, Publishers, Mahwah, New Jersey.

Harkness, J.A., Van de Vijver, F.J.R. & **Mohler, P.Ph** (2003), *Cross-Cultural Survey Methods,* Wiley-Interscience, John Wiley & Sons, Inc., Hoboken, New Jersey.

Harvey-Beavis, A. (2002), Student and School Questionnaire Development, in R.J. Adams and M.L. Wu (eds.), *PISA 2000 Technical Report,* (pp. 33-38), OECD, Paris.

International Labour Organisation (ILO) (1990), *International Standard Classification of Occupations: ISCO-88.* Geneva: International Labour Office.

Jöreskog, K.G. & **Sörbom, Dag** (1993), *LISREL 8 User's Reference Guide*, Chicago: SSI.

Judkins, D.R. (1990), Fay's Method of Variance Estimation, *Journal of Official Statistics*, No. 6 (3), pp. 223-239.

Kaplan, D. (2000), *Structural equation modeling: Foundation and extensions*, Thousand Oaks: SAGE Publications.

Keyfitz, N. (1951), Sampling with probabilities proportionate to science: Adjustment for changes in probabilities, *Journal of the American Statistical Association,* No. 46, American Statistical Association, Alexandria, pp. 105-109.

Kish, L. (1992), Weighting for Unequal, *Pi. Journal of Official Statistics*, No. 8 (2), pp. 183-200.

LISREL (1993), K.G. Jöreskog & D. Sörbom, [computer software], Lincolnwood, IL: Scientific Software International, Inc.

Lohr, S.L. (1999), *Sampling: Design and Analysis*, Duxberry: Pacific Grove.

Macaskill, G., Adams, R.J. & **Wu, M.L.** (1998), Scaling methodology and procedures for the mathematics and science literacy, advanced mathematics and physics scale, in M. Martin and D.L. Kelly, Editors, *Third International Mathematics and Science Study, technical report Volume 3: Implementation and analysis*, Boston College, Chestnut Hill, MA.

Masters, G.N. & **Wright, B.D.** (1997), The Partial Credit Model, in W.J. van der Linden, & R.K. Hambleton (eds.), *Handbook of Modern Item Response Theory* (pp. 101-122), New York/Berlin/Heidelberg: Springer.

Mislevy, R.J. (1991), Randomization-based inference about latent variables from complex samples, *Psychometrika,* No. 56, pp. 177-196.

Mislevy, R.J., Beaton, A., Kaplan, B.A. & **Sheehan, K.** (1992), Estimating population characteristics from sparse matrix samples of item responses, *Journal of Educational Measurement,* No. 29 (2), pp. 133-161.

Mislevy, R.J. & **Sheehan, K.M.** (1987), Marginal estimation procedures, in Beaton, A.E., Editor, 1987. *The NAEP 1983-84 technical report*, National Assessment of Educational Progress, Educational Testing Service, Princeton, pp. 293-360.

Mislevy, R.J. & **Sheehan, K.M.** (1989), Information matrices in latent-variable models, *Journal of Educational Statistics*, No. 14, pp. 335-350.

Mislevy, R.J. & **Sheehan, K.M.** (1989), The role of collateral information about examinees in item parameter estimation, *Psychometrika*, No. 54, pp. 661-679.

Monseur, C. & **Berezner, A.** (2007), The Computation of Equating Errors in International Surveys in Education, *Journal of Applied Measurement,* No. 8 (3), 2007, pp. 323-335.

Monseur, C. (2005), An exploratory alternative approach for student non response weight adjustment, *Studies in Educational Evaluation*, No. 31 (2-3), pp. 129-144.

Muthen, B. & **L. Muthen** (1998), [computer software], *Mplus* Los Angeles, CA: Muthen & Muthen.

Muthen, B., du Toit, S.H.C. & **Spisic, D.** (1997), *Robust inference using weighted least squares and quadratic estimating equations in latent variable modeling with categorical and continuous outcomes*, unpublished manuscript.

OECD (1999), *Classifying Educational Programmes. Manual for ISCED-97 Implementation in OECD Countries,* OECD, Paris.

OECD (2003), *Literacy Skills for the World of Tomorrow: Further results from PISA 2000*, OECD, Paris.

OECD (2004), *Learning for Tomorrow's World – First Results from PISA 2003*, OECD, Paris.

OECD (2005), *Technical Report for the OECD Programme for International Student Assessment 2003*, OECD, Paris.

OECD (2006), *Assessing Scientific, Reading and Mathematical Literacy: A framework for PISA 2006,* OECD, Paris.

OECD (2007), *PISA 2006: Science Competencies for Tomorrow's World*, OECD, Paris.

PISA Consortium (2006), *PISA 2006 Main Study Data Management Manual*, *https://mypisa.acer.edu.au/images/mypisadoc/opmanual/pisa2006_data_management_manual.pdf*

Rasch, G. (1960), Probabilistic models for some intelligence and attainment tests, Copenhagen: Nielsen & Lydiche.

Routitski A. & **Berezner, A.** (2006), Issues influencing the validity of cross-national comparisons of student performance. Data Entry Quality and Parameter Estimation. Paper presented at the Annual Meeting of the American Educational Research Association (AERA) in San Francisco, 7-11 April, *https://mypisa.acer.edu.au/images/mypisadoc/aera06routitsky_berezner.pdf*

Rust, K. (1985), Variance Estimation for Complex Estimators in Sample Surveys, *Journal of Official Statistics*, No. 1, pp. 381-397.

Rust, K.F. & **Rao, J.N.K.** (1996), Variance Estimation for Complex Surveys Using Replication Techniques, *Survey Methods in Medical Research*, No. 5, pp. 283-310.

Shao, J. (1996), Resampling Methods in Sample Surveys (with Discussion), *Statistics*, No. 27, pp. 203-254.

Särndal, C.-E., Swensson, B. & **Wretman, J.** (1992), *Model Assisted Survey Sampling*, New York: Springer-Verlag.

SAS® CALIS (1992), W. Hartmann [computer software], Cary, NC: SAS Institute Inc.

Scheerens, J. (1990), School effectiveness and the development of process indicators of school functioning, *School effectiveness and school improvement,* No. 1, pp. 61-80.

Scheerens, J. & **Bosker, R.J.** (1997), *The Foundations of School Effectiveness*, Oxford: Pergamon.

Schulz, W. (2002), Constructing and Validating the Questionnaire composites, in R.J. Adams and M.L. Wu (eds.), *PISA 2000 Technical Report*, OECD, Paris.

Schulz, W. (2004), Mapping Student Scores to Item Responses, in W. Schulz and H. Sibberns (eds.), *IEA Civic Education Study, Technical Report* (pp. 127-132), Amsterdam: IEA.

Schulz, W. (2006a), *Testing Parameter Invariance for Questionnaire Indices using Confirmatory Factor Analysis and Item Response Theory,* Paper presented at the Annual Meetings of the American Educational Research Association (AERA) in San Francisco, 7-11 April.

Schulz, W. (2006b), *Measuring the socio-economic background of students and its effect on achievement in PISA 2000 and PISA 2003*, Paper presented at the Annual Meetings of the American Educational Research Association (AERA) in San Francisco, 7-11 April.

Thorndike, R.L. (1973), *Reading comprehension in fifteen countries,* New York, Wiley: and Stockholm: Almqvist & Wiksell.

Travers, K.J. & **Westbury, I.** (1989), *The IEA Study of Mathematics I: Analysis of Mathematics Curricula*, Oxford: Pergamon Press.

Travers, K.J., Garden R.A. & **Rosier, M.** (1989), Introduction to the Study, in Robitaille, D. A. and Garden, R. A. (eds), *The IEA Study of Mathematics II: Contexts and Outcomes of School Mathematics Curricula,* Oxford: Pergamon Press.

Verhelst, N. (2002), Coder and Marker Reliabilaity Studies, in R.J. Adams & M.L. Wu (eds.), *PISA 2000 Technical Report*. OECD, Paris.

Walberg, H.J. (1984), Improving the productivity of American schools, *Educational Leadership,* No. 41, pp. 19-27.

Walberg, H. (1986), Synthesis of research on teaching, in M. Wittrock (ed.), *Handbook of research on teaching* (pp. 214-229), New York: Macmillan.

Walker, M. (2006), *The choice of Likert or dichotomous items to measure attitudes across culturally distinct countries in international comparative educational research*. Paper presented at the Annual Meetings of the American Educational Research Association (AERA) in San Francisco, 7-11 April.

Walker, M. (2007), Ameliorating Culturally-Based Extreme Response Tendencies To Attitude items, *Journal of Applied Measurement,* No. 8, pp. 267-278.

Warm, T.A. (1989), Weighted Likelihood Estimation of Ability in Item Response Theory, *Psychometrika*, No. 54 (3), pp. 427-450.

Westat (2007), *WesVar® 5.1* Computer software and manual, Rockville, MD: Author (also see *http://www.westat.com/wesvar/*).

Wilson, M. (1994), Comparing Attitude Across Different Cultures: Two Quantitative Approaches to Construct Validity, in M. Wilson (ed.), *Objective measurement II: Theory into practice* (pp. 271-292), Norwood, NJ: Ablex.

Wolter, K.M. (2007), *Introduction to Variance Estimation*. Second edition, Springer: New York.

Wu, M.L., Adams, R.J. & **Wilson, M.R.** (1997), *ConQuest®: Multi-Aspect Test Software* [computer program manual], Camberwell, Vic.: Australian Council for Educational Research.

List of abbreviations – the following abbreviations are used in this report:

ACER	Australian Council for Educational Research
AGFI	Adjusted Goodness-of-Fit Index
BRR	Balanced Repeated Replication
CBAS	Computer Based Assessment of Science
CFA	Confirmatory Factor Analysis
CFI	Comparative Fit Index
CITO	National Institute for Educational Measurement, The Netherlands
CIVED	Civic Education Study
DIF	Differential Item Functioning
ENR	Enrolment of 15-year-olds
ESCS	PISA Index of Economic, Social and Cultural Status
ETS	Educational Testing Service
IAEP	International Assessment of Educational Progress
I	Sampling Interval
ICR	Inter-Country Coder Reliability Study
ICT	Information Communication Technology
IEA	International Association for the Evaluation of Educational Achievement
INES	OECD Indicators of Education Systems
IRT	Item Response Theory
ISCED	International Standard Classification of Education
ISCO	International Standard Classification of Occupations
ISEI	International Socio-Economic Index
MENR	Enrolment for moderately small school
MOS	Measure of size
NCQM	National Centre Quality Monitor
NDP	National Desired Population
NEP	National Enrolled Population
NFI	Normed Fit Index
NIER	National Institute for Educational Research, Japan
NNFI	Non-Normed Fit Index
NPM	National Project Manager
OECD	Organisation for Economic Cooperation and Development
PISA	Programme for International Student Assessment
PPS	Probability Proportional to Size
PGB	PISA Governing Board
PQM	PISA Quality Monitor
PSU	Primary Sampling Units
QAS	Questionnaire Adaptations Spreadsheet
RMSEA	Root Mean Square Error of Approximation
RN	Random Number
SC	School Co-ordinator
SE	Standard Error
SD	Standard Deviation
SEM	Structural Equation Modelling
SMEG	Subject Matter Expert Group
SPT	Study Programme Table
TA	Test Administrator
TAG	Technical Advisory Group
TCS	Target Cluster Size
TIMSS	Third International Mathematics and Science Study
TIMSS-R	Third International Mathematics and Science Study – Repeat
VENR	Enrolment for very small schools
WLE	Weighted Likelihood Estimates

PISA 2006 TECHNICAL REPORT – ISBN 978-92-64-04808-9 – © OECD 2009

Table of contents

PISA 2006 TECHNICAL REPORT – ISBN 978-92-64-04808-9 – © OECD 2009

PISA 2006 TECHNICAL REPORT – ISBN 978-92-64-04808-9 – © OECD 2009

PISA 2006 TECHNICAL REPORT – ISBN 978-92-64-04808-9 – © OECD 2009

PISA 2006 TECHNICAL REPORT – ISBN 978-92-64-04808-9 – © OECD 2009

PISA 2006 TECHNICAL REPORT – ISBN 978-92-64-04808-9 – © OECD 2009

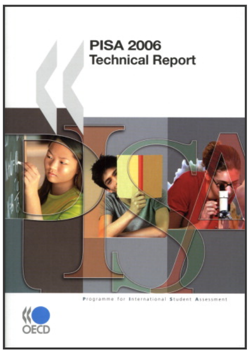
LIST OF BOXES

LIST OF FIGURES

LIST OF TABLES

PISA 2006 TECHNICAL REPORT – ISBN 978-92-64-04808-9 – © OECD 2009

PISA 2006 TECHNICAL REPORT – ISBN 978-92-64-04808-9 – © OECD 2009



From:
PISA 2006 Technical Report

Access the complete publication at:
https://doi.org/10.1787/9789264048096-en

Please cite this chapter as:

OECD (2009), “Field operations”, in *PISA 2006 Technical Report*, OECD Publishing, Paris.

DOI: https://doi.org/10.1787/9789264048096-7-en