



14

# Data adjudication

<b>Introduction</b> .....	264
<b>General outcomes</b> .....	267



## INTRODUCTION

The PISA Technical Standards (see Annex F) specify the way in which PISA must be implemented in each country, economy, and adjudicated region.<sup>1</sup> International contractors monitor the implementation in each of these and adjudicate on their adherence to the standards. This chapter describes the process used to adjudicate the implementation of PISA 2015 in each of the adjudicated entities (i.e. the participating countries, economies, and adjudicated regions<sup>2</sup>) and gives the outcomes of data adjudication that are mainly based on the following aspects:

- the extent to which each adjudicated entity met PISA sampling standards
- the outcomes of the adaptation, translation, and verification process
- the outcomes of the PISA Quality Monitoring visits
- the quality and completeness of the submitted data
- the outcomes of the international coding review.

## PISA 2015 Technical Standards

The areas covered in the PISA 2015 Technical Standards include the following:

### ***Data Standards***

- target population and sampling
- language of testing
- field trial participation
- adaptation of tests, questionnaires, and school-level manuals and scripts
- translation of tests, questionnaires, and school-level manuals and scripts
- test administration
- implementation of national options
- security of the material
- quality monitoring
- printing of material
- response coding
- data submission

### ***Management standards***

- communication with the international contractors
- notification of international and national options
- schedule for submission of materials
- drawing samples
- management of data
- archiving of materials

### ***National involvement standards***

- national feedback

## Implementing the standards – quality assurance

National Project Managers of participating countries, economies, and adjudicated regions are responsible for implementing the standards based on the international contractors' advice as contained in the various operational manuals and guidelines. Throughout the cycle of activities for each PISA survey, the international contractors carried out quality assurance activities in two steps. The first step was to set up quality control using the operational manuals, as well as the agreement processes for national submissions on various aspects of the project. These processes gave the international contractor staff the opportunity to ensure that PISA implementation was planned in accordance with the PISA 2015 Technical Standards and to provide advice on taking rectifying action when required and before critical errors occurred. The second step was quality monitoring, which involved the systematic collection of data that monitored the implementation of the assessment in relation to the standards. For data adjudication, it was the information collected during both the quality control and quality monitoring activities that was used to determine the level of compliance with the standards.



## Information available for adjudication

The international contractors' quality monitoring of a country's data collection is carried out from a range of perspectives during many stages of the PISA cycle. These perspectives include monitoring a country's adherence to the deadlines, communication from the sampling contractor about each country's sampling plan, information from the language verification team, data from the PISA Quality Monitors, and information gathered from direct interviews at National Project Manager and Coder Training meetings. The information was combined together in the database so that:

- indications of non-compliance with the standards could be identified early on in order to enable rectifying measures
- the point at which the problem occurred could be easily identified
- information relating to the same PISA standard could be cross-checked between different areas or sources.

Many of these data collection procedures refer to specific key documents, specified in the National Project Manager's Manual and the Sampling Manual in particular. These are procedures that the international contractors require for field trial and main survey preparation from each National Centre. The data adjudication process provides a motivation for collating and summarising the specific information relating to PISA standards collected in these documents, combined with information collected from specific quality monitoring procedures such as the PISA Quality Monitor visits and from information in the submitted data.

The quality monitoring information was collected from various quality monitoring instruments and procedures and covered the following main administrative areas:

- international contractors' administration and management: information relating to administration processes, agreement of adaptation spreadsheets, submission of information
- data analysis: information from item level reports, from the field trial data, and from data cleaning steps, including consistency checks
- school-level materials: information from the agreement of adaptations to test administration procedures and field operations
- Final Optical Check team: information from the pre- and post-main survey Final Optical Checks of main survey booklets
- main survey review: information provided by the National Project Managers in the main survey Review Questionnaire
- National Centre quality monitoring: information gathered through interviews conducted during meetings of National Project Managers or at other times
- co-ordination of PISA Quality Monitor activities including recruitment
- PISA Quality Monitor country reports: information gathered via the Data Collection Forms from PISA Quality Monitors and through their interactions with School Co-ordinators and Test Administrators
- sampling: information from the submitted data such as school and student response rates, exclusion rates and eligibility problems
- translation: information relating to the verification and translation process
- National Centre Test Administrator or School Associate trainings
- National quality monitoring issues
- data cleaners: issues identified during the data cleaning checks and from data cleaners' reports
- item developers: issues identified in the coder query service and training of coders;
- data processing: issues relating to the eligibility of students tested
- questionnaire data: issues relating to the questionnaire data in the national questionnaire reports provided by the international contractor
- questionnaire Final Optical Check: issues arising from the Final Optical Check of the questionnaires.

### Quality monitoring reports

There were two types of PISA quality monitoring reports: The Session Report Form containing data for each session in each school, and the Data Collection Form detailing the general observations across all schools visited by PQMs. The Session Report Form was completed by the Test Administrator after each test session and also contained data related to test administration. The data from this report were recorded by the National Centre and submitted as part of the national dataset to ETS. The PISA Quality Monitor reports contained data related to test administration in selected schools, and the PISA quality monitoring data were collected independently of the National Project Manager. Additional information on all the standards was also noted in the main survey Review. The main survey Review was self-declared by the National Project Manager.



## Data adjudication process

The main aim of the adjudication process is to make a judgement on each national dataset in a manner that is transparent, based on evidence, and defensible. The data adjudication process achieved this through the following steps:

**Step 1:** Quality control and quality monitoring data were collected throughout the survey administration period.

**Step 2:** Data collected from both quality control and quality monitoring activities were entered into a single quality assurance database.

**Step 3:** Experts compiled country-by-country reports that contained quality assurance data for key areas of project implementation.

**Step 4:** Experts considered the quality assurance data that were collected from both the quality control and quality monitoring activities to make a judgement. In this phase, the experts collaborated with the international contractors to address any identified areas of concern. Where necessary, the relevant National Project Manager was contacted through the contractors. At the end of this phase experts constructed, for each adjudicated dataset, a summary detailing how the PISA Technical Standards had been met.

**Step 5:** The adjudication group, formed by representatives of the OECD and of international contractors, the Technical Advisory Group and the Sampling Referee, reviewed the reports and made a determination with regard to the quality of the data from each adjudicated entity.

Monitoring compliance to any single standard occurred through responses to one or more quality assurance questions regarding test implementation and national procedures which may come from more than one area. For example, the session report data were used in conjunction with the PISA Quality Monitor reports, computer system tracking of timings, and information from the adaptation of national manuals to assess compliance with the PISA session timing standard (Standard 6.1, Annex F).

Information was collected in relation to these standards through a variety of mechanisms:

- through PISA Quality Monitor reports
- through the field trial and main survey reviews
- through information negotiated and stored on the PISA Portal website (the portal which was used in PISA 2015)
- through a system database specific to the implementation of PISA tasks
- through the formal and informal exchanges between the international contractors and National Centres over matters such as sampling, translation and verification, specially requested analyses (such as non-response bias analysis)
- through a detailed post-hoc inspection of all main survey assessment materials (test booklets)
- through the data cleaning and data submission process.

For PISA 2015, an adjudication database was developed to capture, summarise, and store the most important information derived from these various information sources. The staff members of the international contractor who led each area of work were responsible for identifying relevant information and entering it into the database. This means that at the time of data adjudication, relevant information was easily accessible for making recommendations about the fitness of use of data from each PISA adjudicated entity.

The adjudication database captured information related to the major phases of the data operation: field operations, sampling, computer-based problem solving and computer-based assessment of Financial Literacy (where applicable), questionnaires, and tests. Within each of these phases, the specific activities are identified, and linked directly to the corresponding standards.

Within each section of the database, specific comments are entered that describe the situation of concern, the source of the evidence about that situation, and the recommended action. Each entry is classified as serious, minor, or of no importance for adjudication. Typically, events classified as serious would warrant close expert scrutiny and possibly action affecting adjudication outcomes. For example, cognitive data for Kazakhstan were found to be inconsistent across human-coded and machine-scored items, and upon further investigation, the coding of human coded items was invalidated, resulting in the exclusion of Kazakhstan's performance scores from international comparisons and comparisons with results for Kazakhstan from previous years. Events classified as minor would typically not directly affect adjudication outcomes but will be reported back to National Centres to assist them in reviewing their national procedures.



## Data adjudication

It was expected that the data adjudication would result in a range of possible recommendations to the PISA Governing Board. Some possible, foreseen recommendations included:

- that the data be declared fit for use
- that some data be removed for a particular country, economy, or adjudicated region, such as the removal of data for some open-ended items or the removal of data for some schools
- that rectifying action be performed by the National Project Manager, such as providing additional evidence to demonstrate that there was no non-response bias, or rescored open-ended items
- that the data not be endorsed for use in certain types of analyses
- that the data not be endorsed for inclusion in the PISA 2015 database.

Throughout PISA 2015, the international contractors concentrated their quality control activities to ensure that the highest scientific standards were met. However, during data adjudication a wider definition of quality was used, especially when considering data that were at risk. In particular, the underlying criterion used in adjudication was fitness for use; that is, data were endorsed for use if they were deemed to be fit for meeting the major intended purposes of PISA.

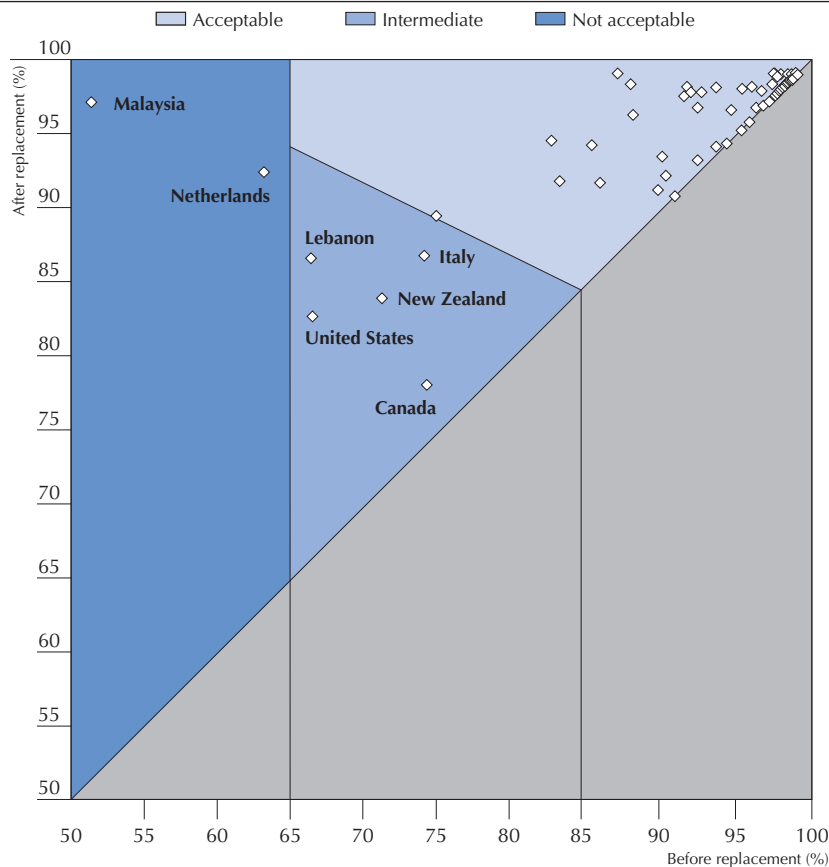
## GENERAL OUTCOMES

### Overview of response rate issues

The PISA school response rate requirements are discussed in Chapter 4. Figure 14.1 is a scatter plot of the attained PISA school response rates before and after replacements. Those countries that are plotted in the light blue shaded region were regarded as fully satisfying the PISA school response rate criterion.

Figure 14.1

Attained school response rates



Source: PISA 2015 Technical Report, Tables 11.3 and 11.5.



Seven countries – Canada, Italy, Lebanon, Malaysia, the Netherlands, New Zealand and the United States – failed to meet the school non-response rate (see Figure 14.1). After reviewing the sampling outcomes, the consortium asked these seven countries to provide additional data that would assist the consortium in making a balanced judgement about the threat of the non-response to the accuracy of inferences which could be made from the PISA data.

Belgium, Hong Kong (China), the United Kingdom and, among adjudicated subnational entities, the Flemish community (Belgium) and Massachusetts (United States), had a response rate below the 85% level before the use of replacement schools but cleared the acceptable level after the replacement schools were included.

One country – Trinidad and Tobago – fell short of the student response-rate standard; however, in consideration of the fact that there were no student-level exclusions and that the achieved response rate (79.4%) was very close to the acceptable rate of 80%, the international contractor determined that the data were acceptable (see Table 11.7).

### **Detailed country comments**

It is important to recognise that PISA data adjudication is a late but not necessarily final step in the quality assurance process. By the time each country was adjudicated at the Technical Advisory Group meeting in June 2016, the quality assurance and monitoring processes outlined earlier in this chapter and in Chapter 7 had been implemented. Data adjudication focused on residual issues that remained after these quality assurance processes had been carried out.

The remaining issues fall under two broad categories: (1) adaptations to the recommended international standard procedures in a country's data collection plan, and (2) a failure to meet international standards at the implementation stage.

#### ***Departures from standard procedures in the national data collection plan***

With such a broad and diverse range of participation, it is to be expected that the international best practice approaches to data collection articulated in the PISA Technical Standards document may not be achieved in all national and local contexts. This may be the case for a number of reasons. For example, it may be contrary to national protocols to have unannounced visits of quality monitors to schools to observe test administration. Or, it may not be possible for teachers from very remote or very small schools to leave their schools to attend training in the mechanics of PISA test administration. Typically these were discussed with international contractor experts in advance of the assessment, and alternative approaches were considered jointly between the National Project Manager and the international contractor. In isolated departures from best practice in cases such as these, a judgement might easily be made by international contractor experts that there was minimal risk to the quality of the data collection plan. Such isolated departures are not reported in the country summaries below.

On the other hand, it may not have been straightforward to determine in advance of the assessment how more extensive or multiple departures from PISA Technical Standards may interact with each other and with other aspects of a country's data collection plan. Cases such as these were considered as part of the data adjudication process and are included in the country summaries below.

#### ***Departures from standards arising from implementation***

Departures from the standards at the implementation stage range from errors within the National Centre (e.g., during the final stages of preparing materials, or in the administration of the coding operation following data collection), a failure to meet documented targets during data collection, for example a shortfall from the minimum school and student sample sizes.

A component of the data adjudication process was to consider the cases of multiple, or more complex departures from the PISA standard procedures, as well as to consider the impact of errors or shortfalls across all aspects of each country's data collection plan and implementation, and make an evaluation with respect to the quality and international comparability of the PISA results. Notable departures from the standards are reported in the country summaries below.

Several countries exceeded the limit on student- and school-level exclusions (5% at most; see Tables 11.1 and 11.2). In countries where other violations of sampling standards were observed or where the combined level of exclusions exceeded 6%, further information was requested to support the case that no bias would result from exclusions. The number of such cases shows a notable increase over the level observed in previous cycles; it is unclear whether the increase in exclusion rates observed in several countries over previous cycles must be attributed to the first implementation of computer-based assessment, to an increase in migrant populations, or to local, idiosyncratic factors.



A small number of countries failed to reach the required minimum sample sizes of 5 400 students and 150 schools (4 500 students and 150 schools for countries that tested in paper mode or did not participate in the testing of collaborative problem solving; the numbers for additional adjudicated entities are 1 800 students and 50 schools and 1 500 students and 50 schools, respectively). Such cases were considered as part of the data adjudication process. Even a minor deviation in sample size might be considered a substantive enough issue to report, for example in countries where standard errors tend to be higher for a given sample size. On the other hand, minor deviations from these minimal sample sizes (i.e. shortfalls of fewer than 50 students or 5 schools, and in countries that nevertheless achieved comparable standard errors on the major survey estimates) are not reported below.

Particular attention has been paid to the achievement of the specified response rates of 85% for schools, 80% for students within schools and no more than 5% of students excluded from the assessment. Seven countries were required to provide additional data to support the case that no bias would result from failure to meet the response-rate standards.

Anomalies in submitted data, particularly inconsistencies and deviations from the expected patterns, were also investigated; most cases could be explained and solved through a resubmission of data. The two cases that could not be solved are noted below.

If a country is not listed below then it fully met the PISA standards. Further, in the case of minor deviations from the standards, unless otherwise noted, additional data were usually available to suggest the data were suitable for use.

## Albania

Analysis of the data submitted by Albania suggested that the PISA Technical Standards were not fully met. Indeed, the relationships between student achievement and student background characteristics collected through self-report questionnaires were, without exceptions, very weak, deviating from associations found in Albania in previous cycles and from the patterns observed in other countries. There was no association, for instance, between reading performance and the reported number of books at home. However, school-level associations and relationships with student characteristics collected with student tracking forms (such as gender or grade) conformed with expectations. A mismatch between cognitive booklets and questionnaire booklets, whereby the same student identifier was used for different students within the same school, is suspected.

The PISA 2015 international database therefore does not include information collected through student questionnaires for Albania. An additional dataset, which uses different student identifiers, contains this information. No attempt should be made to link the student data included in the international PISA database with the additional dataset for Albania.

## Argentina

In Argentina the review of sampling outcomes revealed that the national defined target population (the population of all 15-year-old students enrolled in grade 7 and above in schools listed in the sampling frame for PISA, and not marked for exclusion) deviated significantly from the desired target population.

The sampling frame submitted in February 2015 contained much fewer schools than in previous cycles, but showed good agreement between the sum of school-level enrolment figures and the overall enrolment expected, given the experience of previous cycles and national enrolment statistics. Despite reassurance about the completeness of the sampling frame, the actual enrolment figures collected from sampled schools in June-July 2015 fell significantly short (by about 30%) of the expected number of 15-year-old students in each school. As a result, the coverage of the sampling frame enrolment (Coverage index 4), in 2015, was only 69%. Further information was therefore requested to exclude that the “missing” students were enrolled in schools that were not listed on the sampling frame.

Upon further investigation, it was found that three categories of schools were omitted from the sampling frame:

- Schools that had been created or renamed between 2013 and 2015. Due to an ongoing restructuring policy that merged lower- and upper-secondary schools into new unified secondary schools, the number of students attending newly created or renamed schools was higher than usual.
- Schools that were listed incorrectly in the official registry of educational institutions “administrative headquarters” with no students. This error in the official registry of educational institutions, one of the sources used to identify schools with 15-year-old students, was detected only in October 2015.
- Finally, some provincial authorities did not include all rural schools as they had very few students.



Together, these omissions resulted in the exclusion of well over 10% of the desired target population from the sampling frame, creating a significant threat of non-coverage bias. However, the investigation conducted in June 2016 also highlighted that no eligible school in the adjudicated entity of Ciudad Autónoma de Buenos Aires was missing from the original sampling frame.

The failure to meet standard 1.7 about the definition of the target population and school exclusions implies that Argentina's results may not be representative of the whole country, and may therefore not be comparable to those of other countries or to results for Argentina from previous years. Data for Argentina were therefore not included in the international dataset, and are available as a separate dataset. Data for Ciudad Autónoma de Buenos Aires, on the other hand, were deemed to be acceptable and are fully included in the international dataset.

## **Australia**

There was a total of 5.31% exclusions in Australia; data were included in the final database.

## **Canada**

There was a total of 7.49% exclusions in Canada. Exclusion rates however were only marginally higher than in previous cycles, and, as in 2012, exclusions were mostly due to students with special needs.

Canada had a weighted school response rate of 74.5% before replacement. After replacement, the response rate was 78.6% (a response rate of 90.3% was required in order to fully meet the standard). Further information was sought from Canada to support the case that no notable bias would result from non-responses.

Additional analyses indicated that much of Canada's non-responses came from the relatively large provinces of Alberta, Ontario and Quebec, with the latter province achieving the lowest school response rate among all provinces; all other provinces had acceptable response rates. Consequently, the non-response bias analysis compared the characteristics of the target population and of responding, non-responding, and replacement schools in these three provinces only. In addition to school type and language, assessment results in local assessments were available for all students in Alberta, and at the school level in Ontario (for sampled schools only) and Quebec. Canada presented evidence to show that the characteristics of non-responding schools in Alberta and Ontario were not markedly different from those of respondent schools, while for Quebec the comparison of mean achievement scores at the school level in an assessment of reading and science showed more significant differences.

The adjudication group concluded that no notable bias would result from non-response in the Canadian data, when analysed at the national level, and inclusion in the full range of PISA 2015 reports was recommended. However, caution was invited when reporting data for the province of Quebec in isolation, due to a possible non-response bias.

## **Denmark**

There was a total of 5.04% exclusions in Denmark; data were included in the final database.

## **Estonia**

There was a total of 5.52% exclusions in Estonia; data were included in the final database.

## **Italy**

Italy had a weighted school response rate of 74.4% before replacement. After replacement, the response rate was 87.5% (a response rate of 90.3% was required in order to fully meet the standard).

Additional analysis indicated that a number of characteristics were balanced across respondents and non-responding schools. The higher level of refusals to participate compared to past administrations was explained by the burden created by bigger-than-usual samples, due to the addition of a grade-based sample as national option, in the context of a computer-based administration.

The adjudication group concluded that no notable bias would result from non-response. The data for Italy, therefore, were included in the international database.





## Kazakhstan

During the consistency checks performed on cognitive data prior to scaling, it was found that scores for human-scored items submitted by Kazakhstan were inconsistent with the success rates observed in prior PISA cycles and were almost unrelated to scores on machine-scored (multiple choice) items. Further information was sought to ensure that the coded responses reflected student responses in an authentic way.

Kazakhstan was asked to send 300 randomly selected booklets, in both national languages, for re-coding. This independent coding performed by international coders indicated significant leniency among national coders. However, even the lower scores assigned by international coders deviated from the expected patterns of association with machine-scored items for the same students. The evidence was deemed sufficient to conclude that the data submitted by Kazakhstan did not meet coding standards.

The adjudication group recommended to invalidate all human-scored items. Furthermore, because an assessment that is limited to multiple-choice items would not provide adequate coverage of the constructs measured in PISA and could not be considered comparable to other countries or to past results, the remaining data were not deemed to be fit for inclusion in the international database, and only limited reporting was recommended. Data for Kazakhstan are available as a separate dataset and caution must be exerted when comparing the results for Kazakhstan to past cycles or to results for other countries.

## Latvia

There was a total of 5.07% exclusions in Latvia and there were fewer than the 5400 students specified in the standards for a country or economy participating in the assessment of collaborative problem solving (4845); data were included in the final database.

## Lebanon

Lebanon had a weighted school response rate of 66.6% before replacement. After replacement, the response rate was 87.3% (a response rate of 94.2% was required in order to fully meet the standard). Additional analysis showed that language of instruction and the distribution of school enrolment did not differ significantly between the original sample and the final responding sample using replacements; no other characteristics were available for a non-response bias analysis. The adjudication group nevertheless concluded that it was unlikely that notable bias would result from non-response in the final database. The data for Lebanon, therefore, were included in the international database.

## Lithuania

There was a total of 5.12% exclusions in Lithuania; data were included in the final database.

## Luxembourg

There was a total of 8.16% exclusions in Luxembourg. In consideration of the consistency in the level and nature of exclusions across cycles in Luxembourg, data were deemed to be acceptable and included in the database.

## Malaysia

In Malaysia, the weighted response rate among the initially sampled schools (51.4%) fell short of the 65% minimal threshold, and corresponds to an unacceptable response rate (after the use of replacement schools, the response rate was 98.1%). Malaysia submitted a non-response bias analysis which showed that responding replacement schools had significantly better result, on a national examination, than non-responding schools in the original sample. The adjudication group concluded that in this case, non-response may have introduced bias in comparisons of Malaysia's results with those of other countries or with previous years, and recommended limited reporting. Data for Malaysia are included in a separate database.

## Montenegro

There was a total of 5.17% exclusions in Montenegro; data were included in the final database.

---

## **Netherlands**

In the Netherlands, the weighted response rate among the initially sampled schools (63.3%) fell slightly short of the 65% minimal threshold, and corresponds to an unacceptable response rate (after the use of replacement schools, the response rate was 93.2%). Additional analysis however, using school results on a central examination of mathematics and science subjects, indicated that within each stratum the mean and distribution of results of responding schools (including replacements) did not differ significantly from the target population. The adjudication group concluded that no notable bias would result from non-response in the final database. The data for the Netherlands, therefore, were included in the international database.

## **New Zealand**

There was a total of 6.54% exclusions in New Zealand. Available information indicated that the extra students excluded were all students with limited language ability, recently arrived in the country.

There were fewer than the 5400 students specified in the standards for a country or economy participating in the assessment of collaborative problem solving (4520).

New Zealand also had a weighted school response rate of 71.4% before replacement. After replacement, the response rate was 84.5% (a response rate of 91.8% was required in order to fully meet the standard). Additional analysis, using in particular school results on the New Zealand National Certificate in Educational Achievement, indicated that the mean and distribution of results did not differ significantly between responding schools (including and excluding replacements) and the original sample. The adjudication group concluded that no notable bias would result from non-response in the final database. The data for New Zealand, therefore, were included in the international database.

## **Norway**

There was a total of 6.75% exclusions in Norway. Most of the excess students excluded were students with limited language ability, recently arrived in the country, and data were therefore deemed to be acceptable and included in the final database.

## **Spain**

### ***Castile and Leon***

There was a total of 5.08% exclusions in Castile and Leon; data were included in the final database.

### ***Catalonia***

There was a total of 5.41% exclusions in Catalonia; data were included in the final database.

### ***Valencia***

There was a total of 7.41% exclusions in Valencia; and there were fewer than the 1 800 students specified in the standards for an adjudicated entity participating in the assessment of collaborative problem solving (1611). Data were included in the final database.

## **Sweden**

There was a total of 5.71% exclusions in Sweden. Available information indicated that the extra students excluded were all students with limited language ability, recently arrived in the country, and data were therefore deemed to be acceptable and included in the final database.

## **Trinidad and Tobago**

Trinidad and Tobago had a weighted student response rate of 79.4%, slightly below the standard of 80%. In consideration of the fact that there were no school or student-level exclusions, the adjudication group concluded that no notable bias would result from non-response in the final database. The data for Trinidad and Tobago, therefore, were included in the international database.



## United Kingdom

There was a total of 8.22% exclusions in the United Kingdom, and a marked increase in exclusions due to special needs over previous cycles. The national centre for the United Kingdom (excluding Scotland) explained this as a possible unintended consequence of changes in the timing at which information about special needs was collected (in student tracking forms rather than in student lists collected prior to sampling); this could have led school coordinators to mark more students for exclusion. In consideration of the fact that appropriate actions had been taken to limit exclusions once this issue had been detected, data were deemed to be acceptable and included in the final database.

## Scotland

There was a total of 6.52% exclusions in Scotland; data were included in the final database.

## United States

The United States had a weighted school response rate of 66.7% before replacement. After replacement, the response rate was 83.3% (a response rate of 94.2% was required in order to fully meet the standard). Additional analysis, using data from two school surveys and from the sampling frame, indicated that among originally sampled schools, region and student race/ethnicity differed across responding and non-responding schools. However, after replacement schools were added to the respondents, all available characteristics (school enrolment, control and location, region, race/ethnicity and, for public schools, poverty) were balanced with those of the initially selected school sample, therefore showing that the use of replacement schools substantially reduced the potential for bias. The adjudication group concluded that no notable bias would result from non-response in the final database. The data for the United States, therefore, were included in the international database.

## Massachusetts (Public schools)

There were fewer than the 1 800 students specified in the standards for an additional adjudicated entity participating in the assessment of collaborative problem solving (1391). Data were included in the final database.

## Puerto Rico<sup>3</sup>

There were fewer than the 1 500 students specified in the standards for an additional adjudicated entity testing in paper mode (1398). Data were included in the final database.

## Notes

1. For the remainder of this chapter, we will use the term “country” when referring to a country, economy, or adjudicated region.
2. Not all regions opt to undergo the full adjudication that would allow their results to be compared statistically to all other participating economies and adjudicated regions. For example, the states of Australia are not adjudicated regions, whereas the Flemish Community of Belgium is an adjudicated region.
3. Puerto Rico is an unincorporated territory of the United States. As such, PISA results for the United States do not include Puerto Rico.