

PISA 2022 Technical Report



16 Data Adjudication

Introduction

Data adjudication is the process through which each national dataset is reviewed and a judgement about the appropriateness of the data for the main reporting goals is formed. The PISA Technical Standards (see Annex I) specify the way in which PISA must be implemented in each participating jurisdiction and adjudicated region. 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 PISA 2022 data for each of the adjudicated entities (i.e., the participating countries and economies – hereafter, “jurisdictions” – and the adjudicated regions) 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, including concerns about the quality of the data that were identified during scaling and in preparation for reporting
- the outcomes of the international coding review.

Not all regions (i.e., subnational jurisdictions that report their results separately) 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.

PISA 2022 Technical Standards

The areas covered in the PISA 2022 Technical Standards include several aspects connected to the implementation of PISA, including the definition and sampling of its target population on appropriate languages for testing, translation and adaptation of materials, school and student participation in the Field Trial and Main Survey, test administrations and handling of test materials, coding of responses, data management, privacy, and submissions, to cite a few key aspects. A comprehensive list of Technical Standards used for adjudication is available in Table 16.1.

Implementing the standards – quality assurance

National Project Managers of participating jurisdictions 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- assurance procedures 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 2022 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 the data adjudication, information collected during both the quality-assurance and quality-monitoring activities was used to determine the level of compliance with the standards.

Information available for adjudication

The international contractors' quality monitoring of a participating jurisdiction's data collection is carried out from a range of perspectives during many stages of the PISA cycle. These perspectives include monitoring a participating jurisdiction's adherence to the deadlines, communication from the sampling contractor about each participating jurisdiction's sampling plan, information from the linguistic 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 said non-compliance 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 Technical 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 areas:

- international contractors' administration and management: information relating to administration processes, agreement of adaptation spreadsheets, submission of information.
- translation: information from linguistic verification of test items, questionnaire items, and the test administration script.
- sampling: information from the submitted data such as school and student response rates, exclusion rates and eligibility problems.
- school-level materials: information from the agreement of adaptations to test administration procedures and field operations.
- student materials: information from the pre- and post- Main Survey final optical checks of MS test booklets and background questionnaires.
- National Centre operations: School Coordinator, Test Administrator or School Associate trainings; information gathered through interviews conducted during meetings of National Project Managers or at other times.
- PISA Quality Monitors (PQMs): co-ordination of PISA Quality Monitor activities including recruitment; information gathered via the Data Collection Forms from PQMs and through their interactions with School Co-ordinators and Test Administrators.
- data cleaners: issues identified during the data cleaning checks and from data cleaners' reports.
- data processing: issues relating to the eligibility of students tested; issues identified in the coder query service and training of coders.

- data analysis: information from item level reports, from the Field Trial data, and from data cleaning steps, including consistency checks.
- questionnaire data: issues relating to the questionnaire data in the national questionnaire reports provided by the international contractor.
- Main Survey and Field Trial Reviews: information provided by the National Project Managers in the Field Trial and Main Survey Review 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 international contractor in charge of coordinating PISA implementation (Core A, see Chapter 1) where it was aggregated by the international project manager at the contractor. The PQM reports contained data related to test administration in selected schools, and the PISA quality monitoring data were collected independently of the National Project Manager.

Data adjudication process

Data adjudication is the process through which each national dataset is reviewed and a judgement about the appropriateness of the data for the main reporting goals is formed. The different steps in the data adjudication process ensure that the final judgement is transparent, based on evidence, and defensible.

The data adjudication process achieved this through the following steps:

Step 1: International contractors collected quality-assurance and quality monitoring data throughout the survey administration period. The international project manager compiled this information into an adjudication database that was updated or amended as new information arose and provided an overview of the national implementation of PISA throughout the cycle.

Step 2: The international project manager compiled individual reports for each jurisdiction that contained quality-assurance data for key areas of project implementation.

Step 3: The international project director, together with the international contractor leads, identified data issues that were in need of adjudication. Where necessary, the relevant National Project Manager was contacted to provide additional information. After this stage, for each dataset, a summary report detailing whether and how the PISA Technical Standards had been met was drafted.

Step 4: The PISA Adjudication Group, formed by representatives of the OECD, of international contractors, the Technical Advisory Group and the Sampling Referee, reviewed the summary reports to recommend adequate treatment of the data from each adjudicated entity in international PISA products (database and reports).

Step 5: The recommendations of the PISA Adjudication Group were presented to the PISA Governing Board representatives and to the countries concerned.

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 I and Table 16.1).

Information was collected in relation to these standards through a variety of instruments and information sources:

- through PISA Quality Monitor reports
- through the Field Trial and Main Survey reviews submitted by National Centres
- through information negotiated and stored on the communications portal for PISA 2022
- 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
- through the data cleaning and data submission process.

For PISA 2022, an adjudication database was developed to capture, summarise, and store the most important information derived from these various sources. International contractor staff 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 appropriate and comparable 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, 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. 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.

The adjudication process for PISA 2022 had an increased challenge imposed by the COVID-19 global pandemic that caused school closures worldwide. The onset of the pandemic was in early 2020 – right as early testing countries were about to implement their Field Trial data collections for the – then named – PISA 2021 cycle.

School closures affected education systems differently, but overall, significantly disrupted survey operations, which led to a decision by the PISA Governing Board to postpone data collection for one year, thus renaming the cycle PISA 2022. As schools reopened and instruction was normalized throughout 2021 and 2022, data collection for PISA resumed, but nonetheless required a degree of reactivity and flexibility from education systems and international contractors. Indeed, as schools reopened at an uneven pace throughout jurisdictions and attempted to get back on track for the rest of the (sometimes expedited) school year, participation of said schools in PISA 2022 Main Survey data collection could not be taken for granted, neither could access from Test Administrators to students, or even high attendance of the latter.

These limitations compelled some changes to Standard 1.3 (assessment period), namely:

- Extension of the assessment period beyond 56 days where students remain within the PISA-eligible age range would be agreed to with the OECD's implicit approval.
- Extension of the assessment period that would not exceed the allowed 56 days but would result in assessed students who are outside of the PISA-eligible age range **by less than a week** would be agreed to with the OECD's implicit approval.
- Extension of the assessment period that would both exceed 56 days AND result in assessed students who are outside of the PISA-eligible age range will require further consultation with the contractors and the OECD before approval of such a deviation would be granted.

The changes were proposed by the international contractors, endorsed by the PISA Technical Advisory Group (TAG) on its December 2021 meeting, and implemented throughout PISA 2022 MS data collection. All participating jurisdictions managed to successfully conclude MS data collection using this added flexibility.

Data adjudication outcomes

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 the data be declared fit for use with explicit cautions advised regarding its representativity of the jurisdiction's student cohort, or international comparability of its results
- that some data be removed for a particular participating jurisdiction 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 rescoring 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 2022 database.

Throughout PISA 2022, 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

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 participating jurisdiction was adjudicated at the adjudication group meeting in July 2023, the quality assurance and monitoring processes outlined earlier in this chapter, foreseen in the Technical Standards, and described throughout this report had been implemented. Data adjudication focused on residual issues that remained after these quality assurance processes had been carried out.

Overall, the Adjudication Group's review suggests good adherence of national implementations of PISA to the technical standards in spite of the challenging circumstances that affected not only PISA operations but schooling more generally during the COVID-19 pandemic. Thanks to the reactivity and flexibility of participating countries and international contractors, to carefully constructed instruments, to a test design that is aligned to the main reporting goals and is supported by adequate sample design, and to the use of appropriate statistical methods for scaling, population estimates are highly reliable and comparable across countries and time, and particularly with 2018 results.

Nevertheless, a number of deviations from standards were noted and their consequences for data quality were reviewed in depth. The following overall patterns of deviations from standards were identified:

- About one in five of all adjudicated entities had exclusion rates exceeding the limits set by the technical standards (Standard 1.7).
- Seven entities failed to meet the required school response rates, with three of them failing to meet the stricter level of 65% before replacement (Standard 1.11). This is in line with earlier cycles of PISA.

- There was a significant increase in the number of entities that failed to meet the required student response rates (Standard 1.12): 10 entities did not meet this standard.
- There were delays in data submission in a significant number of entities (Standard 19.1): 14 entities did not meet this standard, and 13 only partially met it. The Adjudication Group noted that delayed submissions may affect the quality of the international contractors' work; and if shorter reporting timelines are expected, it may no longer be possible to accommodate such delays.
- A large number of entities did not conduct the field trial as intended (Standard 3.1) or did not attend all meetings (Standard 23.1). While this may also be a consequence of the pandemic, the Adjudication Group noted that these violations may be particularly consequential for new participants and for less-experienced teams. The Group underlined the importance of attendance at coder training sessions for ensuring comparability of the data.

At the international level, these frequent deviations should guide future efforts of the PISA Governing Board, the OECD Secretariat and Contractors to review the corresponding standards, prevent future deviations from standards, or mitigate the consequences of such violations.

At the level of individual adjudicated countries, economies and regions, in most cases, these issues did not result in major threats to the validity of reports, and the data could be declared fit for use. Where school or student participation rates fell short of the standard and created a potential threat for non-response/non-participation bias, countries/economies were requested to submit non-response-bias analyses. The evidence produced by countries/economies (and in some cases, by the sampling contractor) was reviewed by the Adjudication Group.

The Adjudication Group reviewed and discussed major adjudication issues in June 2023. The major adjudication issues reviewed by the group fall under three broad categories: (1) exclusions and response rates, (2) invariance of item parameters, and (3) issues originating from the Chromebook online administration pilot.

Overview of exclusion and response rate issues

Exclusion rates

PISA Technical Standard 1.7 states that the target population - the population that sits PISA - covers at least 95% of the desired population, the one for which broader conclusions from the assessment are sought. This means that overall exclusions, at both school and student levels, cannot exceed 5%. Sixteen jurisdictions excluded students in excess of this threshold at varying degrees.

Data collection was severely disrupted by the onset of Russia's war of aggression on February 2022, meaning that only data for 18 out of Ukraine's 27 regions could be collected. Exclusions were computed with respect to the original sampling frame, covering the entire country. After February 2022, however, survey operations could not be completed successfully in the regions most affected by wartime disruptions adding to an exclusion rate of 36.1%. Results for the remaining regions were deemed fit for reporting, but comparisons with previous results should be made only with great caution, and with due consideration to the differences in target populations. The Adjudication Group recommended that the results are presented in such a way to alert readers to the difference in the target population between prior cycles and PISA 2022.

Exclusions in Denmark increased by a large margin, presenting a marked increase compared to previous cycles, at 11.6% in PISA 2022. The Adjudication Group noted that high levels of student exclusions may bias performance results upwards. In Denmark, a major cause behind the rise appears to be the increased share of students with diagnosed dyslexia, and the fact that more of these students are using electronic assistive devices to help them read on the screen, including during exams. The lack of such

accommodation in PISA led schools to exclude many of these students, meaning that rates are likely to fall should said accommodations, especially those supporting dyslexic students, are allowed.

Rather elevated rates were also observed in the Netherlands and in Latvia, at 8.4% and 7.9% of schools respectively. On the former participant, the Dutch National Centre submitted non-response bias analysis was submitted, analysing differences in performance and in other characteristics between responding schools and the total population of schools, as well as differences between replacement schools and originally sampled, but non-responding schools. This supported the case that no large bias would result from non-response; furthermore, given the available evidence, there is no clear indication about the direction of any residual bias.

The Adjudication Group noted that exclusions exceeded the acceptable rate by a small margin in:

- Croatia (5.4%), Lithuania (6.5%), and in the United States (6.1%), which showed a marked increase in exclusions due to students with functional or intellectual disabilities, which are also bound to fall in the presence of increased accessibility in future cycles. The Adjudication Group invited the national centres to investigate the reasons for this increase in exclusion rates and take remedial action for future cycles. It is expected that exclusion rates will fall again in the future, as a result.
- Australia (6.9%), Canada (5.8%), Estonia (5.9%), New Zealand (5.8%), Norway (7.3%), Scotland (6.6%), Switzerland (5.8%), and Türkiye (5.6%) where the exclusion rates observed in 2022 remained relatively close to exclusion rates observed in 2018.
- Sweden, (7.4%), which showed a marked decrease from the high levels observed in 2018. The Adjudication Group noted that this might be the combined result of falling rates of refugee students and of the national centre's further effort to ensure uniform application of guidelines across schools.

School response rates

The PISA school response rate requirements are foreseen by Technical Standard 1.11, stating that all jurisdictions are to reach a weighted school response rate of 85%, which can be accomplished by administering PISA in replacement schools as needed. A comprehensive account of both jurisdictions and adjudicated regions' response rates either before and after replacement can be found in Tables 13.3, 13.4, 13.5, and 13.6 respectively in Chapter 13 of this report.

Seven jurisdictions did not meet the 85% school participation threshold before replacements, with only two meeting the standard with the use of replacement schools. Nonetheless, the increase in participation of replacement schools was rather heterogeneous across jurisdictions. This lack of response, in particular at the school level, has the potential to induce bias in observed results, and thus further investigations in the form of a non-response bias analysis (NRBA) were conducted by affected jurisdictions supported by the international sampling contractor.

Such is the case of the Netherlands, where 66% of sampled schools responded, a share that increased to 90% upon replacement. A NRBA was submitted, analysing differences in performance and in other characteristics between responding schools and the total population of schools, as well as differences between replacement schools and originally sampled, but non-responding schools. This supported the case that no large bias would result from non-response; furthermore, given the available evidence, there is no clear indication about the direction of any residual bias.

Similarly, in Chinese Taipei, where the standard was nearly met before and after replacement (83% before replacement, 84% after), a thorough NRBA was produced, using school-level achievement data as auxiliary information, which provided convincing evidence that the potential bias is minimal after non-response adjustments are considered.

The effect of replacement schools was less pronounced in the United States, where 51% of schools responded before replacement, with the school response rate going up to 62% once replacement schools were invited to participate. Participation rates thus missed the standard by a significant margin, with particularly low participation rates among private schools (representing about 7% of the student population). A NRBA was submitted, indicating that, after replacement schools and non-response adjustments are considered, a number of characteristics (not including direct measures of school performance) are balanced across respondents and non-respondents. Based on the available information, it was not possible for the Adjudication Group to exclude the possibility of bias, nor to determine its most likely direction.

Four other jurisdictions: Canada (81% before replacement, 86% after), Hong Kong (China; 60% before replacement, 80% after), New Zealand (61% before replacement, 72% after), and the United Kingdom (excl. Scotland; 66% before replacement, 80% after) have also submitted NRBA's supporting the case that any bias resulting from school non-participation is most likely be negligible, and are discussed in further detail below, as student participation was also a concern.

Student response rates

Technical Standard 1.12 states that student response rates must be in excess of 80% across responding schools. Students are not replaced in PISA, and thus only those sampled and present at the testing sessions are to sit the test. Student response rates for all jurisdictions and adjudication entities can be found at Tables 13.7 and 13.8 in Chapter 13 of this report.

Ten jurisdictions did not meet this standard. Albeit some observed rates were close to the 80% threshold, a downwards trend in student participation is of particular concern. Checking for this potential bias is particularly significant, as students' absence from school for the PISA 2022 cycle comes in the aftermath of a global pandemic with severe economic consequences, which might affect some students more than others..

In Malta the student response rate observed in PISA 2022 (79%), fell short of the standard by a small margin, but decreased significantly (from 86%) compared to PISA 2018. A thorough non-response bias analysis (NRBA) was produced, using student-level academic track variable as auxiliary information, along with demographic characteristics. Because students were tracked in the previous academic years based on their grades in mathematics, track information can be expected to correlate strongly with performance on the PISA test. The NRBA provides convincing evidence that the potential bias is minimal after non-response adjustments are taken into account.

Similarly, in the United Kingdom (excl. Scotland), student response rates decreased to 75% from 83% with respect to PISA 2018. School response rates also fell short of the target as discussed above. An informative NRBA was submitted, using external achievement data at student level as auxiliary information, along with demographic characteristics; the analysis was limited to England as the largest subnational entity within the UK, and thus covered over 90% of the intended sample. The analysis provided evidence to suggest a small residual upwards bias, after non-response adjustments are considered, driven entirely by student non-response while school non-participation did not result in significant bias. On the PISA scale, considering that the standard deviation in Scotland (in 2018) was about 95 score points in reading and mathematics, this could translate in a bias of approximately 9 or 10 points.

On the other hand, in Scotland, student response rates missed the standard by one percentage point but were otherwise similar to response rates in PISA 2018 (81%). A thorough non-response bias analysis was submitted, using several external achievement variables at student level as auxiliary information, along with demographic characteristics. The analysis provided evidence to suggest a residual upwards bias of about 0.1 standard deviations, after non-response adjustments are taken into account. On the PISA scale, considering that the standard deviation in Scotland (in 2018) was about 95 score points in reading and mathematics, this could translate in an estimated bias of approximately 9 or 10 points. Given the similarity

of response rates between 2018 and 2022, it cannot be excluded that a similar bias might be present in 2018 as well, and in many PISA 2022 participants whose response rates were similarly close to the target. For this reason, data were deemed to be comparable to previous cycles.

Decreases in student participation were more severe for other jurisdictions. In the challenging circumstances surrounding schooling in Panama in 2022 (teacher strikes, road blockades, and student absenteeism), student response rates decreased from 90% in PISA 2018 to 77% in PISA 2022. However, no NRBA was submitted; the PISA national centre explained that non-response was potentially related to the agitated school climate the students found themselves when returning to their schools after the strikes. A limited NRBA was prepared by the international sampling contractor, to compare respondent characteristics (both before and after nonresponse adjustment) to characteristics of the full eligible sample of students. This analysis suggested that (before non-response adjustments were taken into account), non-response was related to students' grade level, and to special needs status. Based on the available information, it is not possible to exclude the possibility of bias; considering the analyses on student non-response conducted in other countries, the residual bias after non-response adjustments are taken into account is likely to correspond to an upward bias.

In Canada, response rates decreased to 77% in PISA 2022 from 84% observed in PISA 2018. A thorough NRBA was submitted, with analyses conducted separately for each Canadian province, using students' academic achievement data as auxiliary information. School response rates also fell short of the target, driven by low participation rates in two provinces (Québec and Alberta). For these provinces, non-response bias was also examined at the school level. The analyses clearly indicate that school nonresponse has not led to any appreciable bias, but student nonresponse has given rise to a small upwards bias.

A similar decrease in participation was observed in Ireland, where student response rates decreased to 77% in PISA 2022 from 86% with respect to PISA 2018. A thorough NRBA was submitted, using external achievement data at student level as auxiliary information. The analysis provided evidence to suggest a residual upwards bias of about 0.1 standard deviations, after non-response adjustments are taken into account. On the PISA scale, considering that the standard deviation in Ireland ranged (in 2018) from 78 score points in mathematics to 91 score points in reading, this could translate in an estimated bias of approximately 8 or 9 points. The Adjudication Group also noted that the bias associated with trend and cross-country comparisons might be smaller, if past data or data for other countries are biased in the same direction.

Australia also observed a decline in student response rates, from 85% in PISA 2018 to 76% in PISA 2022. A technically sound NRBA was submitted; however, the strength of the evidence was limited by the fact that no external student-level achievement variables could be used in the analysis. Based on the available evidence, and on the experience of other countries participating in PISA, the Adjudication Group considered that while non-response adjustments likely limited the severity of non-response biases, a small residual upward bias could not be excluded.

Hong Kong (China) had a similar decrease from 85% in PISA 2018 to 75% in PISA 2022. School response rates also fell short of the standard (as they did in 2018). At the school level, the fact that a raw, but direct measure of school performance is used to assign schools to sampling strata (and therefore, differential non-response across strata is unlikely to cause bias), limits the risk of bias due to non-response. A NRBA was submitted; however, the strength of the evidence was limited by the fact that no external student-level achievement variables could be used in the analysis (only student grade information, already used in non-response adjustments, was available). The proxies for school and student achievement (school size and student grade) that were used in the analyses showed no or very limited relationship with participation rates. Nevertheless, based on the available evidence, and on the experience of other countries participating in PISA, the Adjudication Group considered that while non-response adjustments likely limited the severity of non-response biases, a small residual upward bias could not be excluded.

New Zealand also experienced a decline in student participation and did not meet this standard in PISA 2022. Indeed, student response rates decreased from 83% in PISA 2018 to 72% in PISA 2022. School response rates also fell short of the target as shown above. A thorough and detailed NRBA was submitted, using external achievement data at student level, but also information on chronic absenteeism along with demographic characteristics to further support comparisons. The analysis provided evidence to suggest a residual upwards bias of about 0.1 standard deviations, after non-response adjustments are considered, driven entirely by student non-response with no discernible bias due to school non-response. The analysis also suggested that chronically absent students are over-represented among non-respondents in PISA. On the PISA scale, considering that the standard deviation in New Zealand ranged (in 2018) from 93 score points in mathematics to 106 score points in reading, this could translate in an estimated bias of approximately 10 points. The Adjudication Group also noted that the bias associated with trend and cross-country comparisons might be smaller, if past data or data for other countries are biased in the same direction.

A new participant in PISA 2022, Jamaica also observed student participant rates well below the standard, at 66%. A simple NRBA was submitted by the National Centre, analysing student response rates by school characteristics: this showed in particular lower response rates in schools located in rural areas. A limited NRBA was also prepared by the international sampling contractor, to compare respondent characteristics (both before and after nonresponse adjustment) to characteristics of the full eligible sample of students. This suggested that non-response was also related to students' grade level and gender (both variables are used in non-response adjustments). Based on the available information, it is not possible to exclude the possibility of bias; considering the analyses on student non-response conducted in other countries, the residual bias after non-response adjustments are taken into account is likely to correspond to an upward bias.

The Adjudication Group noted that a number of issues encountered during the Main Survey data collection could have been prevented, had Jamaica been able to do a full Field Trial, which was not possible due to COVID-related disruptions to schooling in 2021. In particular, enrolment information available to the national centre for school-level sampling often turned out to be imprecise; and low student participation rates could have been anticipated, had a regular Field Trial been conducted. As a result of inaccurate sampling frames and low student response rates, the achieved sample size for the Main Survey was well below target, and sampling errors for Jamaica are larger than desired. In spite of the violations of sampling standards, the Adjudication Group considered the data of sufficient quality for reporting if reports are annotated with appropriate notes of caution.

Invariance of item parameters

Albeit not a formal PISA Technical Standard for the 2022 cycle, the share of non-invariant items (i.e., presenting differential item functioning) was considered for adjudication, as measurement invariance underpins the international comparability of results, and thus, one of the main goals of PISA itself. During its 2021 December meeting, the PISA Technical Advisory Group (TAG) provided guidance on this matter¹ and set the share of two thirds of invariant items for each significant language group within a jurisdiction as a threshold for adjudication on whether there is sufficient alignment with the international PISA scale for results to be deemed comparable. Viet Nam did not meet this psychometric threshold.

In Viet Nam, mathematics and science scores were considered fit for reporting, given that for the vast majority of items, student responses were in line with the expectations derived from the experience of other countries participating in PISA. In reading however, a strong linkage to the international PISA scale could not be established as 40% of items in reading (35 of 87) were assigned unique (group-specific) parameters. The Adjudication Group noted that this lack of fit in reading might also reflect differences in construct coverage of the PISA paper-based instrument used in Viet Nam and noted that this instrument will no longer be available in PISA 2025 and further cycles. Furthermore, in addition to item invariance, the

Adjudication Group also noted that the response patterns in all subjects deviated significantly from those observed in Viet Nam in earlier cycles; for this reason, the Adjudication Group recommended breaking the trend for Viet Nam, and avoiding comparisons of scale scores to those reported in past cycles.

Chromebook pilot administration issues

In PISA 2022, Iceland, and Norway participated in a pilot administration of online data collection using Chromebooks. Schools in both countries, especially those tested at the first or last days of testing windows in both countries, were affected by server outages during testing. This outage caused slowness and unresponsiveness for some students taking the cognitive assessment, thus resulting in inferior test conditions.

While the PISA Consortium solved this problem during the testing period, 579 students in Iceland (17.2% of the final student sample, unweighted) and 584 students in Norway (8.8%) were assessed on Chromebooks before the problem was solved. According to Iceland, test administrators reported the issue having affected at most 13% of the unweighted final sample (438 students). Data analyses for both countries indicated noticeable differences in the overall response time and overall response rate between students that took the test on days affected by technical problems and those tested on other days, but no noticeable differences were observed in the fit statistics, proficiency estimates or performance between the two testing sessions (first and second hour). In December 2022, the TAG reviewed these results and supported the proposal to keep the entire data for both countries, including that of days when Chromebook issues were reported, also considering that in PISA students are not penalised for having non-reached items at the end of each session. The Adjudication Group, in June 2023, confirmed that overall, the data, including those of students who took the test in these circumstances, were considered to be fit for reporting. The group noted that while it is not possible to exclude that the issue affected students' engagement and motivation to give their best effort, and therefore may have resulted in a small negative impact, their responses did show good fit with the model, and were not remarkably different from the performance of students in other schools.

Notes

1. Namely, the summary record of the PISA TAG meeting reads as follows: *For each major language of assessment within a participating country/economy, over two-thirds of items per domain are expected to be invariant from the international item parameters for the Field Test and the Main Survey. Cases with less than two-thirds of common items will undergo further review of their items and response data.*

Annex 16.A. Tables

Table 16.1. PISA 2022 Technical Standards considered in data adjudication

Area	Standard
Target population and sampling	Standard 1.2: Unless otherwise agreed upon only PISA-Eligible students participate in the test.
	Standard 1.3: Unless otherwise agreed upon, the testing period: <ul style="list-style-type: none"> • is no longer than eight consecutive weeks in duration for computer-based testing participants, • is no longer than six consecutive weeks in duration for paper-based testing participants, • does not coincide with the first six weeks of the academic year, and • begins exactly three years from the beginning of the testing period in the previous PISA cycle NOTE: TAG approved deviations to the testing period when necessary due to covid-19. a) Extension of assessment period beyond 8 weeks, students still in PISA-eligible age range b) Extension of assessment period does not exceed 8 weeks but students assessed may be outside of PISA-eligible age range by less than one week c) OECD and contractors pre-approved extension beyond 8 weeks that resulted in students outside the PISA-eligible age range being assessed.
	Standard 1.7: The PISA Defined Target Population covers 95% or more of the PISA Desired Target Population. That is, school-level exclusions and within-school exclusions combined do not exceed 5%.
	Standard 1.8: The student sample size for the computer-based mode is a minimum of 6300 assessed students, and 2100 for additional adjudicated entities, or the entire PISA Defined Target Population where the PISA Defined Target Population is below 6300 and 2100 respectively. The student sample size of assessed students for the paper-based mode is a minimum of 5250.
	Standard 1.9: The school sample size needs to result in a minimum of 150 participating schools, and 50 participating schools for additional adjudicated entities, or all schools that have students in the PISA Defined Target Population where the number of schools with students in the PISA Defined Target Population is below 150 and 50 respectively. Countries not having at least 150 schools, but which have more students than the required minimum student sample size, can be permitted, if agreed upon, to take a smaller sample of schools while still ensuring enough sampled PISA students overall.
	Standard 1.10: The minimum acceptable sample size in each school is 25 students per school (all students in the case of school with fewer than 25 eligible students enrolled).
	Standard 1.11: The final weighted school response rate is at least 85% of sampled eligible and non-excluded schools. If a response rate is below 85% then an acceptable response rate can still be achieved through agreed upon use of replacement schools.
	Standard 1.12: The final weighted student response rate is at least 80% of all sampled students across responding schools.
	Standard 1.13: The final weighted teacher response rate is at least 75% of all sampled teachers across responding schools.
	Standard 1.14: The final weighted sampling unit response rate for any optional cognitive assessment is at least 80% of all sampled students across responding schools.
	Standard 1.16: Unless otherwise agreed upon, the international contractors will draw the school sample for the Main Survey.
	Standard 1.17: Unless otherwise agreed upon, the National Centre will use the sampling contractor's software to draw the student sample, using the list of eligible students provided for each school.
Language of Testing	Standard 2.1: The PISA test is administered to a student in a language of instruction provided by the sampled school to that sampled student in the major domain (Mathematics) of the test.
	If the language of instruction in the major domain is not well defined across the set of sampled students then, if agreed upon, a choice of language can be provided, with the decision being made at the student, school, or National Centre level. Agreement with the international contractor will be subject to the principle that the language options provided should be languages that are common in the community and are common languages of instruction in schools in that adjudicated entity.
	If the language of instruction differs across domains then, if agreed upon, students may be tested using assessment instruments in more than one language on the condition that the test language of each domain matches the language of instruction for that domain. Information obtained from the Field Trial will be used to gauge the suitability of using assessment instruments with more than one language in the Main Survey.
Field Trial Participation	Standard 3.1: PISA participants participating in the PISA2021 Main Survey will have successfully implemented the Field Trial. Unless otherwise agreed upon:

Area	Standard
	<p>A Field Trial should occur in an assessment language if that language group represents more than 5% of the target population. For the largest language group among the target population, the Field Trial student sample should be a minimum of 200 students per item.</p> <p>For all other assessment languages that apply to at least 5% of the target population, the Field Trial student sample should be a minimum of 100 students per item.</p> <p>For additional adjudicated entities, where the assessment language applies to at least 5% of the target population in the entity, the Field Trial student sample should be a minimum of 100 students per item.</p>
Adaptation of tests, questionnaires and manuals	<p>Standard 4.1: The majority of test items used in previous cycles will be administered unchanged from their previous administration, unless amendments have been made to source versions, or outright errors have been identified in the national versions.</p> <p>Standard 4.2: All assessment instruments are equivalent to the source versions. Agreed upon adaptations to the local context are made if needed.</p> <p>Standard 4.3: National versions of questionnaire items used in previous cycles will be administered unchanged from their previous administration, unless amendments have been made to source versions, outright errors have been identified in the national versions, or a change in the national context calls for an adjustment.</p> <p>Standard 4.4: The questionnaire instruments are equivalent to the source versions. Agreed upon adaptations to the local context are made if as needed.</p> <p>Standard 4.5: School-level materials are equivalent to the source versions. Agreed upon adaptations to the local context are made as needed.</p>
Translation of assessment instruments, questionnaires, and manuals	<p>Standard 5.1: The following documents are translated into the assessment language in order to be linguistically equivalent to the international source versions.</p> <ul style="list-style-type: none"> • All administered assessment instruments • All administered questionnaires • The Test Administrator script from the Test Administrator (or School Associate) Manual • The Coding Guides <p>Standard 5.2: Unless otherwise agreed upon, school-level materials are translated/adapted into the assessment language to make them functionally equivalent to the international source versions.</p>
Testing of national software versions	<p>Standard 6.1: The international contractors must test all national software versions prior to their release to ensure that they were assembled correctly and have no technical problems.</p> <p>Standard 6.2: Once released, countries must test the national software versions following testing plans to ensure the correct implementation of national adaptations and extensions, display of national languages, and proper functioning on computers typically found in schools in each country. Testing results must be submitted to the international contractors so that any errors can be promptly resolved.</p>
Test administration	<p>Standard 8.1: All test sessions follow international procedures as specified in the PISA operations manuals, particularly the procedures that relate to:</p> <ul style="list-style-type: none"> • test session timing, • maintaining test conditions, • responding to students' questions, • student tracking, and • assigning assessment materials. <p>Standard 8.2: The relationship between Test Administrators and participating students must not compromise the credibility of the test session. In particular, the Test Administrator should not be the reading, mathematics, or science instructor of any student in the assessment sessions he or she will administer for PISA.</p> <p>Standard 8.3: National Centres must not offer rewards or incentives that are related to student achievement in the PISA test to students, teachers, or schools.</p>
Training Support	<p>Standard 9.1: Qualified contractor staff will conduct trainer training sessions with NPMs or designees on PISA materials and procedures to prepare them to train PISA test administrators.</p> <p>Standard 9.2: NPMs or designees shall use the comprehensive training package developed by the contractors and provided on the PISA Portal to train PISA test administrators.</p> <p>Standard 9.3: All test administrator training sessions should be scripted to ensure consistency of presentations across training sessions and across countries. Failure to do so could cause errors in data collection and make results less comparable.</p> <p>Standard 9.4: In-person and/or web based test administrator trainings should be conducted by the NPMs or designees, unless a suitable alternative is agreed upon.</p> <p>Standard 9.5: PQMs need to successfully complete self-training materials, attend webinars to review and enhance the self-training, and attend the test administrator training, unless otherwise agreed upon.</p>
National Options	<p>Standard 10.1: Only national options that are agreed upon between the National Centre and the international contractors are implemented.</p> <p>Standard 10.2: Any national option instruments that are not part of the core components of PISA are administered after all the test and questionnaire instruments of the core component of PISA have been administered to students that are part of the international PISA sample.</p>

Area	Standard
Security of the material	<p>Standard 11.1: PISA materials designated as secure are kept confidential at all times. Secure materials include all test materials, data, and draft materials. In particular:</p> <ul style="list-style-type: none"> • no-one other than approved project staff and participating students during the test session is able to access and view the test materials, • no-one other than approved project staff will have access to secure PISA data and embargoed material, and • formal confidentiality arrangements will be in place for all approved project staff.
	<p>Standard 11.2: Participating schools, students and/or teachers should only receive general information about the test prior to the test session, rather than formal content-specific training. In particular, it is inappropriate to offer formal training sessions to participating students, in order to cover skills or knowledge from PISA test items, with the intention to raise PISA scores.</p>
Quality Monitoring	<p>Standard 12.1: PISA Main Survey test administration is monitored using site visits by trained independent quality monitors.</p>
	<p>Standard 12.2: Fifteen site visits to observe test administration sessions are conducted in each PISA participating country/economy, and five site visits in each adjudicated region.</p>
	<p>Standard 12.3: Test administration sessions that are the subject of a site visit are selected by the international contractors to be representative of a variety of schools in a country/economy.</p>
Printing of Materials	<p>Standard 13.1: All paper-based student assessment material will be centrally assembled by the international contractors and must be printed using the final printready file and agreed upon paper and print quality. New countries/entities must submit a printed copy of all Field Trial instruments (booklets and questionnaires) for approval of the printing quality for the Main Survey. The same printing standard must be used for both the Field Trial and the Main Survey.</p>
	<p>Standard 13.2: The cover page of all national PISA test paper-based materials used for students and schools must contain all titles and approved logos in a standard format provided in the international version.</p>
	<p>Standard 13.3: The layout and pagination of all test paper-based material is the same as in the source versions, unless otherwise agreed upon.</p>
	<p>Standard 13.4: The layout and formatting of the paper-based questionnaire material is equivalent to the source versions, with the exception of changes made necessary by national adaptations.</p>
Response Coding	<p>Standard 14.1: The coding scheme described in the coding guides is implemented according to instructions from the international contractors' item developers.</p>
	<p>Standard 14.3: Both the single and multiple coding procedures must be implemented as specified in the PISA operations manuals (see Note 14.1). These procedures are implemented in all software that countries will be required to use.</p>
	<p>Standard 14.4: Coders are recruited and trained following agreed procedures.</p>
	<p>Standard 14.2: Representatives from each National Centre attend the international PISA coder training session for both the Field Trial and the Main Survey.</p>
Data Submission	<p>Standard 15.1: Each PISA participant submits its data in a single complete database, unless otherwise agreed upon.</p>
	<p>Standard 15.2: All data collected for PISA will be imported into a national database using the Data Management Expert (DME) data integration software provided by the international contractors following specifications in the corresponding operational manuals and international/national record layouts (codebooks). Data are submitted in the DME format.</p>
	<p>Standard 15.3: Data for all instruments are submitted. This includes the assessment data, questionnaires data, and tracking data as described in the PISA operations manuals.</p>
	<p>Standard 15.4: Unless agreed upon, all data are submitted without recoding any of the original response variables.</p>
	<p>Standard 15.5: Each PISA participating country's database is submitted with full documentation as specified in the PISA operations manuals.</p>
Communication with the International Contractors	<p>Standard 16.2: The National Centre ensures that qualified staff are available to respond to requests by the international contractors during all stages of the project. The qualified staff:</p>
	<ul style="list-style-type: none"> • Are authorized to respond to queries, • Are able to communicate in English, • Acknowledge receipt of queries within one working day, • Respond to queries from international contractors within five working days, or, if processing the query takes longer, give an indication of the amount of time required to respond to the query.
Schedule for submission of materials	<p>Standard 18.1: An agreed upon Translation Plan will be negotiated between each National Centre and the international contractors.</p>
	<p>Standard 18.2: The following items are submitted to the international contractors in accordance with agreed timelines:</p> <ul style="list-style-type: none"> • the Translation Plan • a print sample of booklets prior to final printing, for new countries/entities using the paper-based instruments (where this is required, see Standard 13.1), • results from the national checking of adapted computer-based assessment materials and questionnaires, • adaptations to school-level materials, • sampling forms (see Standard 1), • demographic tables, • completed Field Trial and Main Survey Review Forms, and • documents related to PISA Quality Monitors: nomination information, Test Administrator training schedules, translated school-

Area	Standard
	<p>level materials, school contact information, test dates, and</p> <ul style="list-style-type: none"> • other documents as specified by the PISA operations manuals <p>Standard 18.3: Questionnaire materials are submitted for linguistic verification only after all adaptations have been agreed upon.</p> <p>Standard 18.4: All adaptations to those elements of the school-level materials that are required to be functionally equivalent to the source as specified in Standard 5.2, need to be agreed upon.</p>
Management of data	<p>Standard 19.1: The timeline for submission of national databases to the international contractors is within eight weeks of the last day of testing for the Field Trial and within eight weeks of the last day of testing for the Main Survey, unless otherwise agreed upon.</p> <p>Standard 19.2: National Centres execute data checking procedures as specified in the PISA operations manuals before submitting the database.</p> <p>Standard 19.3: National Centres make a data manager available upon submission of the database. The data manager:</p> <ul style="list-style-type: none"> • is authorized to respond to international contractor data queries, • is available for a three-month period immediately after the database is submitted unless otherwise agreed upon, • is able to communicate in English, • is able to respond to international contractor queries within three working days, and • is able to resolve data discrepancies. <p>Standard 19.5: To enable the PISA participant to submit a single dataset, all instruments for all additional adjudicated entities will contain the same variables as the primary adjudicated entity of the PISA participant.</p>
Archiving of Materials	<p>Standard 19.4: A complete set of PISA paper-based instruments as administered and including any national options, is forwarded to the international contractors on or before the first day of testing. The submission includes the following:</p> <ul style="list-style-type: none"> • hard copies of instruments, • electronic PDF copies of instruments <p>Standard 20.2: The National Project Manager must submit one copy of each of the following translated and adapted Main Survey materials to the international contractors:</p> <ul style="list-style-type: none"> • electronic versions (Word and/or PDF) of all administered Test Instruments, including international and national options • electronic versions (Word and/or PDF) of all administered Questionnaires, including international and national options (paper-based countries only); • electronic versions of the school-level materials; and • electronic versions of the Coding Guides.
Data Suppression for Privacy Rights	<p>Standard 21.4: Each National Centre must facilitate requests from participants to exercise their data rights.</p> <ul style="list-style-type: none"> • Data access requests will be possible using the raw data from the assessment. No scaled data will be provided in breach of the PISA data embargo. • Data erasure requests will be possible for a limited period before submission to the Contractors. This is to be decided by each National Centre, with two options, up to the submission of ST12 or to upload of student data files to the OECS. • Each National Centre will retain and update a log of completed data requests for data erasure, to facilitate quality control processes. This information must be submitted to the PISA contractors in a timely manner to comply with the requests and for the purpose of data management and sampling processes.
Meeting Attendance	<p>Standard 23.1: Representatives from each National Centre are required to attend all PISA international meetings including National Project Manager meetings, coder training, and any separate within-school sampling training, and data management training, as necessary. Up to 6 international meetings are planned per cycle.</p> <p>Standard 23.2: Representatives from each National Centre who attend international meetings must be able to work and communicate in English.</p>
Invariant Item Parameters	<p><i>TAG Memo Dec 2021:</i> For each major language of assessment within a participating country/economy, over two-thirds of items per domain are expected to be invariant from the international item parameters for the Field Test and the Main Survey.</p>

Note: Albeit not a Technical Standard *per se*, a significant proportion of common items are essential for the linking of PISA national versions among jurisdictions and thus central for the comparability of assessment results. The PISA Technical Advisory Group (TAG) has fixed a threshold for invariant item parameters that will be incorporated into PISA technical documents in future cycles.

Source: PISA 2022 Technical Standards (Annex I)

This work is published under the responsibility of the Secretary-General of the OECD. The opinions expressed and arguments employed herein do not necessarily reflect the official views of the Member countries of the OECD.

Note by the Republic of Türkiye

The information in this document with reference to “Cyprus” relates to the southern part of the Island. There is no single authority representing both Turkish and Greek Cypriot people on the Island. Türkiye recognises the Turkish Republic of Northern Cyprus (TRNC). Until a lasting and equitable solution is found within the context of the United Nations, Türkiye shall preserve its position concerning the “Cyprus issue”.

Note by all the European Union Member States of the OECD and the European Union

The Republic of Cyprus is recognised by all members of the United Nations with the exception of Türkiye. The information in this document relates to the area under the effective control of the Government of the Republic of Cyprus.

The statistical data for Israel are supplied by and under the responsibility of the relevant Israeli authorities. The use of such data by the OECD is without prejudice to the status of the Golan Heights, East Jerusalem and Israeli settlements in the West Bank under the terms of international law.

The use of this work, whether digital or print, is governed by the Terms and Conditions to be found at: <https://www.oecd.org/termsandconditions>

