1

How does PISA assess reading?

Reading was the focus of the OECD Programme for International Student Assessment (PISA) in 2018. This chapter discusses how PISA defined and measured reading literacy. Differences between the PISA 2018 reading test and that of previous PISA assessments are highlighted. The chapter also explains what is meant by adaptive testing – the new way students progress through the assessment.

The OECD Programme for International Student Assessment (PISA) is a triennial survey that assesses what students know and what they can do with what they know. In addition to an innovative domain developed expressly for each new round of PISA, the survey measures students' proficiency in three foundational domains of competence – reading, mathematics and science – one of which, the so-called major domain, is the particular focus of that assessment. The major domain is rotated with each round of PISA.

The major domain in the first year PISA was conducted, 2000, was reading. Reading was the major domain again in 2009 and in 2018. However, the nature of reading has evolved significantly over the past decade, notably due to the growing influence and rapid evolution of technology. Reading now involves not only the printed page but also electronic formats (i.e. digital reading). Moreover, readers must now engage in a greater variety of tasks. In the past, when students did not know the answer to a question, they could look it up in an encyclopaedia and generally trust that the answer they found was accurate. Today, digital search engines give students millions of answers, and it will be up to them to figure out which are accurate, true and relevant and which are not. Now, more than ever before, literacy requires triangulating different sources, navigating through ambiguity, distinguishing between fact and opinion, and constructing knowledge. The ways PISA measures competency in reading, or reading literacy, have had to adapt to these changes, some of which are described in Box I.1.1.

Box I.1.1. **The changing nature of reading**

The past decade has been a period of rapid digitalisation. In 2009, the most recent year reading was the major domain of assessment in PISA, about 15% of students in OECD countries, on average, reported that they did not have access to the Internet at home. By 2018, that proportion had shrunk to less than 5% (Tables I.B1.54, I.B1.55 and I.B1.56). The growth in access to online services is likely to be even larger than suggested by these percentages, which hide the exponential growth in the quality of Internet services and the explosion of mobile Internet services over the past decade. OECD statistics indicate, for example, that between 2009 and 2018, the number of mobile broadband subscriptions per capita increased by more than a factor of three across OECD countries, on average. At the end of 2018, there were more mobile broadband subscriptions than inhabitants, on average (109.7 per 100 inhabitants) (OECD, 2019[1]).

The rapid digitalisation of communication is having a profound impact on the kind of information literacy that young adults will need to demonstrate in their future jobs and in their wider social interactions. Evolving technologies have, for example, changed the ways people read and exchange information, whether at home, at school or in the workplace.

Figure I.1.1 **Change between 2009 and 2018 in what and why students read**

Percentage of students; OECD average
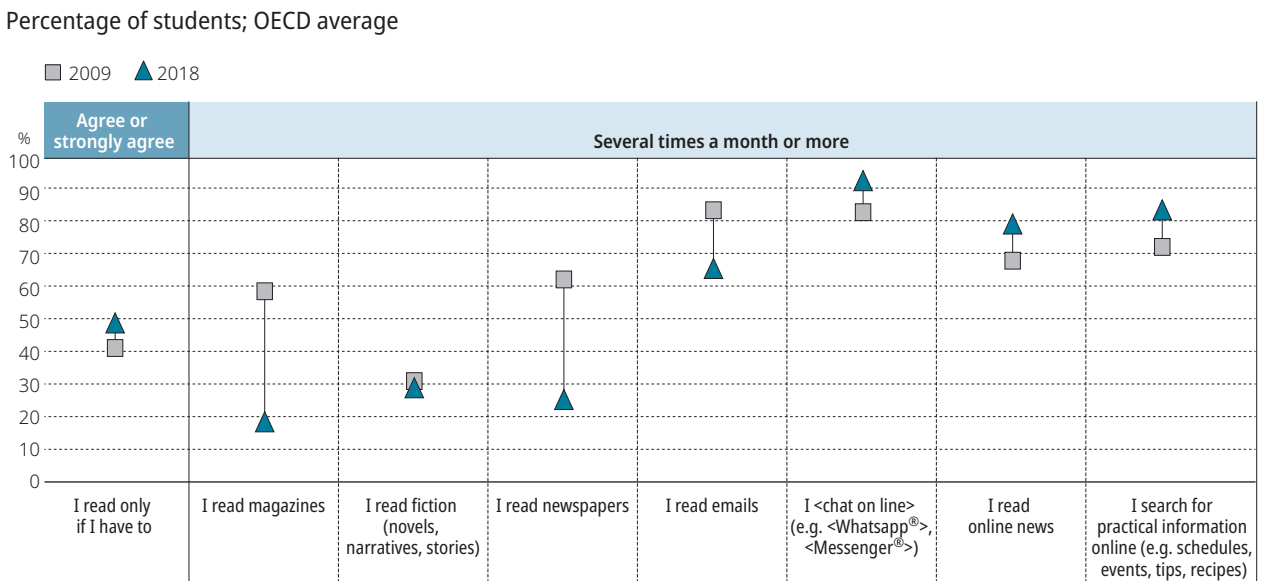
Note: All changes between PISA 2018 and PISA 2009 are statistically significant (see Annex A3).
Source: OECD, PISA 2018 Database, Tables I.B1.57, I.B1.58 and I.B1.59.
StatLink https://doi.org/10.1787/888934028159

...

Some of these changes are already apparent in what 15-year-olds do and read. In all countries and economies that distributed the optional ICT familiarity questionnaire, the amount of time that 15-year-old students spent on line outside of school increased between 2012 and 2018. The average increase across OECD countries was of more than one hour per day (on both weekdays and weekends). Students now spend about 3 hours on line outside of school on weekdays, on average, and almost 3.5 hours on line on weekend days (Tables I.B1.51, I.B1.52 and I.B1.53).[1] In particular, students in Costa Rica, Ireland, Italy and Turkey more than doubled the amount of time they spent on line on both weekdays and weekends, on average.

In parallel, students seem to read less for leisure and to read fewer fiction books, magazines or newspapers because they want to do so (as opposed to needing to do so). Instead, they read more to fulfil their practical needs, and they read more in online formats, such as chats, online news or websites containing practical information (e.g. schedules, events, tips, recipes) (Figure I.1.1). More students consider reading "a waste of time" (+5 percentage points, on average) and fewer students read for enjoyment (-5 percentage points) (Table I.B1.59).

As the medium through which people access textual information expands from print to computer screens to smartphones, the variety of the structure and formats of texts has also expanded. Reading remains a practical necessity, and perhaps more so than in the past, it requires the use of complex information-processing strategies, including the analysis, synthesis, integration and interpretation of relevant information from multiple sources. The nature of texts and the type of problems included in the PISA 2018 assessment of reading reflect the evolving nature of reading in increasingly digital societies.

The changes in the assessment of reading described in this chapter apply to the countries/economies that delivered the PISA test on computer, which comprised the vast majority of the countries/economies that participated in PISA 2018. However, nine countries – Argentina, Jordan, Lebanon, the Republic of Moldova, the Republic of North Macedonia, Romania, Saudi Arabia, Ukraine and Viet Nam – assessed their students' knowledge and skills in PISA 2018 using paper-based instruments. The paper-based test of reading was based on the PISA 2009 reading framework (see Annex A5) and only included items previously used in the PISA assessment; no new items were developed for the paper-based test. Box I.1.2 summarises the changes in the reading framework and assessment between PISA 2009 and PISA 2018.

Box I.1.2. **Changes between 2009 and 2018 in the PISA assessment of reading literacy**

This chapter describes the PISA 2018 reading literacy framework. This framework is similar in many respects to the PISA 2009 reading literacy framework, which was also used in PISA 2012 and 2015. The chapter also discusses some changes in how the reading assessment was implemented. The major differences between the 2009 and 2018 assessments are:

- a greater emphasis on multiple-source texts, i.e. texts composed of several units of text, created separately by different authors (Rouet, Britt and Potocki, 2019[2]). These types of text are more prevalent in the information-rich digital world, and the digital delivery of the PISA reading assessment made it possible to present them to students. While the availability of multiple sources does not necessarily imply greater difficulty, including multiple-source units helped to expand the range of higher-level reading processes and strategies measured by PISA. In 2018, these included searching for information across multiple documents, integrating across texts to generate inferences, assessing the quality and credibility of sources, and handling conflicts across sources (List and Alexander, 2018[3]; Barzilai, Zohar and Mor-Hagani, 2018[4]; Stadtler and Bromme, 2014[5]; Magliano et al., 2017[6])

- the explicit assessment of reading fluency, defined as the ease and efficiency with which students can read text

- the use of adaptive testing, whereby the electronic test form that a student saw depended on his or her answers to earlier questions

- the digital, on-screen delivery of text, which facilitated the first and third changes listed above. The 2009 assessment was conducted on paper while the 2018 assessment was conducted (by default) on computer.[2, 3] Students had to use navigational tools to move between passages of text, as there was often too much text to fit onto one screen.

These changes are all described in this chapter. An analysis of whether and how such changes might have affected results is provided in Chapter 9, which analyses changes in performance between 2015 and 2018. While a few countries/economies may have been affected more than others, the analysis in Box I.8.1 in Chapter 8 shows that effects on country mean scores were not widespread.

HOW DOES PISA DEFINE READING LITERACY?

PISA assesses reading literacy, as opposed to reading. Reading is often interpreted, in a general, non-academic context, as reading aloud or simply converting text into sounds. PISA conceives of reading literacy as a broader set of competencies that allows readers to engage with written information, presented in one or more texts, for a specific purpose (RAND Reading Study Group and Snow, 2002[7]; Perfetti, Landi and Oakhill, 2005[8]).

To engage with what they read, readers must understand the text and integrate this with their pre-existing knowledge. They must examine the author's (or authors') point of view and decide whether the text is reliable and truthful, and whether it is relevant to their goals or purpose (Bråten, Strømsø and Britt, 2009[9]).

PISA also recognises that reading is a daily activity for most people, and that education systems need to prepare students to be able to adapt to the variety of scenarios in which they will need to read as adults. These scenarios range from their own personal goals and development initiatives, to their experiences in further and continuing education, and to their interactions at work, with public entities, in online communities and with society at large. It is not enough to be a proficient reader; students should also be motivated to read and be able to read for a variety of purposes (Britt, Rouet and Durik, 2017[10]; van den Broek et al., 2011[11]).

All of these considerations are reflected in the PISA 2018 definition of reading literacy:

> *Reading literacy is understanding, using, evaluating, reflecting on and engaging with texts in order to achieve one's goals, to develop one's knowledge and potential, and to participate in society.*

THE PISA 2018 FRAMEWORK FOR ASSESSING READING LITERACY

The PISA 2018 framework for reading guided the development of the PISA 2018 reading literacy assessment (OECD, 2019[12]). It conceptualises reading as an activity where the reader interacts with both the text that he or she reads and with the tasks[4] that he or she wants to accomplish during or after reading the text. To be as complete as possible, the assessment covers different types of texts and tasks over a range of difficulty levels. The assessment also requires students to use a variety of processes, or different ways in which they cognitively interact with the text.

Texts

The PISA 2009 reading framework classified texts along four dimensions:

- **Medium**: Is the text delivered in print format or in electronic format?

- **Environment**: Was the text composed by an author or group of authors alone, without the participation of the reader, or was the text composed in a collaborative way with the potential contribution of the reader?

- **Text format**: Is it a piece of continuous prose, a non-continuous (usually list-like) matrix of writing, or a mixture of these two formats?[5]

- **Text type**: Why was the text written and how is it organised? Six major text types were identified:[6]
  - *descriptions* identify a tangible object and where it is located in space
  - *narrations* detail when and in what sequence events occurred
  - *expositions* explain or summarise an object or concept, and describe how objects and concepts relate to one another
  - *argumentations* try to persuade or convince the reader of the writer's viewpoint
  - *instructions* provide directions as to what to do
  - *transactions* aim to achieve a specific purpose (and are often in the form of letters or messages between two interlocutors).

In the PISA 2018 computer-based assessment of reading, all texts were presented on screens; as such, the "medium" dimension was no longer relevant for classification purposes as it no longer distinguished between texts. The four dimensions used to classify texts in the PISA 2018 reading literacy framework are:

- **Source** (related to the previous classification of "environment"): Is the text composed of a single unit (a single-source text) or of multiple units (a multiple-source text)?[7]

- **Organisational and navigational structure**: How do readers read and move through all of the text when only a certain portion can be displayed on the screen at any given time? **Static texts** have a simple, often linear organisational structure and make use of a low density of straightforward navigational tools, such as scroll bars and tabs. **Dynamic texts**, on the other hand, have a more intricate organisational structure and a higher density and complexity of navigational tools, such as a table of contents, hyperlinks to switch between segments of text, or interactive tools that allow the reader to communicate with others (as in social networks).

- **Text format** (unchanged from the previous framework): Is it a piece of continuous prose, a non-continuous (usually list-like) matrix of writing, or a mixture of these two formats?

- **Text type** (unchanged from the previous framework): Why was the text written and how is it organised?[8]

Processes

The PISA 2018 framework identifies four processes that readers activate when engaging with a piece of text. Three of these processes were also identified, in various guises, in previous PISA frameworks: "locating information", "understanding", and "evaluating and reflecting". The fourth process, "reading fluently", underpins the other three processes. The inclusion of tasks that assess reading fluency independently of other processes is new to the PISA 2018 assessment. Table I.1.1 presents a breakdown of the PISA 2018 reading literacy assessment by process assessed.

Table I.1.1. **Approximate distribution of tasks, by process and text source**

2015 Framework	2018 Framework		
		Single-source text 65%	Multiple-source text 35%
Accessing and retrieving 25%	Locating information 25%	Scanning and locating 15%	Searching for and selecting relevant text 10%
Integrating and interpreting 50%	Understanding 45%	Representing literal meaning 15% Integrating and generating inferences 15%	Integrating and generating inferences 15%
Reflecting and evaluating 25%	Evaluating and reflecting 30%	Assessing quality and credibility, and Reflecting on content and form 20%	Corroborating and handling conflict 10%

Note: Reading fluency is not included in the above table. Reading-fluency items were included at the beginning of the assessment and considered in the computation of students' overall score. However, these items were not included in the computation of subscale scores (neither the text-source subscale nor the reading-process subscale) and are not part of any of the percentages in this table.
Source: OECD (2019[12]), *PISA 2018 Assessment and Analytical Framework,* PISA, OECD Publishing, Paris, https://dx.doi.org/10.1787/b25efab8-en.

Reading fluently

PISA defines reading fluency as the ease and efficiency with which one can read and understand a piece of text. More specifically, it includes the ability to read words and text accurately and automatically, and then to parse, phrase and process them to comprehend the overall meaning of the text (Kuhn and Stahl, 2003[13]).

Reading fluency is positively correlated with reading comprehension (Annex A8). Indeed, students who can easily and efficiently read a piece of text free up cognitive resources for higher-level comprehension tasks (Cain and Oakhill, 2004[14]; Perfetti, Marron and Foltz, 1996[15]).

PISA 2018 evaluated reading fluency by presenting students with a variety of sentences, one at a time, and asking them whether they made sense. These sentences were all relatively simple, and it was unambiguous whether they made sense or not. Example sentences include:[9]

- Six birds flew over the trees.
- The window sang the song loudly.
- The man drove the car to the store.

Locating information

The first cognitive process involved in reading is "locating information" (known in previous frameworks as "accessing and retrieving"). Readers often search for a particular piece of information, without considering the rest of the text (White, Chen and Forsyth, 2010[16]). Locating information when reading digitally also demands skills different from those used when reading in print format. For example, readers need to be able to handle new text formats, such as search engine results, and websites with multiple tabs and various navigational features.

In order to locate information as quickly and efficiently as possible, readers must be able to judge the relevance, accuracy and credibility of passages. They need to be able to modulate their reading speed, skimming through sections deemed to be irrelevant until arriving at a promising passage, whereupon they read more carefully. Readers must finally make use of text organisers, such as headers, that may suggest which sections are relevant.

PISA 2018 breaks "locating information" into two specific cognitive processes, depending on the number of texts involved:

- **Scanning and locating**, where readers need to scan only a single piece of text to retrieve a few words, phrases or numerical values. There is little need to comprehend the overall text as the target information appears essentially verbatim in the text.

- **Searching for and selecting relevant text**, where readers need to deal with several pieces of text. This is particularly relevant in digital reading, where the total amount of text available far exceeds the amount that readers can or need to process. In order to locate the desired information, readers need first to identify the appropriate piece of text, which adds to the complexity of this process. Text organisers, such as headers, source information (e.g. author, medium and date of publication) and links (e.g. search engine result pages) are particularly important for this process.

A task involving multiple sources of texts is not necessarily more difficult than one involving a single source of text. In PISA 2018, care was taken to include in the assessment some easy search tasks involving multiple texts of limited length and complexity (such as short notes on a bulletin board, or lists of document titles or search engine results). In contrast, it was not possible (due to time limits, and the offline nature of the assessment) to include more complex and open-ended search scenarios that readers may encounter on the Internet. As a consequence, both types of processes can be found at all levels of difficulty. Simple scan-and-locate or search-and-select tasks involve little information, salient targets and literal matches, whereas more complex tasks involve more information, non-literal matches, targets located in non-salient positions and a high density of distractors.

Understanding

"Understanding" (known in previous frameworks as "integrating and interpreting", and commonly referred to as "reading comprehension") involves constructing a mental representation of the content of a piece of text or a set of texts (Kintsch, 1998[17]). In other words, readers must recognise the meaning conveyed in the passage. The PISA 2018 reading literacy framework identifies two specific cognitive processes involved in understanding, distinguished by the length of the text to be understood:

- **Representing literal meaning**, where readers must paraphrase sentences or short passages so that they match the target information desired by the task.

- **Integrating and generating inferences**, where readers must work with longer passages to establish their overall meaning. They may have to connect information across various passages or texts, and infer how they are connected to each other (e.g. spatially, temporally or causally) and potentially also to the statement in the question. Readers may also have to resolve conflicts between different texts. Constructing an integrated text representation is associated with tasks such as identifying the main idea of a piece of text or a set of texts, summarising a long passage or giving a title to a piece of text or set of texts. Inter-textual inferences tend to require a high level of proficiency, perhaps because they involve distinct and demanding cognitive processes (Barzilai, Zohar and Mor-Hagani, 2018[4]). This process can be engaged when reading multiple pieces of text or when reading just one, typically longer, piece of text.

Evaluating and reflecting

The highest-level process identified by the PISA 2018 reading literacy framework is "evaluating and reflecting". Here, readers must go beyond understanding the literal or inferred meaning of a piece of text or a set of texts to assess the quality and validity of its content and form.

Three specific cognitive processes are classified under evaluating and reflecting:

- **Assessing quality and credibility**, where readers judge whether the content is valid, accurate and/or unbiased. This may also involve identifying the source of the information and thereby identifying the author's intentions and judging whether the author is competent and well-informed. Assessing quality and credibility, in other words, requires the reader to combine the content of what is said in the text with peripheral cues, such as who wrote it, when, for what purpose and so forth.

- **Reflecting on content and form**, where readers evaluate the quality and the style of the text. They need to assess whether the content and form adequately express the author's purpose and point of view. In order to do so, they may need to draw from their real-world knowledge and experience in order to be able to compare different perspectives.

- **Corroborating and handling conflict**, where readers need to compare information across texts, recognise contradictions between pieces of text and then decide how best to manage such contradictions. They can do so by evaluating the credibility of the sources, and the logic and soundness of their claims (Stadtler and Bromme, 2014[5]). This cognitive process is commonly used when examining multiple-source texts.

Evaluating and reflecting has always been a part of reading literacy. However, its importance has grown in the era of digital reading as readers are now confronted with ever-growing amounts of information, and must be able to distinguish between what is trustworthy and what is not. Indeed, only the first two cognitive processes above, "assessing quality and credibility" and "reflecting on content and form", were included in previous PISA reading literacy frameworks under the overall process of "reflecting and evaluating".

Tasks

Readers engage with texts for a purpose; in PISA, the purpose is to respond to questions about these texts in order to provide evidence of their level of reading literacy. Such questions, or tasks, require students to perform at least one of the cognitive processes discussed in the previous section (see Table I.1.1 above). They are arranged in units, which are based on one piece of or several texts. Within each unit, tasks are often arranged in order of difficulty. For example, the first task in a unit could ask students to locate the most relevant piece of text; the second task could ask students to consider information that is specifically stated in the text; and the third task could ask students to compare the points of view in two different pieces of text.

PISA tasks have usually been presented in the form of discrete, unrelated units, each with its own set of texts. However, in order to better engage students, PISA 2018 also presented some tasks using scenarios, each of which had an overarching purpose and was supported by a collection of thematically related texts that may have come from a variety of sources.

As in traditional units, students responding to these scenarios must realise what is being asked of them, set out how they will achieve what is being asked of them, and monitor their progress along this path. Instead of reading a clearly assigned passage, as is done in traditional units, students responding to scenarios have a greater choice of the sources they use to respond to questions. Students are therefore required to search for relevant pieces or passages of text.

Regardless of whether an item is part of an individual unit or a broader scenario, one of a small set of response formats is applied: selected response (e.g. multiple choice, true/false, yes/no) and short constructed response (or open response).[10] Some 87 items, or about one-third of the 245 items,[11] asked students for short constructed responses, which the students usually had to type into an open text-entry field. For 82 of those 87 items, human coders scored students' responses as correct or incorrect after the assessment was completed. Automatic, real-time scoring could be used for five items, such as when the correct response consisted of a simple number.

Although writing and reading are correlated skills, and although students had to construct some short, human-coded responses, PISA is a reading assessment, not a writing assessment. As such, writing skills (spelling, grammar, organisation and quality) were not evaluated by human coders.

Illustrative examples of reading tasks, including some actually used in the PISA 2018 assessment, and a discussion of the texts and processes required to solve these tasks are provided in Chapter 5 and Annex C.

HOW DOES THE PISA ADAPTIVE TEST OF READING WORK?

Most students in OECD countries perform near the middle of the score distribution, or at around 500 points. Most of the test material in previous PISA assessments was also targeted to middle-performing students, which allowed for more refined differentiation of student ability at this level. However, this meant that there was a relative lack of test material at the higher and lower ends of student ability, and that the scores of both high- and low-performing students were determined with less accuracy than the scores of middle-performing students.

This was generally not a problem (or less of a problem) when examining country averages or when examining countries and economies that scored at around 500 points. Many PISA analyses, however, examine high- or low-performing groups of students in more detail. For example, students from advantaged families (who typically have high scores in PISA) are compared to students from disadvantaged families (who typically have low scores in PISA) when determining the impact of socio-economic status on performance. It is hence important that PISA be able to accurately gauge student ability at the ends of the distribution.

In order to improve the accuracy of such measurements, PISA 2018 introduced adaptive testing in its reading assessment. Instead of using fixed, predetermined test clusters, as was done through PISA 2015, the reading assessment given to each student was dynamically determined, based on how the student performed in prior stages of the test.

There were three stages to the PISA 2018 reading assessment: Core, Stage 1 and Stage 2.[12,13] Students first saw a short non-adaptive Core stage, which consisted of between 7 and 10 items.[14] The vast majority of these items (at least 80% and always at least 7 items) were automatically scored. Students' performance in this stage was provisionally classified as low, medium or high, depending on the number of correct answers to these automatically scored items.[15]

The various Core blocks of material delivered to students did not differ in any meaningful way in their difficulty. Stages 1 and 2, however, both existed in two different forms: comparatively easy and comparatively difficult.[16] Students who displayed medium performance in the Core stage were equally likely to be assigned an easy or a difficult Stage 1. Students who displayed low performance in the Core stage had a 90% chance of being assigned to an easy Stage 1 and a 10% chance of being assigned to a difficult Stage 1. Students who displayed high performance in the Core stage had a 90% chance of being assigned to a difficult Stage 1 and a 10% chance of being assigned to an easy Stage 1.

Students were assigned to easy and difficult Stage 2 blocks of material in much the same way. In order to classify student performance as precisely as possible, however, responses to automatically scored items from both the Core stage and Stage 1 were used.[17]

This contrasts with how the PISA reading test was conducted in previous assessments, when test material was divided into several fixed 30-minute clusters, which were then assembled into electronic test forms or paper booklets. In PISA 2015, for example, each student received a two-hour test form or booklet composed of two 30-minute clusters of test material in the major domain along with two clusters in one or two of the other domains. As they were fixed, the test form did not change over the course of the assessment, irrespective of student performance.[18]

As with many of the new features in the reading framework, adaptive testing was made possible through the use of computers. Adaptive testing could not have been used in the paper-based assessment as there would have been no way of ascertaining performance while the student was completing the test. One potential drawback of an adaptive design is that students are unable to come back to a question in a previous stage. This was already the case in the PISA 2015 computer-based assessment, where students could navigate between items in a unit but not across units. However, with adaptive testing, students' responses in the Core stage and in Stage 1 affected not only their performance but also the questions that they saw later in the assessment.[19] The *PISA 2018 Technical Report* (OECD, forthcoming[18]) and Annex A8 present further indicators of the impact of adaptive testing on students' test-taking behaviour.

....

Notes

1. These times are lower estimates of the average amount of time students spent on line. Students were asked to report the time they spent on line as, for example, "between 1 and 30 minutes", "between 31 and 60 minutes", or "between 1 and 2 hours". The average amount of time was calculated by using the lower bounds of each range, i.e. 1, 31 and 61 minutes for the three aforementioned options.

2. PISA was delivered via computer in most countries and economies in the 2015 assessment. However, all of the questions used in the PISA 2015 reading assessment were recycled from previous assessments; they were all based on either the PISA 2000 or PISA 2009 reading literacy frameworks. The 2018 reading framework was the first framework specifically developed with computer delivery in mind, and thus the first to take into account the new opportunities made possible by computer delivery.

3. A first attempt at measuring students' ability to read digital texts was conducted in PISA 2009, which contained a separate digital reading assessment in addition to the standard (paper) reading assessment. However, it was more limited in scope, with only 19 countries/economies participating in this assessment. This assessment of digital reading was repeated in 2012, with the participation of an additional 13 countries and economies (32 in total).

4. Although tasks, especially as conceived in the PISA assessment, can refer to specific goals, such as locating information or identifying the main points of an argument, a task may also be simply reading for pure enjoyment.

5. The PISA 2009 reading literacy framework also included multiple texts as a possible text format. In the PISA 2018 framework, the distinction between multiple texts (or multiple-source texts) and single texts (which can be continuous, non-continuous or mixed texts) is captured by the "source" dimension and is discussed later in this section.

6. Many texts in the real world can be classified under multiple text types. This assessment generally classifies each piece of text into one text type, based on its predominant properties, to ensure a wide coverage of text types; however, a small number of items are classified as having "multiple" text types.

7. A unit of text is characterised as having been written by a definite author or group of authors at a specific time. It often has a specific title. Long pieces of text with several sections and subtitles, and websites that span multiple pages (without any indication as to their date of creation or publication) are both considered to be single units of text. However, a newspaper with multiple articles and an online forum with multiple posts are both considered to be multiple units of text.

8. One new text type was included in 2018: *interactions* show conversations and discussions between people, often without the same sense of purpose, as in a transaction.

9. The first and third sentences make sense, while the second sentence does not.

10. Computer delivery of the assessment also allows for new digital response formats that involve interaction with the text, such as highlighting passages or dragging and dropping words and passages into place. Such formats were used only to a limited extent in PISA 2018 (see item #6 in unit *Rapa Nui*, in Annex C, for an example) but remain a possibility for future assessments.

11. There were 245 items in the test; however, one item was not considered in scaling due to a technical problem with the recording of student responses.

12. For a more detailed description of the adaptive testing design, and a discussion of the considerations that guided its development, see the *PISA 2018 Technical Report* (OECD, forthcoming[18]) and Yamamoto, Shin and Khorramdel (2018[19], 2019 [20]).

13. Reading-fluency items were given at the beginning of the PISA reading assessment; student responses to these items were not used in the determination of the Stage 1 and Stage 2 blocks that a student would see and had no effect on the adaptive testing aspect of the assessment. However, these items were used in the overall determination of students' proficiency in reading.

14. The non-adaptive Core stage was delivered as one of eight possible Core blocks of material. Each Core block was composed of two units, and each unit comprised a set of items developed around shared stimulus material.

15. To select an adequate test form (i.e. Stage 1 and Stage 2 blocks) while students sat the assessment, their performance was classified using only the automatically scored items already seen (i.e. Core stage items to decide the Stage 1 block, and Core stage and Stage 1 items to decide the Stage 2 block). However, all items, including those that required human coding, were used to evaluate overall performance and report students' proficiency in reading.

16. More specifically, both Stages 1 and 2 were delivered as one of 16 possible blocks of material, 8 of which were comparatively easy and 8 of which were comparatively difficult. Each Stage 1 block was built from 3 units that, in total, amounted to between 12 and 15 items, of which between 8 and 11 were automatically scored. Similarly, each Stage 2 block was built from 2 units that, in total, amounted to between 12 and 15 items, of which between 6 and 12 were automatically scored.

17. Some 75% of students were first presented with Stage 1 after the Core stage, after which they were presented with Stage 2. The other 25% of students were presented with a Stage 2 block immediately after the Core stage, after which they were directed to an easier or more difficult Stage 1 block, depending on their performance in the Core stage and Stage 2. Using two complementary test designs allowed for greater accuracy when calibrating the parameters that described item difficulty and discrimination. See Annex A1 and and the *PISA 2018 Technical Report* (OECD, forthcoming[18]).

18. See Annex A8 and Chapter 8 for a discussion on whether and how adaptive testing may have affected results.

19. Adaptive testing allows for a more accurate measurement of student performance by asking students questions that are better suited to their ability. This process does not bias student scores, when compared to the ideal scenario where students would answer all questions over a longer testing period.

References

Barzilai, S., A. Zohar and **S. Mor-Hagani** (2018), "Promoting Integration of Multiple Texts: a Review of Instructional Approaches and Practices", *Educational Psychology Review*, Vol. 30/3, pp. 973-999, http://dx.doi.org/10.1007/s10648-018-9436-8.   [4]

Bråten, I., H. Strømsø and **M. Britt** (2009), "Trust Matters: Examining the Role of Source Evaluation in Students' Construction of Meaning Within and Across Multiple Texts", *Reading Research Quarterly*, Vol. 44/1, pp. 6-28, http://dx.doi.org/10.1598/rrq.44.1.1.   [9]

Britt, M., J. Rouet and **A. Durik** (2017), *Literacy beyond Text Comprehension*, Routledge, http://dx.doi.org/10.4324/9781315682860.   [10]

Cain, K. and J. Oakhill (2004), "Reading Comprehension Difficulties", in *Handbook of Children's Literacy*, Springer Netherlands, Dordrecht, http://dx.doi.org/10.1007/978-94-017-1731-1_18.   [14]

Kintsch, W. (1998), *Comprehension: A Paradigm for Cognition*, Cambridge University press.   [17]

Kuhn, M. and **S. Stahl** (2003), "Fluency: A review of developmental and remedial practices.", *Journal of Educational Psychology*, Vol. 95/1, pp. 3-21, http://dx.doi.org/10.1037/0022-0663.95.1.3.   [13]

List, A. and **P. Alexander** (2018), "Toward an Integrated Framework of Multiple Text Use", *Educational Psychologist*, Vol. 54/1, pp. 20-39, http://dx.doi.org/10.1080/00461520.2018.1505514. [3]

Magliano, J. et al. (2017), "The Modern Reader", in *The Routledge Handbook of Discourse Processes*, Routledge, http://dx.doi.org/10.4324/9781315687384-18. [6]

OECD (2019), *Broadband Portal*, https://www.oecd.org/sti/broadband/broadband-statistics/ (accessed on 10 July 2019). [1]

OECD (2019), *PISA 2018 Assessment and Analytical Framework*, PISA, OECD Publishing, Paris, https://dx.doi.org/10.1787/b25efab8-en. [12]

OECD (forthcoming), *PISA 2018 Technical Report*, OECD Publishing, Paris. [18]

Perfetti, C., N. Landi and **J. Oakhill** (2005), "The Acquisition of Reading Comprehension Skill", in Snowling, M. and C. Hulme (eds.), *The Science of Reading: A Handbook*, Blackwell Publishing Ltd, Oxford, UK, http://dx.doi.org/10.1002/9780470757642.ch13. [8]

Perfetti, C., M. Marron and **P. Foltz** (1996), "Sources of Comprehension Failure: Theoretical Perspectives and Case Studies", in Cornoldi, C. and J. Oakhill (eds.), *Reading Comprehension Difficulties: Processes and Remediation*, Erlbaum. [15]

RAND Reading Study Group and **C. Snow** (2002), *Reading for Understanding: Toward an R&D Program in Reading Comprehension*, RAND Corporation, Santa Monica, CA; Arlington, VA; Pittsburgh, PA, https://www.jstor.org/stable/10.7249/mr1465oeri (accessed on 30 August 2019). [7]

Rouet, J., M. Britt and **A. Potocki** (2019), "Multiple-Text Comprehension", in *The Cambridge Handbook of Cognition and Education*, Cambridge University Press, http://dx.doi.org/10.1017/9781108235631.015. [2]

Stadtler, M. and **R. Bromme** (2014), "The content–source integration model: A taxonomic description of how readers comprehend conflicting scientific information", in Rapp, D. and J. Braasch (eds.), *Processing Inaccurate Information: Theoretical and Applied Perspectives from Cognitive Science and the Educational Sciences*, MIT Press. [5]

van den Broek, P. et al. (2011), "When a reader meets a text: The role of standards of coherence in reading comprehension", in McCrudden, M., J. Magliano and G. Schraw (eds.), *Text relevance and learning from text*, Information Age Publishing. [11]

White, S., J. Chen and **B. Forsyth** (2010), "Reading-Related Literacy Activities of American Adults: Time Spent, Task Types, and Cognitive Skills Used", *Journal of Literacy Research*, Vol. 42/3, pp. 276-307, http://dx.doi.org/10.1080/1086296x.2010.503552. [16]

Yamamoto, K., H. Shin and **L. Khorramdel** (2019), "Introduction of multistage adaptive testing design in PISA 2018", *OECD Education Working Papers*, OECD Publishing, Paris, https://dx.doi.org/10.1787/19939019. [20]

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, http://dx.doi.org/10.1111/emip.12226. [19]
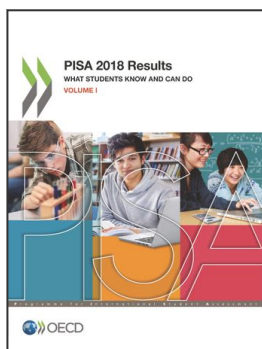
From:
PISA 2018 Results (Volume I)
What Students Know and Can Do

Access the complete publication at:
https://doi.org/10.1787/5f07c754-en