

Annex A9. Adaptive testing in PISA 2022

To improve the accuracy of measurements of student ability at the ends of the score distribution (i.e. high- or low-performing groups of students), PISA introduced adaptive testing in its reading assessment in 2018, and expanded its use to mathematics in PISA 2022. Instead of using fixed, predetermined test clusters, and rotating them at random, as was done through PISA 2015, the test items given to each student in an adaptive test are dynamically determined, based on how the student performed in prior stages of the test. Adaptive testing allows for a more accurate measurement of student performance by asking students questions that are better suited to their ability (Yamamoto, Shin and Khorramdel, 2018^[1]).

PISA 2022 implemented adaptive testing in mathematics and reading. In mathematics, a new hybrid adaptive testing design was developed and used. In reading, a reduced version of the PISA 2018 adaptive test was used.

Adaptive testing was used in every participating country/economy that took PISA 2022 using computer-based administration (CBA) as the primary mode of delivery of the test. A non-adaptive version of the test was used in countries/economies that took PISA 2022 in paper-based administration (PBA) mode.

A summary of adaptive testing in PISA 2022 is provided in this Annex. For a more detailed description of the adaptive testing design, and a discussion of the considerations that guided its development, see the *PISA 2022 Technical Report* (OECD, Forthcoming^[2]).

Adaptive Testing Design for Mathematics in PISA 2022

A hybrid multistage adaptive testing design (MSAT) was used for mathematics in PISA 2022. The design was “hybrid” because it combined an adaptive testing design with non-adaptive random-rotation design (in the latter, item assignment is not conditional on prior performance).

The MSAT design for Mathematics partitioned the item pool of 234 items (99 units) into three mutually exclusive item sets (each with 78 items). For each of the item sets, Stage 1 “core” testlets of medium difficulty, Stage 2 high- or low-difficulty testlets and Stage 3 high-, medium-, or low-difficulty testlets were assembled, each comprising 9 or 10 items. The sequence of the item sets was rotated in the final instruments (each constituting one “core”, one “Stage 1” and one “Stage 2” testlet), in order to constitute three sets of equivalent instruments to be assigned to three groups of randomly selected students (A, B, and C).

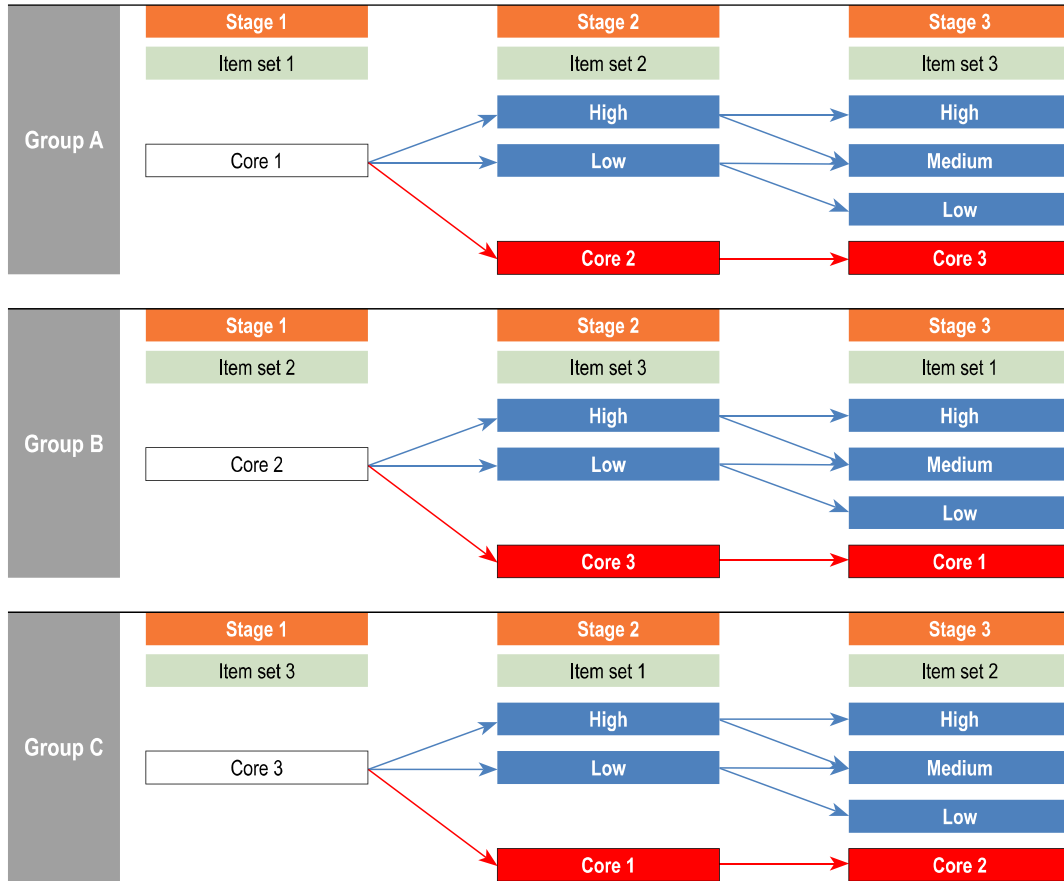
As shown in Figure I.A9.1, for the students assigned to the adaptive part of the design, Group A starts with a medium-difficulty core testlet (“Core 1” in the figure) from the first item set, followed by either a high- or low-difficulty testlet from the second item set, followed by either a high-, medium-, or low-difficulty testlet from the third item set. Similarly, group B starts with a core testlet from the second item set (“Core 2” in the figure) and group C starts with a core testlet from the third item set (“Core 3” in the figure). For the students assigned to the non-adaptive part of the design, after the core testlet their path continued in Stage 2 and Stage 3 to core testlets from the other item sets (as highlighted in red in Figure 1).

From each item set, 16 testlets of either 9 or 10 items were created within each stage. Therefore, across the three item sets and three groups, there are a total of 144 testlets (16*3*3). Each student takes one testlet in each stage; the total number of mathematics items administered to each student ranges from 28 to 30.

Simulation studies using the mathematics item pool and the Field Trial item parameters were conducted to refine the design and determine the optimum operational parameters (e.g. routing thresholds, which are the number of correct

responses on automatically scored items that determines whether students are routed to a "high", "medium" or "low" testlet in the next stage). These studies led to the decision to assign 75% of the students to the adaptive and 25% to the non-adaptive part of the hybrid design.

Figure I.A9.1. Multistage Adaptive Testing Design for PISA 2022 Mathematics



Source: PISA 2022 Technical Report (OECD, forthcoming).

Source: PISA 2022 Technical Report (OECD, Forthcoming_[2]).

Adaptive Testing Design for Reading in PISA 2022

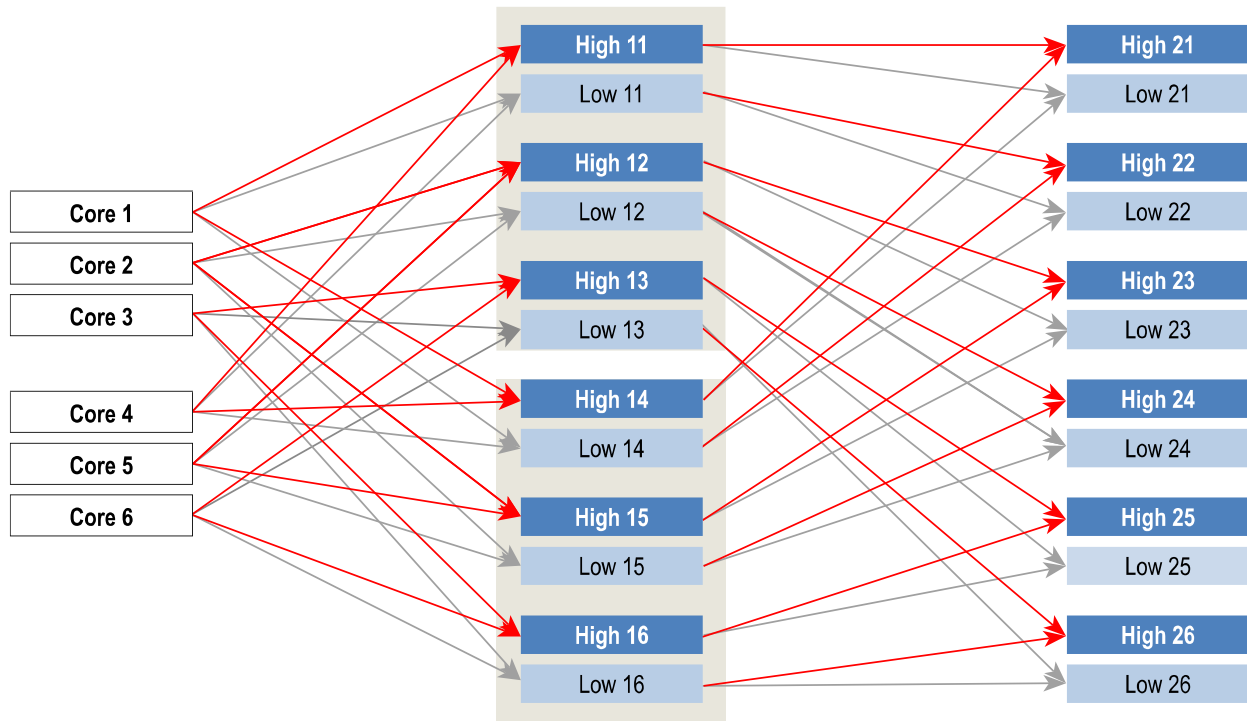
The PISA 2022 Reading MSAT design was a reduced version of the PISA 2018 Main Survey Design. It used the same adaptive structure (e.g. number of stages) as in 2018, but the 2018 Main Survey Reading item pool was reduced by approximately 25%.

In the PISA 2022 reading assessment, there were three stages: Core, Stage 1 and Stage 2. At the Core stage, six testlets were assembled. At Stage 1 and Stage 2, twelve testlets were assembled (the six more-difficult testlets were labeled as “high” and the six easier testlets were labelled as “low”).

As shown in Figure I.A9.2, at the Core stage, students were assigned to a core testlet based on a random number (between 1 and 6). At Stage 1, testlet assignment was based on three criteria: i) the Core testlet assigned, ii) the students’ performance on the Core (i.e. total number correct on automatically scored items on the given testlet), and iii) a random number and a set of rules (probability layer matrix) to overwrite the adaptive assignment for a certain proportion of students. Similarly, at Stage 2, testlet assignment was based on: i) the testlet taken at Stage 1, ii) the

performance at Core and Stage 1 (i.e. total number correct on automatically scored items on previously taken testlets), and iii) a random number and a probability layer matrix.

Figure I.A9.2. Multistage Adaptive Testing Design for PISA 2022 Reading: Standard Design



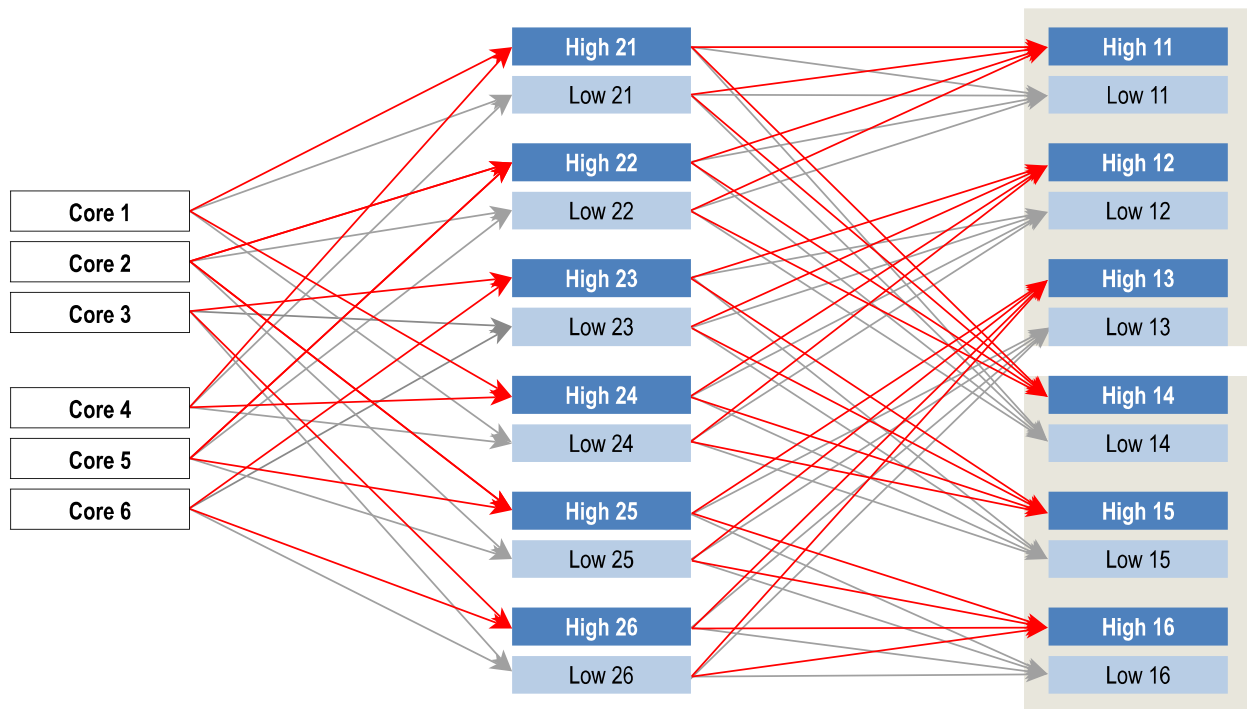
Source: *PISA 2022 Technical Report* (OECD, Forthcoming_[2]).

The routing paths shown in Figure I.A9.3 are called the Standard Design, in which students first answered a Core testlet, then a Stage 1 testlet, and finally a Stage 2 testlet. In each country, some 75% of students were expected to follow this Standard Design.

An additional set of routing paths was created and called the Alternate Design, which is shown in Figure 3. In the Alternate Design, students first answered a Core testlet, then a Stage 2 testlet, and finally a Stage 1 testlet. These additional routing paths double the number of paths from 48 in the Standard Design to 96 paths in total, with the Alternate Design.

In each country, 75% of students were expected to follow the Standard Design routing paths shown in Figure 2 (Core>Stage 1>Stage 2, with 48 paths in total) and 25% of students were expected to follow the swapped routing paths of the Alternate Design shown in Figure I.A9. (Core>Stage 2>Stage 1, with 96 paths in total).

Figure I.A9.3. Multistage Adaptive Testing Design for PISA 2022 Reading: Alternate Design

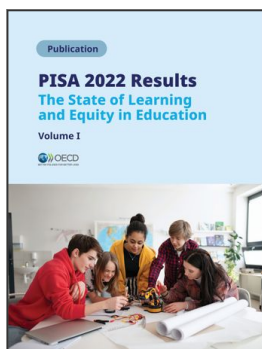


Source: *PISA 2022 Technical Report* (OECD, Forthcoming^[2]).

References

OECD (Forthcoming), *PISA 2022 Technical Report*, PISA, OECD Publishing, Paris. [2]

Yamamoto, K., H. Shin and L. Khorramdel (2018), "Multistage Adaptive Testing Design in International Large-Scale Assessments", *Educational Measurement: Issues and Practice*, Vol. 37/4, pp. 16-27, <https://doi.org/10.1111/emip.12226>. [1]



From:
PISA 2022 Results (Volume I)
The State of Learning and Equity in Education

Access the complete publication at:

<https://doi.org/10.1787/53f23881-en>

Please cite this chapter as:

OECD (2023), "Adaptive testing in PISA 2022", in *PISA 2022 Results (Volume I): The State of Learning and Equity in Education*, OECD Publishing, Paris.

DOI: <https://doi.org/10.1787/89c0f253-en>

This document, as well as any data and map included herein, are without prejudice to the status of or sovereignty over any territory, to the delimitation of international frontiers and boundaries and to the name of any territory, city or area. Extracts from publications may be subject to additional disclaimers, which are set out in the complete version of the publication, available at the link provided.

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