



What can students do in reading?

This chapter presents the various levels of proficiency that students exhibited in the PISA 2018 reading assessment. It describes what students can do at each level of proficiency using items from the actual assessment and the field trial that preceded it. The chapter presents how many students performed at each proficiency level. It then discusses student performance in various specific aspects of reading.



What can students do in reading?

Reading proficiency is essential for a wide variety of human activities – from following instructions in a manual; to figuring out the who, what, when, where and why of a situation; to the many ways of communicating with others for a specific purpose or transaction. Moreover, reading is a component of many other domains of knowledge. For example, real-life problems often require people to draw on their knowledge of mathematics and science, the two other core subjects that PISA tests. Yet in order to do so, people have to be able to read well to obtain the information they need, whether that means reading the nutritional labels on prepared food or comparing car-insurance contracts. People also need to engage in the critical and analytical thinking inherent in reading as they make use of written information for their own purposes.¹

While digitalisation has made sharing non-text-based sources of information, such as videos and images, easier, it has not necessarily done so at the expense of text-based information. In fact, even access to visual or spoken information today often requires some reading: virtually every screen application contains written words (e.g. titles, summaries or comments). If anything, digitalisation has resulted in the emergence and availability of new forms of text. These range from the concise (text messages; memes that combine text with video or images; annotated search engine results; and some online forum posts) to the lengthy (tabbed, multipage websites; newly accessible archival material scanned from microfiches; and some other online forum posts). In other words, reading proficiency will be just as essential in tomorrow's highly digitised world as it is today. Indeed, education systems are increasingly incorporating digital (reading) literacy into their programmes of instruction (Erstad, 2006_[1]; Common Core State Standards Initiative, 2010_[2]).

This chapter describes what students were able to do in the PISA 2018 reading assessment. It focuses, in particular, on the computer-delivered reading assessment. This computer-based test included new text and assessment formats made possible through digital delivery. The test aimed to assess reading literacy in the digital environment while retaining the ability to assess more traditional forms of reading literacy.

What the data tell us

- Some 77% of students, on average across OECD countries, attained at least Level 2 proficiency in reading. At a minimum, these students are able to identify the main idea in a text of moderate length, find information based on explicit, though sometimes complex, criteria, and reflect on the purpose and form of texts when explicitly directed to do so. Over 85% of students in Beijing, Shanghai, Jiangsu and Zhejiang (China), Canada, Estonia, Finland, Hong Kong (China), Ireland, Macao (China), Poland and Singapore performed at this level or above.
- Around 8.7% of students, on average across OECD countries, were top performers in reading, meaning that they attained Level 5 or 6 in the PISA reading test. At these levels, students are able to comprehend lengthy texts, deal with concepts that are abstract or counterintuitive, and establish distinctions between fact and opinion, based on implicit cues pertaining to the content or source of the information. In 20 education systems, including those of 15 OECD countries, over 10% of 15-year-old students were top performers.

THE RANGE OF PROFICIENCY COVERED BY THE PISA READING TEST

Chapter 4 describes students' performance through their placement on the reading, mathematics and science scales. The higher a student scored on the scale, the stronger he or she performed in that particular subject. However, these scores do not indicate what students are actually capable of accomplishing in each subject. This chapter describes what students are able to do in reading; the next two chapters (Chapters 6 and 7) describe students' ability in mathematics and science in greater detail.

As in previous PISA cycles, the reading scale was divided into a range of proficiency levels. Seven of these levels – Levels 1b, 1a, 2, 3, 4, 5 and 6, in ascending order of proficiency – were used to describe reading proficiency in PISA 2009, 2012 and 2015. While the score cut-offs between reading proficiency levels have not changed, the descriptions for all proficiency levels were updated to reflect new aspects of reading that were assessed for the first time in 2018. For example, Levels 3, 4, 5 and 6, as defined in PISA 2018, capture students' ability to assess the quality and credibility of information, and to manage conflict across texts, an aspect of reading literacy that was not highlighted in past assessments (see Chapter 1 for detailed descriptions).

In previous cycles of PISA, there were no tasks to describe the capabilities of students who performed below Level 1b. It was clear that these students could not, in general, successfully perform tasks that were classified at Level 1b, but it was not clear what they actually *could* do. However, all countries, and low-achieving countries in particular, have some 15-year-old students who perform below Level 1b. The PISA for Development programme, in operation between 2015 and 2018 to help eight medium- and low-income countries prepare for full participation in PISA, introduced less-difficult items that were more suitable for students

in these countries (OECD, 2018_[3]). Building on this experience, PISA 2018 introduced new items (beyond those used in PISA for Development) and was able to add a new level, Level 1c, to describe the proficiency of some students who would previously have simply been classified as below Level 1b.

Proficiency scales not only describe student performance; they also describe the difficulty of the tasks presented to students in the assessment. The descriptions of what students at each proficiency level can do and of the typical features of tasks and texts at each level (Table I.5.1) were obtained from an analysis of the tasks located at each proficiency level.² These descriptions were updated from those used in previous PISA cycles to reflect the new reading framework. In particular, Table I.5.1 takes into account the new items created for this assessment (including those at Level 1c) and their increased emphasis on certain forms of text, such as non-continuous texts, texts that span multiple screens and cannot be viewed simultaneously, and multiple-source texts.

Table I.5.1 ^[1/2] **Summary description of the eight levels of reading proficiency in PISA 2018**

Level	Lower score limit	Percentage of students able to perform tasks at each level or above (OECD average)	Characteristics of tasks
6	698	1.3%	<p>Readers at Level 6 can comprehend lengthy and abstract texts in which the information of interest is deeply embedded and only indirectly related to the task. They can compare, contrast and integrate information representing multiple and potentially conflicting perspectives, using multiple criteria and generating inferences across distant pieces of information to determine how the information may be used.</p> <p>Readers at Level 6 can reflect deeply on the text's source in relation to its content, using criteria external to the text. They can compare and contrast information across texts, identifying and resolving inter-textual discrepancies and conflicts through inferences about the sources of information, their explicit or vested interests, and other cues as to the validity of the information.</p> <p>Tasks at Level 6 typically require the reader to set up elaborate plans, combining multiple criteria and generating inferences to relate the task and the text(s). Materials at this level include one or several complex and abstract text(s), involving multiple and possibly discrepant perspectives. Target information may take the form of details that are deeply embedded within or across texts and potentially obscured by competing information.</p>
5	626	8.7%	<p>Readers at Level 5 can comprehend lengthy texts, inferring which information in the text is relevant even though the information of interest may be easily overlooked. They can perform causal or other forms of reasoning based on a deep understanding of extended pieces of text. They can also answer indirect questions by inferring the relationship between the question and one or several pieces of information distributed within or across multiple texts and sources.</p> <p>Reflective tasks require the production or critical evaluation of hypotheses, drawing on specific information. Readers can establish distinctions between content and purpose, and between fact and opinion as applied to complex or abstract statements. They can assess neutrality and bias based on explicit or implicit cues pertaining to both the content and/or source of the information. They can also draw conclusions regarding the reliability of the claims or conclusions offered in a piece of text.</p> <p>For all aspects of reading, tasks at Level 5 typically involve dealing with concepts that are abstract or counterintuitive, and going through several steps until the goal is reached. In addition, tasks at this level may require the reader to handle several long texts, switching back and forth across texts in order to compare and contrast information.</p>
4	553	27.6%	<p>At Level 4, readers can comprehend extended passages in single or multiple-text settings. They interpret the meaning of nuances of language in a section of text by taking into account the text as a whole. In other interpretative tasks, students demonstrate understanding and application of <i>ad hoc</i> categories. They can compare perspectives and draw inferences based on multiple sources.</p> <p>Readers can search, locate and integrate several pieces of embedded information in the presence of plausible distractors. They can generate inferences based on the task statement in order to assess the relevance of target information. They can handle tasks that require them to memorise prior task context.</p> <p>In addition, students at this level can evaluate the relationship between specific statements and a person's overall stance or conclusion about a topic. They can reflect on the strategies that authors use to convey their points, based on salient features of texts (e.g., titles and illustrations). They can compare and contrast claims explicitly made in several texts and assess the reliability of a source based on salient criteria.</p> <p>Texts at Level 4 are often long or complex, and their content or form may not be standard. Many of the tasks are situated in multiple-text settings. The texts and the tasks contain indirect or implicit cues.</p>

...

Table I.5.1 [2/2] Summary description of the eight levels of reading proficiency in PISA 2018

Level	Lower score limit	Percentage of students able to perform tasks at each level or above (OECD average)	Characteristics of tasks
3	480	53.6%	<p>Readers at Level 3 can represent the literal meaning of single or multiple texts in the absence of explicit content or organisational clues. Readers can integrate content and generate both basic and more advanced inferences. They can also integrate several parts of a piece of text in order to identify the main idea, understand a relationship or construe the meaning of a word or phrase when the required information is featured on a single page.</p> <p>They can search for information based on indirect prompts, and locate target information that is not in a prominent position and/or is in the presence of distractors. In some cases, readers at this level recognise the relationship between several pieces of information based on multiple criteria.</p> <p>Level 3 readers can reflect on a piece of text or a small set of texts, and compare and contrast several authors' viewpoints based on explicit information. Reflective tasks at this level may require the reader to perform comparisons, generate explanations or evaluate a feature of the text. Some reflective tasks require readers to demonstrate a detailed understanding of a piece of text dealing with a familiar topic, whereas others require a basic understanding of less-familiar content.</p> <p>Tasks at Level 3 require the reader to take many features into account when comparing, contrasting or categorising information. The required information is often not prominent or there may be a considerable amount of competing information. Texts typical of this level may include other obstacles, such as ideas that are contrary to expectation or negatively worded.</p>
2	407	77.4%	<p>Readers at Level 2 can identify the main idea in a piece of text of moderate length. They can understand relationships or construe meaning within a limited part of the text when the information is not prominent by producing basic inferences, and/or when the text(s) include some distracting information.</p> <p>They can select and access a page in a set based on explicit though sometimes complex prompts, and locate one or more pieces of information based on multiple, partly implicit criteria.</p> <p>Readers at Level 2 can, when explicitly cued, reflect on the overall purpose, or on the purpose of specific details, in texts of moderate length. They can reflect on simple visual or typographical features. They can compare claims and evaluate the reasons supporting them based on short, explicit statements.</p> <p>Tasks at Level 2 may involve comparisons or contrasts based on a single feature in the text. Typical reflective tasks at this level require readers to make a comparison or several connections between the text and outside knowledge by drawing on personal experience and attitudes.</p>
1a	335	92.3%	<p>Readers at Level 1a can understand the literal meaning of sentences or short passages. Readers at this level can also recognise the main theme or the author's purpose in a piece of text about a familiar topic, and make a simple connection between several adjacent pieces of information, or between the given information and their own prior knowledge.</p> <p>They can select a relevant page from a small set based on simple prompts, and locate one or more independent pieces of information within short texts.</p> <p>Level 1a readers can reflect on the overall purpose and on the relative importance of information (e.g. the main idea vs. non-essential detail) in simple texts containing explicit cues.</p> <p>Most tasks at this level contain explicit cues regarding what needs to be done, how to do it, and where in the text(s) readers should focus their attention.</p>
1b	262	98.6%	<p>Readers at Level 1b can evaluate the literal meaning of simple sentences. They can also interpret the literal meaning of texts by making simple connections between adjacent pieces of information in the question and/or the text.</p> <p>Readers at this level can scan for and locate a single piece of prominently placed, explicitly stated information in a single sentence, a short text or a simple list. They can access a relevant page from a small set based on simple prompts when explicit cues are present.</p> <p>Tasks at Level 1b explicitly direct readers to consider relevant factors in the task and in the text. Texts at this level are short and typically provide support to the reader, such as through repetition of information, pictures or familiar symbols. There is minimal competing information.</p>
1c	189	99.9%	<p>Readers at Level 1c can understand and affirm the meaning of short, syntactically simple sentences on a literal level, and read for a clear and simple purpose within a limited amount of time.</p> <p>Tasks at this level involve simple vocabulary and syntactic structures.</p>

However, these descriptions of student proficiency only apply to the computer-based assessment. While the results from countries that conducted the PISA 2018 assessment using pen and paper can be compared to those from countries that delivered the test on computer, countries that used the paper version of the test included only items that were developed for PISA 2009 according to the previous reading framework.³ A description of the proficiency levels that describe what students who sat the paper-based assessment can do can be found in the *PISA 2009 Initial Report* (OECD, 2010_[4]).

Table I.5.2 presents the difficulty level of several released items from both the PISA 2018 main study (i.e. items that were actually used in the assessment) and the PISA 2018 field trial. These items are presented in full in Annex C. Items that illustrate the proficiency levels applicable to the paper-based assessment were presented in the *PISA 2009 Initial Report* (OECD, 2010_[4]).

Table I.5.2 **Map of selected reading questions, illustrating the proficiency levels**

Level	Lower score limit	Question (in descending order of difficulty)	Question difficulty (in PISA score points)
6	698		
5	626	RAPA NUI – Released item 6 (CR551Q10)	665
		COW'S MILK - Released item 5 (CR557Q12)	662
		RAPA NUI – Released item 3 (CR551Q06)	654
		RAPA NUI – Released item 4 (CR551Q08)	634
4	553	RAPA NUI – Released item 5 (CR551Q09)	597
		RAPA NUI – Released item 7 (CR551Q11)	588
		RAPA NUI – Released item 1 (CR551Q01)	559
3	480	COW'S MILK - Released item 3 (CR557Q07)	539
		RAPA NUI – Released item 2 (CR551Q05)	513
		COW'S MILK - Released item 7 (CCR557Q14)	506
		COW'S MILK - Released item 4 (CR557Q10)	498
2	407	CHICKEN FORUM - Released item 7 (CR548Q09)	466
		CHICKEN FORUM - Released item 3 (CR548Q01)	458
		COW'S MILK - Released item 2 (CR557Q04)	452
		CHICKEN FORUM - Released item 6 (CR548Q07)	409
1a	335	COW'S MILK - Released item 6 (CR557Q13)	406
		CHICKEN FORUM - Released item 2 (CR548Q03)	357
		CHICKEN FORUM - Released item 5 (CR548Q05)	347
1b	262	CHICKEN FORUM - Released item 1 (CR548Q02)	328
		CHICKEN FORUM - Released item 4 (CR548Q04)	328
		COW'S MILK - Released item 1 (CR557Q03) <i>Most reading-fluency tasks calling for a "no" response (meaningless sentences, such as "Airplanes are made of dogs")</i>	323
1c	189	<i>Most reading-fluency tasks calling for a "yes" response (meaningful sentences, such as "The red car had a flat tyre") are located at Level 1c or below</i>	

Note: The units COW'S MILK and CHICKEN FORUM were only used in the field trial; estimates of the difficulty level of these items were thus based only on data from the field trial and are reported in italics. Only items in the computer-based assessment (either in the PISA 2018 main survey or its field trial) are included.

PERCENTAGE OF STUDENTS AT THE DIFFERENT LEVELS OF READING PROFICIENCY

Figure I.5.1 presents the distribution of students across the eight levels of reading proficiency. The percentage of students performing at Level 1a or below (i.e. below Level 2) is shown on the left side of the vertical axis.

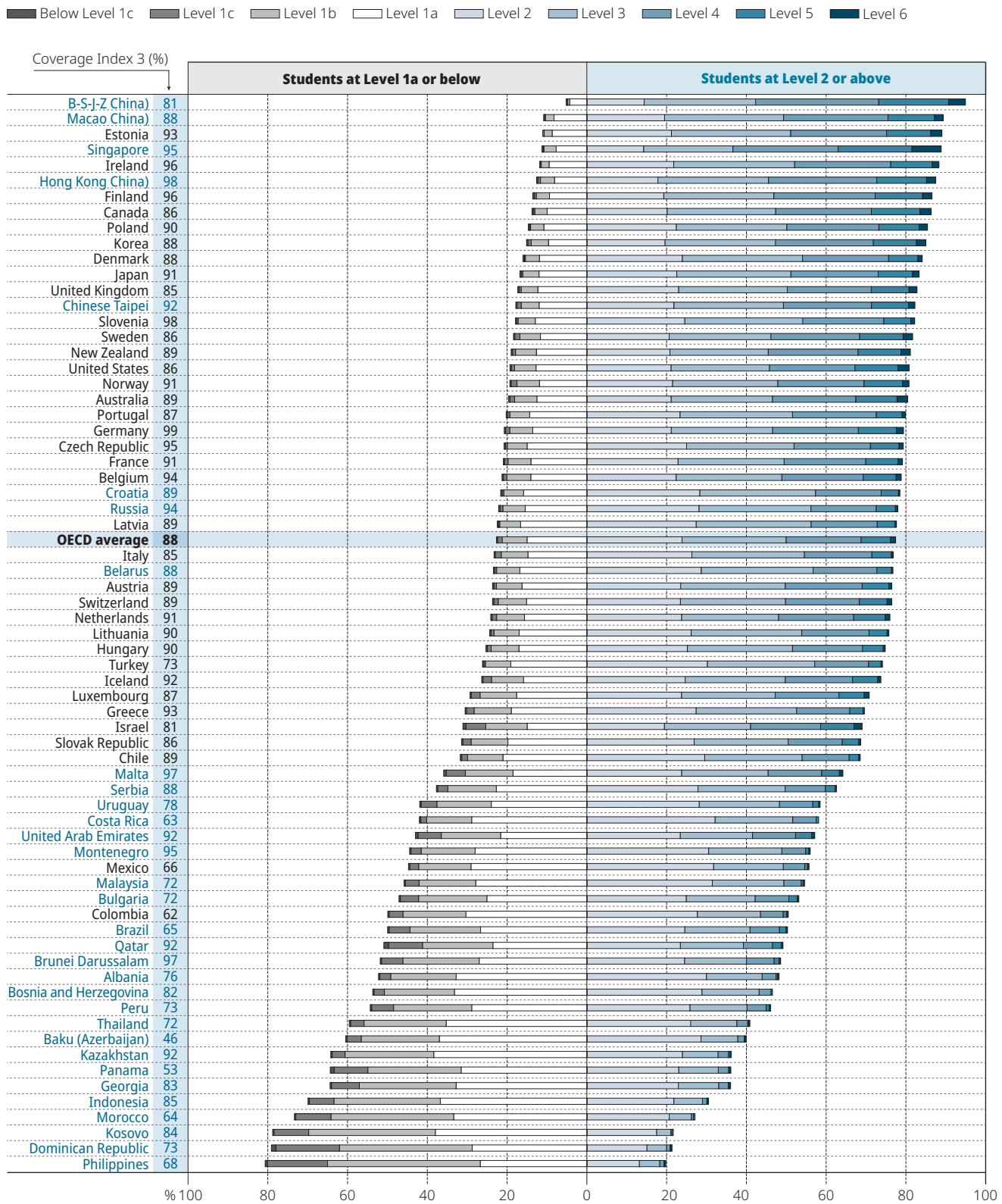
Proficiency at Level 2 or above

At Level 2, students begin to demonstrate the capacity to use their reading skills to acquire knowledge and solve a wide range of practical problems. Students who do not attain Level 2 proficiency in reading often have difficulty when confronted with material that is unfamiliar to them or that is of moderate length and complexity. They usually need to be prompted with cues or instructions before they can engage with a text. In the context of the United Nations Sustainable Development Goals, Level 2 proficiency has been identified as the “minimum level of proficiency” that all children should acquire by the end of secondary education (see Chapter 10).



What can students do in reading?

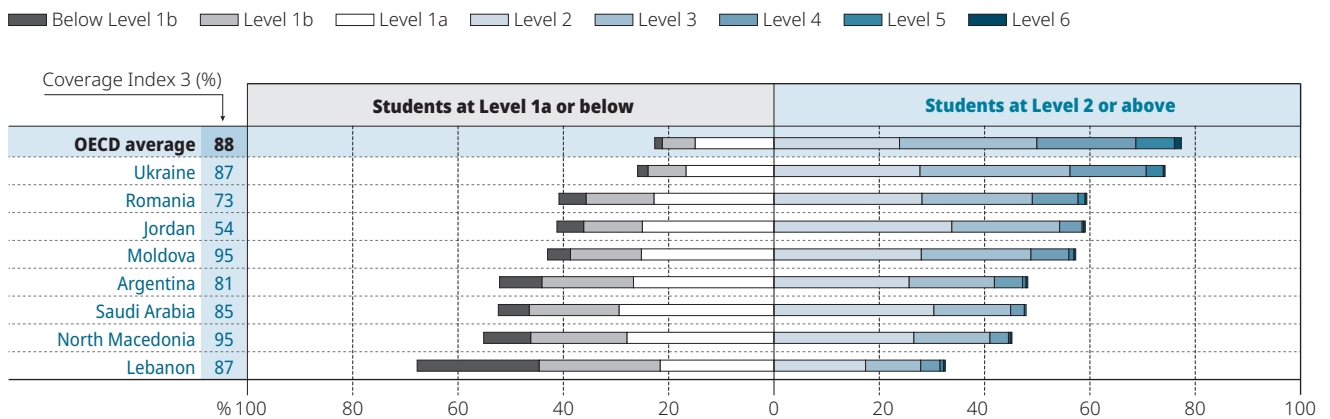
Figure I.5.1 Students' proficiency in reading (computer-based assessment)



Note: Coverage Index 3 is shown next to the country/economy name.
 Countries and economies are ranked in descending order of the percentage of students who performed at or above Level 2.

Source: OECD, PISA 2018 Database, Tables I.B1.1 and I.A2.1.


StatLink <https://doi.org/10.1787/888934028558>

Figure I.5.2 **Students' proficiency in reading (paper-based assessment)**

Note: Coverage Index 3 is shown next to the country name.

Countries are ranked in descending order of the percentage of students who performed at or above Level 2.

Source: OECD, PISA 2018 Database, Tables I.B1.1 and I.A2.1.

StatLink  <https://doi.org/10.1787/888934028577>

But as skill requirements and the contexts in which skills are applied evolve, no particular level of proficiency can be identified as “the one” that signals that students can participate effectively and productively in society. In fact, success in the workplace today, and even more so in the future, may require increasingly higher levels of reading proficiency. Computer scientists interviewed for a recent OECD report (Elliott, 2017^[5]) largely agreed that today’s computers are already capable of solving most of the reading “problems” that students at lower levels of proficiency are capable of solving. Although these artificial intelligence and machine learning technologies may already exist, their diffusion and adoption in the economy is not yet widespread. The effects of such technologies on the demand for reading skills (and for other general cognitive skills) may only become apparent in a few decades.

By acknowledging how our societies are evolving, PISA invites educators and policy makers to consider the proposition that a good education is a moving target: it can never be considered to have been fully attained. While PISA proficiency Level 2 can be considered to be a minimum or baseline level, it is neither a “starting point” from which individuals develop their reading skills nor the “ultimate goal”.

Proficiency at Level 2

At Level 2, students can identify the main idea in a piece of text of moderate length. They can understand relationships or construe meaning within a limited part of the text when the information is not prominent by producing basic inferences, and/or when the information is in the presence of some distracting information. They can select and access a page in a set based on explicit though sometimes complex prompts, and locate one or more pieces of information based on multiple, partly implicit criteria. Readers at Level 2 can, when explicitly cued, reflect on the overall purpose, or on the purpose of specific details, in texts of moderate length. They can reflect on simple visual or typographical features. They can compare claims and evaluate the reasons supporting them based on short, explicit statements.

Typical tasks at Level 2 may involve comparisons or contrasts based on a single feature in the text, or require readers to make a comparison or several connections between the text and outside knowledge by drawing on personal experience and attitudes.

Question 6 from the field-trial unit *CHICKEN FORUM* is a typical “reflecting” task at Level 2. In this unit, students are presented with a series of posts on a forum called “Chicken Health: Your online resource for healthy chickens”. One user, Ivana_88, started a thread asking other users of the forum for advice about her injured hen. Question 6 asked students to identify the person who posted the most reliable answer to her question and to provide a written response justifying their answer. Options A, B and D were all accepted as correct as long as a reasonable justification was provided (e.g. Frank was the most reliable because he said he is a veterinarian or he said he specialises in birds; or NellieB79 was the most reliable because she said that she asks her vet first). This item was classified as “assessing quality and credibility”.

Question 7 in *CHICKEN FORUM* illustrates the capacity of students who are proficient at (at least) Level 2 to generate basic inferences. Ivana_88 asked whether she could give aspirin to her injured hen. In responding to Ivana_88, Frank said that she could, but was unable to give her an exact amount of aspirin to give. Students responding to this “integrate and generate inferences across multiple sources” item were asked to explain why he was unable to do this. Any answer that related to the lack



What can students do in reading?

of information on the size or weight of the hen was accepted as correct (Frank provided the dosage of aspirin per kilogram of body weight but Ivana_88 did not provide the weight of the chicken). As each source of text (i.e. each individual forum post) was short, this was one of the easier items amongst those that required students to use multiple sources of text.

On average across OECD countries in 2018, 77% of students were proficient at Level 2 or higher. In Beijing, Shanghai, Jiangsu and Zhejiang (China) (hereafter “B-S-J-Z [China]”), almost 95% of students performed at or above this benchmark, as did between 88% and 90% of students in Estonia, Ireland, Macao (China) and Singapore. Between 85% and 88% of students in another 4 education systems (Canada, Finland, Hong Kong [China] and Poland) achieved at least Level 2 proficiency, as did between 80% and 85% of students in 11 more education systems (Australia, Denmark, Japan, Korea, New Zealand, Norway, Slovenia, Sweden, Chinese Taipei, the United Kingdom and the United States) (Figure I.5.1).

At the other end of the performance spectrum, over 25% of students, or more than 1 in 4 students, in 10 OECD countries – Chile, Colombia, Greece, Hungary, Iceland, Israel, Luxembourg, Mexico, the Slovak Republic and Turkey – performed below Level 2. However, in all OECD countries, at least 50% of students were still able to attain Level 2 proficiency in reading (Figure I.5.1).

By contrast, in 15 partner education systems that delivered the assessment via computer, including many low- and middle-income countries/economies, more than one in two students scored below Level 2 (Figure I.5.1). Fewer than 1 in 5 students in the Philippines, fewer than 1 in 4 in the Dominican Republic and Kosovo, and fewer than 1 in 3 in Indonesia and Morocco were able to perform at Level 2 or above. This is also true in four countries that assessed students using the pen-and-paper test, which was based on the PISA 2009 test: Argentina, Lebanon, the Republic of North Macedonia (hereafter “North Macedonia”) and Saudi Arabia (Figure I.5.2). All these countries are still far from the objective of equipping all students with the minimum level of reading skills that enables further education and full participation in knowledge-based societies.

Proficiency at Level 3

Tasks at Level 3 require students to take many features into account when comparing, contrasting or categorising information. The required information is often not prominent or there may be a considerable amount of competing information. Texts typical of this level may include other obstacles, such as ideas that are contrary to expectation or negatively worded.

Question 2 of the unit *RAPA NUI* illustrates an “understanding” task at Level 3. The text available to students in this task is a blog post by a professor conducting field work on Easter Island (also known as *Rapa Nui*). The text is illustrated with a picture and contains a couple of short comments by blog readers at the bottom. Question 2 requires students to represent the literal meaning of a particular paragraph in the text (“In the last paragraph of the blog, the professor writes ‘Another mystery remained...’; To what mystery does she refer?”). The open-response format of this question and the fact that to access the paragraph, students must use the scroll bar or mouse (the paragraph is initially hidden) both contribute to the difficulty of the question. Students who answered this question correctly by copying a sentence from the blog post (“What happened to these plants and large trees that had been used to move the *moai*?”) or by paraphrasing it (“Where are the large trees?”) demonstrated the ability to locate target information that is not in a prominent position and to represent the literal meaning of a text.

On average across OECD countries, 54% of students, or just over 1 in 2, were proficient at Level 3 or higher. This describes over 80% of students in B-S-J-Z (China), almost 75% of students in Singapore, and between 65% and 70% of students in Canada, Estonia, Finland, Hong Kong (China), Ireland, Korea and Macao (China). In contrast, fewer than 1 in 5 students in 13 countries and economies that delivered the assessment by computer (all of which are partner countries and economies) was able to perform at Level 3 or higher (Figure I.5.1).

Proficiency at Level 4

A typical Level 4 task might involve texts that are long or complex, whose content or form may not be standard. Many of the tasks are situated in multiple-text settings. They may require students to compare perspectives; evaluate the relationship between specific statements and a person’s overall stance or conclusion about a topic; compare and contrast claims explicitly made in several texts; or assess the reliability of a source based on salient criteria.

At Level 4, readers can comprehend extended passages. They interpret the meaning of nuances of language in a section of text by taking into account the text as a whole.

Question 1 of the unit *RAPA NUI* represents a difficult “scanning and locating” task, demonstrating proficiency at Level 4 (although it is near the lower limit of Level 4 proficiency in difficulty). Students need to consider the blog post provided to them and answer the question “When did the professor start her field work?”. The question is made difficult by the length of the text provided and by the presence of plausible distractors. The correct answer is “Nine months ago” (the blog states: “the *moai* that I have been studying for the past nine months”), but at least two of the possible responses (“One year ago” and “During the 1990s”) are literal matches to distractors in the text (“If you have been following my blog this year”, or “It remained a mystery until the 1990s”).

Question 7 of unit *RAPA NUI* is a typical task measuring students' capacity in "corroborating and handling conflict". In this task, students must consider all three sources provided in the unit – the professor's blog post, a review of the book *Collapse* that is linked in the professor's blog, and an article entitled "Did Polynesian Rats Destroy Rapa Nui's Trees?" which refers to the theory espoused by *Collapse* and presents an alternative theory. The question asks students: "What do you think caused the disappearance of the large trees on Rapa Nui? Provide specific information from the sources to support your answer". There is no single correct answer to this question; rather, answers that received full credit (such as "I think it is because so many trees were cut down to move the statues" or "It's too hard to know based on what I've read. I need more information") demonstrate students' ability to compare and contrast claims explicitly made in several texts. Vague responses (such as "Both", or "We don't know") or responses that did not refer to the theories presented in the source texts (such as "civil war") did not receive credit.

On average across OECD countries in 2018, 28% of students, or just over 1 in 4, attained at least Level 4 in the reading assessment. Over half of the students in the high-performing education systems of B-S-J-Z (China) and Singapore were able to attain this level, while between 35% and 42% of students in a further 10 countries and economies (Canada, Estonia, Finland, Hong Kong [China], Ireland, Korea, Macao [China], New Zealand, Poland and Sweden) performed at Level 4 or above. However, less than 1% of students in the Dominican Republic, Kosovo and Morocco, and only between 1% and 5% of students in another 10 education systems, were proficient at this level or above (Figure I.5.1).

Proficiency at Level 5

Tasks at Level 5 typically involve dealing with concepts that are abstract or counterintuitive, and going through several steps until the goal is reached. In addition, tasks at this level may require the reader to handle several long texts, switching back and forth across texts in order to compare and contrast information.

Question 3 of the unit *RAPA NUI* is a typical Level 5 task, asking students to distinguish between facts and opinions that are expressed in complex and abstract statements. The ability to distinguish fact from opinion is part of the process "reflecting on content and form". In this item, students must classify five distinct statements taken from a review of the book *Collapse* as either "fact" or "opinion". Only students who classified all five statements correctly were given full credit; partial credit was given to students who classified four out of five statements correctly (this corresponds to Level 3 proficiency). The most difficult statement in this list is the first statement ("In the book, the author describes several civilisations that collapsed because of the choices they made and their impact on the environment"). It presents a fact (what the book is about), but some students, particularly those who are proficient below Level 5, may have misclassified this as "opinion" based on the embedded clause, which summarises the book author's theory (the civilisations "collapsed because of the choices they made and their impact on the environment").

Some 8.7% of students performed at Level 5 or above, on average across OECD countries. These students are referred to as top performers in reading. In Singapore, about triple that percentage (26%) were top performers in reading, while in B-S-J-Z (China), 22% of students were top performers. In 18 other countries and economies (including 15 OECD countries), between 10% and 15% of students were top performers in reading. By contrast, in 18 education systems, including Colombia and Mexico, less than 1% of students were classified as top performers in reading (Figure I.5.1).

In countries that used the pen-and-pencil assessment of reading, only single-source processes were assessed. In five of these countries (Argentina, Jordan, Lebanon, North Macedonia and Saudi Arabia), less than 1% of students were classified as top performers (Figure I.5.2).

Proficiency at Level 6

Tasks at Level 6, the highest level of proficiency on the PISA scale, require students to set up elaborate plans in order to achieve a particular goal with the text(s). Readers at Level 6 can comprehend lengthy and abstract texts in which the information of interest is deeply embedded and only indirectly related to the task. They can compare, contrast and integrate information representing multiple and potentially conflicting perspectives, using multiple criteria and generating inferences across distant pieces of information to determine how the information may be used.

Readers at Level 6 can reflect deeply on the text's source in relation to its content, using criteria external to the text. They can compare and contrast information across texts, identifying and resolving inter-textual discrepancies and conflicts through inferences about the sources of information, their explicit or vested interests, and other cues as to the validity of the information.

There are no released items from the PISA 2018 main survey or field trial to illustrate proficiency at Level 6. Altogether, there were ten tasks in the computer-based assessment of reading of Level 6 difficulty. Question 3 in the unit *THE PLAY'S THE THING*, released after the PISA 2009 main study, illustrates some of the competences of students who score at this level. It is based on a long, literary text, a scene from a theatre play. The text describes a fictional world that is remote from the experience



What can students do in reading?

of most 15-year-olds. The theme of the dialogues is abstract (the relationship between life and art, and the challenges of writing for the theatre). Question 3 is particularly difficult because it requires a significant effort of interpretation. The question refers to what the characters (not the actors) were doing “just before the curtain went up”. This requires students to shift between the real world (where there is a curtain and a stage) and the fictional world of the characters, who were in the dining room having dinner just before they entered the guest room, the scene of the play’s action. The task is also difficult because the information about what the characters were doing “before” is not located at the beginning of the text, as one would expect it to be, but about halfway through the text (OECD, 2010, pp. 107-108_[4]).

On average across OECD countries, only 1.3% of students were proficient at Level 6 in reading. This proportion was much higher in some education systems – 7.3% in Singapore, 4.2% in B-S-J-Z (China) and over 2.5% (or over 1 in 40 students) in Australia, Canada, Estonia and the United States. However, in 20 of the 70 PISA-participating education systems that conducted the assessment on computer, fewer than 1 in 1 000 students (0.1%) attained Level 6 in reading. In 5 of these 20 education systems, none of the students who were assessed scored at Level 6 (Figure I.5.1).

Proficiency below Level 2

The PISA 2018 reading assessment identified three proficiency levels below Level 2. PISA considers students who scored at or below these three levels to be low performers in reading.

Proficiency at Level 1a

Tasks at Level 1a, ask students to understand the literal meaning of sentences or short passages, recognise the main theme or the author’s purpose in a piece of text about a familiar topic, or make a simple connection between several adjacent pieces of information, or between the given information and their own prior knowledge. Most tasks at this level point to relevant factors in the task and in the text. Students who perform at Level 1a can select a relevant page from a small set based on simple prompts, and can locate one or more independent pieces of information within short texts. At this level, “reflecting” tasks typically contain explicit cues.

Question 2 in the field-trial unit *CHICKEN FORUM* is a typical Level 1a task. The text in this unit consists of a set of short posts on a web forum, written by distinct authors at different times. Question 2 in this unit asks students: “Why does Ivana_88 decide to post her question on an Internet forum?”. To answer this question correctly, the student must go beyond the literal meaning of the opening post in this forum (signed by user Ivana_88), which states “I can’t get to the veterinarian until Monday, and the vet isn’t answering the phone”, and also consider the full context of her post to identify the correct answer. The process required to identify the correct answer (Option C: “Because she wants to help her hen as soon as possible”) is therefore “integrating and generating inferences”.

Some 15% of students, on average across OECD countries, displayed proficiency at Level 1a but no higher, meaning that they could solve tasks at Level 1a but not those considered to be more difficult; another 7.7% of students did not even attain Level 1a. In 16 education systems – Albania, Baku (Azerbaijan), Bosnia and Herzegovina, Brazil, Brunei Darussalam, Bulgaria, Colombia, Georgia, Indonesia, Kazakhstan, Kosovo, Morocco, Panama, Peru, Qatar and Thailand – Level 1a was the modal level, or the level at which the largest proportion of students scored (Figure I.5.1). This was also true of Argentina and North Macedonia amongst countries that assessed students using the pen-and-paper test (Figure I.5.2).

Proficiency at Level 1b

Tasks located at Level 1b typically use short texts, with minimal competing information, and provide support to the reader, through repetition of information, pictures, familiar symbols or other means. They may require students to evaluate the literal meaning of simple sentences or to interpret the literal meaning of texts by making simple connections between adjacent pieces of information in the question and/or the text.

Readers at this level can scan for and locate a single piece of prominently placed, explicitly stated information in a single sentence, a short text or a simple list. They can access a relevant page in a small set based on simple prompts when explicit cues are present.

Question 1 in the field-trial unit *CHICKEN FORUM* is a typical Level 1b task. The first question in this unit simply asks students to understand the literal meaning of the opening post in this forum thread (“What does Ivana_88 want to know?”). To answer this question correctly, the student must match the paraphrase of Ivana_88’s initial question (“Is it okay to give aspirin to my hen?”) to the options in the item (Option A: “If she can give aspirin to an injured hen”). This is not simply an “accessing and retrieving information within a text” item, but is classified as measuring the process of “understanding the literal meaning”, because there is not a direct, verbatim match between the item options and the stimulus. Some of the most difficult reading-fluency tasks, which ask students to identify whether a single, syntactically simple sentence makes sense, also correspond to Level 1b proficiency (see below, under *Proficiency at Level 1c*).

In addition, question 3 in the unit *BRUSHING YOUR TEETH*, released after the PISA 2009 main study, illustrates the capacity of students who were proficient at Level 1b to find information within short texts based on explicit cues (OECD, 2010, pp. 91-92_[4]). The unit is based on a short text (eight sentences, arranged in three short paragraphs, using familiar syntax) around a topic that most students encounter every day. The question asks: “Why should you brush your tongue, according to Bente Hansen?”, and both “Bente Hansen” and “tongue” can be used to identify the relevant paragraph within the text. Students can quote directly from the text or paraphrase to get credit, but they need to understand that the question is asking about the cause (why?). This task, as well as Question 1 in the field-trial unit *CHICKEN FORUM* described above, show that students described as performing at Level 1b demonstrate a basic degree of understanding, which goes beyond mere decoding skills.

On average across OECD countries, 6.2% of students were able to display proficiency at Level 1b but no higher; 1.4% of students were not even able to complete tasks at this level. Indeed, in 20 education systems, fewer than 1% of students were only able to perform tasks at Level 1b. This proportion was below 0.5% in B-S-J-Z (China), Estonia, Ireland and Macao (China) (Figure I.5.1).

In both the Dominican Republic and the Philippines, the largest share of students scored at Level 1b. In these two countries, between 30% and 40% of students performed at this level; more than 15% of students in these countries could not complete tasks at this level (Figure I.5.1).

Proficiency at Level 1c

Level 1c tasks are the simplest tasks included in the PISA test and involve simple vocabulary and syntactic structures (no task at this level was included in the pen-and-paper test, which was used in nine countries). Readers at Level 1c can understand and affirm the meaning of short, syntactically simple sentences on a literal level, and read for a clear and simple purpose within a limited amount of time.

The simple literal understanding tasks included in the “reading fluency” section at the beginning of the reading test are typical tasks at Level 1c (or below Level 1c). These tasks required test-takers to decide as quickly as possible whether a simple sentence has meaning. Students who score at Level 1c can typically affirm that a meaningful sentence (such as “The red car had a flat tyre” or “The student read the book last night”) indeed has meaning, but some are hesitant to reject meaningless sentences as such (for example, “Airplanes are made of dogs” or “The window sang the song loudly”). These latter items that call for a “no” response are mostly at the 1b level.⁴

Only some 1.4% of students (or roughly 1 in 75) were able to display proficiency at Level 1c but no higher; less than 0.1% of students (or fewer than 1 in 1 000) were unable to display even Level 1c proficiency, on average across OECD countries. By contrast, over 1% of students in the Dominican Republic and Qatar were unable to display even Level 1c proficiency (Figure I.5.1).

Box I.5.1. Accounting for out-of-school 15-year-olds

When evaluating countries’ success in equipping young adults with solid reading, mathematics and science skills, it is also important to consider whether these comparisons could change if 15-year-olds who are not part of the PISA target population were also included. For this reason, Figure I.5.1 reports, next to the name of each country/economy, the proportion of 15-year-olds who were covered by the PISA sample (Coverage Index 3).

In many middle- and low-income countries, less than 75% of 15-year-olds were covered by the PISA sample; indeed, in these countries, a significant portion of 15-year-olds were not eligible to participate in PISA because they had dropped out of school, had never attended school, or were in school but enrolled in grade 6 or below (see Chapter 3). It is not possible to know for certain, in any country, how the 15-year-olds who are not represented by the PISA sample would have scored had they sat the assessment. However, for countries where many 15-year-olds are not enrolled or are retained in grade 6 or below, mean performance and the percentage of students reaching Level 2 or higher would likely be lower than the estimates in this report suggest.

In order to further delimit the possible impact of the 15-year-olds not covered by the PISA sample on skills distributions, it is necessary to make certain assumptions about who they are, and how they would have scored had they sat the PISA test. It is not necessary to attribute an exact score to these 15-year-olds to estimate lower and upper bounds for most results of interest, including the mean score, the median score and other percentiles, or the proportion of 15-year-olds reaching minimum levels of proficiency (Horowitz and Manski, 1995_[6]; Lee, 2009_[7]; Blundell et al., 2007_[8]). For example, several researchers have suggested that out-of-school 15-year-olds, and students who are retained below grade 7, would have scored in the bottom part of a country’s performance distribution (Spaull and Taylor, 2015_[9]; Taylor and Spaull, 2015_[10]).⁵

...

Under a best-case scenario (the distribution of reading, mathematics and science skills in the population not covered by the sample is the same as that of the covered population), the estimates of mean scores and percentiles derived from PISA samples represent an *upper bound* on the means, percentiles and proportions of students reaching minimum proficiency amongst the entire population of 15-year-olds. A *lower bound* can be estimated by assuming a plausible worst-case scenario, such as that all 15-year-olds not covered by the sample would score below a certain point in the distribution. For example, if all of those 15-year-olds would have scored below Level 2, then the lower bound on the proportion of 15-year-olds reaching minimum levels of proficiency would simply be this proportion in the PISA target population multiplied by Coverage Index 3.

Accounting for changing rates of out-of-school 15-year-olds is particularly important when comparing countries' performance over time (see Chapter 8), or when assessing countries' performance against global development goals for the education of all children (see Chapter 9).

STUDENTS' PERFORMANCE IN DIFFERENT ASPECTS OF READING COMPETENCE

In general, scores in any section of the PISA reading assessment are highly correlated with the overall reading score and with scores in other sections. Students who perform well in one aspect of reading also tend to perform well in others. However, there was some variation in performance across different subscales at the country level, which may reflect differences in emphasis in education systems' curriculum and teaching. This section analyses each country's/economy's relative strengths and weaknesses by looking at differences in mean performance across the PISA reading subscales.

Reporting subscales in reading

Two sets of subscales for the reading assessment were developed:

- Process: the main cognitive process required to solve the item (locating information, understanding, or evaluating and reflecting; see Chapter 1 for more details)
- Source: the number of text sources required to construct the correct answer to the item (single source or multiple source).

Subscale scores can be compared within a particular classification of assessment tasks, although not between subscales related to different classifications (i.e. between a process subscale and a source subscale).

However, just like reading and mathematics scales, subscale scores cannot be directly compared even within the same classification (process or source), as each scale measures something different. In order to identify relative strengths and weaknesses, the scores are first standardised by comparison to the mean and standard deviation across all PISA-participating countries. When the standardised score in one subscale is significantly higher than that in another subscale in a country/economy, it can be said to be relatively stronger in the first subscale compared to the average across PISA-participating education systems.

The results that follow only concern countries that conducted the assessment on computer, as the pen-and-paper assessment was based on an earlier framework with different subscales and did not include a sufficient number of tasks to ensure reliable and comparable estimates of subscale proficiency.

Countries' and economies' strengths and weaknesses, by reading process

Each item in the PISA 2018 computer-based reading assessment was classified into one of the three reading processes of "locating information", "understanding" or "evaluating and reflecting". This classification applied at the level of the individual item, not the unit; indeed, items in the same unit could test and emphasise different processes. For example, Questions 1 and 4 in *RAPA NUI* were classified as "locating information"; Questions 2 and 6 as "understanding" ("representing literal meaning" and "integrating and generating inferences"); and Questions 3, 5 and 7 as "evaluating and reflecting" (Question 3: "reflecting on content and form"; Questions 5 and 7: "corroborating and handling conflict").

Table I.5.3 shows the country/economy mean for the overall reading scale and for each of the three reading-process subscales. It also includes an indication of which differences along the (standardised) subscale means are significant, through which a country's/economy's relative strengths and weaknesses can be inferred.

For example, in Norway, mean performance in reading was 499 score points; but performance in the process of "locating information" was 503 points; in the process of "understanding", the score was 498 points; and in "evaluating and reflecting", the score was 502 points. There were no significant differences in how students in Norway performed across different subscales (compared to differences in how students, on average across PISA-participating countries/economies, performed in different subscales) (Table I.5.3).

Table I.5.3^[1/2] Comparing countries and economies on the reading-process subscales

	Mean performance in reading (overall reading scale)	Mean performance on each reading-process subscale			Relative strengths in reading: Standardised mean performance on the reading-process subscale... ¹		
		Locating information	Understanding	Evaluating and reflecting	... locating information (li) is higher than on...	... understanding (un) is higher than on evaluating and reflecting (er) is higher than on...
B-S-J-Z (China)	555	553	562	565		li	li
Singapore	549	553	548	561			li un
Macao (China)	525	529	529	534			
Hong Kong (China)	524	528	529	532			
Estonia	523	529	526	521	er	er	
Canada	520	517	520	527			li un
Finland	520	526	518	517	un er	er	
Ireland	518	521	510	519	un er		un
Korea	514	521	522	522		er	
Poland	512	514	514	514		er	
Sweden	506	511	504	512	un		un
New Zealand	506	506	506	509			
United States	505	501	501	511			li un
United Kingdom	504	507	498	511	un		un
Japan	504	499	505	502		li er	
Australia	503	499	502	513			li un
Chinese Taipei	503	499	506	504		li er	
Denmark	501	501	497	505	un		un
Norway	499	503	498	502			
Germany	498	498	494	497	un er		
Slovenia	495	498	496	494	un er	er	
Belgium	493	498	492	497	un		
France	493	496	490	491	un er		
Portugal	492	489	489	494			un
Czech Republic	490	492	488	489	un er		
OECD average	487	487	486	489	un		
Netherlands	485	500	484	476	un er	er	
Austria	484	480	481	483			
Switzerland	484	483	483	482			
Croatia	479	478	478	474	er	er	
Latvia	479	483	482	477	er	er	
Russia	479	479	480	479		er	
Italy	476	470	478	482		li	li
Hungary	476	471	479	477		li er	li
Lithuania	476	474	475	474			
Iceland	474	482	480	475	er	er	
Belarus	474	480	477	473	un er	er	
Israel	470	461	469	481		li	li un

1. Relative strengths that are statistically significant are highlighted; empty cells indicate cases where the standardised subscale score is not significantly higher compared to other subscales, including cases in which it is lower. A country/economy is relatively stronger in one subscale than another if its standardised score, as determined by the mean and standard deviation of student performance in that subscale across all participating countries/economies, is significantly higher in the first subscale than in the second subscale. Process subscales are indicated by the following abbreviations: li – locating information; un – understanding; er – evaluating and reflecting.


Notes: Only countries and economies where PISA 2018 was delivered on computer are shown.

Although the OECD mean is shown in this table, the standardisation of subscale scores was performed according to the mean and standard deviation of students across all PISA-participating countries/economies.

The standardised scores that were used to determine the relative strengths of each country/economy are not shown in this table.

Countries and economies are ranked in descending order of mean reading performance.

Source: OECD, PISA 2018 Database.

StatLink  <https://doi.org/10.1787/888934028596>

...

Table I.5.3 [2/2] Comparing countries and economies on the reading-process subscales

	Mean performance in reading (overall reading scale)	Mean performance on each reading-process subscale			Relative strengths in reading: Standardised mean performance on the reading-process subscale... ¹		
		Locating information	Understanding	Evaluating and reflecting	... locating information (li) is higher than on...	... understanding (un) is higher than on evaluating and reflecting (er) is higher than on...
Luxembourg	470	470	470	468	er	er	
Turkey	466	463	474	475		li	li
Slovak Republic	458	461	458	457	un er		
Greece	457	458	457	462			
Chile	452	441	450	456		li	li
Malta	448	453	441	448	un er		un
Serbia	439	434	439	434	er	li er	
United Arab Emirates	432	429	433	444		li	li un
Uruguay	427	420	429	433		li	li
Costa Rica	426	425	426	411	er	er	
Cyprus	424	424	422	432	un		li un
Montenegro	421	417	418	416	er	er	
Mexico	420	416	417	426			li un
Bulgaria	420	413	415	416			
Malaysia	415	424	414	418	un er		un
Brazil	413	398	409	419		li	li un
Colombia	412	404	413	417		li	li un
Brunei Darussalam	408	419	409	411	un er		
Qatar	407	404	406	417			li un
Albania	405	394	403	403		li	li
Bosnia and Herzegovina	403	395	400	387	er	er	
Peru	401	398	409	413		li	li un
Thailand	393	393	401	398		li er	
Baku (Azerbaijan)	389	383	386	375	er	er	
Kazakhstan	387	389	394	389	er	er	
Georgia	380	362	374	379		li	li un
Panama	377	367	373	367	er	er	
Indonesia	371	372	370	378	un		un
Morocco	359	356	358	363			li un
Kosovo	353	340	352	353		li	li
Dominican Republic	342	333	342	351		li	li un
Philippines	340	343	335	333	un er		

1. Relative strengths that are statistically significant are highlighted; empty cells indicate cases where the standardised subscale score is not significantly higher compared to other subscales, including cases in which it is lower. A country/economy is relatively stronger in one subscale than another if its standardised score, as determined by the mean and standard deviation of student performance in that subscale across all participating countries/economies, is significantly higher in the first subscale than in the second subscale. Process subscales are indicated by the following abbreviations: li – locating information; un – understanding; er – evaluating and reflecting.


Notes: Only countries and economies where PISA 2018 was delivered on computer are shown.

Although the OECD mean is shown in this table, the standardisation of subscale scores was performed according to the mean and standard deviation of students across all PISA-participating countries/economies.

The standardised scores that were used to determine the relative strengths of each country/economy are not shown in this table.

Countries and economies are ranked in descending order of mean reading performance.

Source: OECD, PISA 2018 Database.

StatLink  <https://doi.org/10.1787/888934028596>

As another example, the mean performance in reading in the Netherlands was 485 score points. However, there was a large range of reading process-subscale scores: 500 points in “locating information”, 484 points in “understanding”, and 476 points in “evaluating and reflecting”. Relative to the average across PISA-participating countries/economies, students in the Netherlands were strongest in “locating information” and weakest in “evaluating and reflecting” (Table I.5.3).

For a final example, although the mean performance in “understanding” and in “evaluating and reflecting” differed by less than 0.1 of a score point in both Korea and Poland, students in both countries performed relatively better in “understanding”, because the average across all PISA-participating countries/economies in “understanding” was lower than it was for “evaluating and reflecting” (Table I.5.3).

Students were relatively stronger in “locating information” than in “understanding”, on average across OECD countries, compared to the worldwide average; this was particularly true in Brunei Darussalam, Ireland, Malaysia, Malta, the Netherlands and the Philippines. By contrast, students in Brazil, Georgia, Kosovo, Peru and Turkey were relatively stronger in “understanding” than in “locating information” (Table I.5.3).

Across OECD countries, there was no significant difference in the relative strength of students in “locating information” and in “evaluating and reflecting”. Students in Brunei Darussalam, Costa Rica, Finland, the Netherlands and the Philippines were relatively stronger in “locating information” than in “evaluating and reflecting”, while the reverse was true in Brazil, the Dominican Republic, Kosovo and Qatar (Table I.5.3).

There was also no significant difference across OECD countries between the relative strength of students in “understanding” and in “evaluating and reflecting”. Students in Bosnia and Herzegovina, Costa Rica, Croatia and Latvia were relatively stronger in “understanding” than in “evaluating and reflecting”, while students in Brazil, the Dominican Republic, Qatar, the United Arab Emirates and the United Kingdom were relatively stronger in “evaluating and reflecting” than in “understanding” (Table I.5.3).

It is also possible to compare mean subscale scores between two countries/economies in the same way as mean reading scores can be compared. For instance, while there was no significant difference in performance between the two highest-performing education systems in the PISA 2018 reading assessment, B-S-J-Z (China) and Singapore, and no significant difference in either the “locating information” or the “evaluating and reflecting” subscales, B-S-J-Z (China) performed significantly higher than Singapore in “understanding” (Tables I.4.1, I.B1.21, I.B1.22, I.B1.23).

Relative strengths and weaknesses of countries/economies, by text source

Each item in the PISA 2018 computer-based reading assessment was assigned to either the single-source or multiple-source text category, depending on the number of sources required to construct the correct answer. In some cases, a unit started with a single stimulus text, and after some initial questions, the scenario was updated to introduce a second text. This was the case, for example, of the field-trial unit *COWS MILK* (see Annex C). Initially, the student was only provided with the “Farm to Market Dairy” webpage. Several questions that focused only on the content of this webpage were initially presented. Then, the scenario was updated, and the student was able to view the second webpage. In other cases, multiple sources were available to students from the outset (as was the case of all questions in the unit *RAPA NUI*), but some questions required only a single source to construct the answer. For example, for the first question in the unit *RAPA NUI*, students were directed to a particular paragraph within the first text although multiple texts were available. In all cases, items were classified by the number of sources required to construct the correct answer, not by the number of sources available in the unit; items in the same unit could be classified differently.

In designing the assessment, care was taken not to confound multiple document settings with the amount of information to be read or the intrinsic complexity of the tasks. Thus, multiple document tasks involving very short simple texts, such as short notes on a bulletin board or mere lists of document titles or search engine results, were also included. These tasks are not intrinsically more difficult than tasks involving single texts of comparable length and complexity.

Table I.5.4 shows the country/economy mean for the overall reading scale and for each of the text-source subscales. It also includes an indication of which differences along the (standardised) subscale means are significant, through which a country's/economy's relative strengths and weaknesses can be inferred.

Standardisation was particularly important for the text-source subscales because in the large majority of countries/economies the multiple-source scores were higher than the single-source scores (such raw differences have no practical meaning). This means that a simple difference in the subscale scores would not show which education systems were relatively stronger in each subscale. Indeed, although the mean multiple-source subscale scores in Australia and Chinese Taipei were both five score points higher than the mean single-source subscale scores, students in neither Australia nor Chinese Taipei were deemed to be relatively stronger at multiple-source reading. By contrast, students in Slovenia were found to be relatively stronger at single-source reading, even though their single-source score was two score points lower than their multiple-source score (Table I.5.4).

Students in OECD countries were found to be relatively stronger in multiple-source reading than students across all PISA-participating countries/economies, on average. This was particularly true of students in Belgium, the Czech Republic, Luxembourg, the Slovak Republic and Switzerland, while students in Colombia, Greece, Indonesia, Montenegro and Morocco were relatively stronger in single-source reading (or relatively weaker in multiple-source reading) (Table I.5.4).

Students develop their competences in all of the reading processes simultaneously; there is no inherent order to the process subscales. On the other hand, the source subscales have a natural sequence: reading single-source texts is a basic skill that precedes the development of competences specific to multiple-source texts. This may explain why countries/economies that are relatively stronger at multiple-source items tend to be higher-performing, on average, than the countries/economies that are relatively weaker at reading multiple-source texts.⁶



What can students do in reading?

Table I.5.4 [1/2] Comparing countries and economies on the single- and multiple-source subscales

	Mean performance in reading (overall reading scale)	Mean performance on each reading text-source subscale		Relative strengths in reading: Standardised mean performance on the reading ... ¹	
		Single text	Multiple text	... single-source text subscale is higher than on the multiple-source texts subscale (ml)	... multiple-source texts subscale is higher than on the single-source text subscale (sn)
B-S-J-Z (China)	555	556	564		sn
Singapore	549	554	553	ml	
Macao (China)	525	529	530		
Hong Kong (China)	524	529	529	ml	
Estonia	523	522	529		sn
Canada	520	521	522		
Finland	520	518	520		
Ireland	518	513	517		
Korea	514	518	525		sn
Poland	512	512	514		
Sweden	506	503	511		sn
New Zealand	506	504	509		sn
United States	505	502	505		
United Kingdom	504	498	508		sn
Japan	504	499	506		sn
Australia	503	502	507		
Chinese Taipei	503	501	506		
Denmark	501	496	503		sn
Norway	499	498	502		
Germany	498	494	497		
Slovenia	495	495	497		
Belgium	493	491	500		sn
France	493	486	495		sn
Portugal	492	487	494		
Czech Republic	490	484	494		sn
OECD average	487	485	490		sn
Netherlands	485	488	495		sn
Austria	484	478	484		sn
Switzerland	484	477	489		sn
Croatia	479	475	478		
Latvia	479	479	483		
Russia	479	477	482		
Italy	476	474	481		sn
Hungary	476	474	480		
Lithuania	476	474	475	ml	
Iceland	474	479	479	ml	
Belarus	474	474	478		
Israel	470	469	471		
Luxembourg	470	464	475		sn
Turkey	466	473	471	ml	
Slovak Republic	458	453	465		sn
Greece	457	459	458	ml	

1. Relative strengths that are statistically significant are highlighted; empty cells indicate cases where the standardised subscale score is not significantly higher compared to other subscales, including cases in which it is lower. A country/economy is relatively stronger in one subscale than another if its standardised score, as determined by the mean and standard deviation of student performance in that subscale across all participating countries/economies, is significantly higher in the first subscale than in the second subscale. Text-source subscales are indicated by the following abbreviations: sn - single text; ml - multiple text.

Notes: Only countries and economies where PISA 2018 was delivered on computer are shown.

Although the OECD mean is shown in this table, the standardisation of subscale scores was performed according to the mean and standard deviation of students across all PISA-participating countries/economies.

The standardised scores that were used to determine the relative strengths of each country/economy are not shown in this table.

Countries and economies are ranked in descending order of mean reading performance.

Source: OECD, PISA 2018 Database.

StatLink <https://doi.org/10.1787/888934028615>

...

Table I.5.4 [2/2] Comparing countries and economies on the single- and multiple-source subscales

	Mean performance in reading (overall reading scale)	Mean performance on each reading text-source subscale		Relative strengths in reading: Standardised mean performance on the reading ... ¹	
		Single text	Multiple text	... single-source text subscale is higher than on the multiple-source texts subscale (ml)	... multiple-source texts subscale is higher than on the single-source text subscale (sn)
Chile	452	449	451	ml	
Malta	448	443	448		
Serbia	439	435	437	ml	
United Arab Emirates	432	433	436		
Uruguay	427	424	431		
Costa Rica	426	424	427		
Cyprus	424	423	425	ml	
Montenegro	421	417	416	ml	
Mexico	420	419	419	ml	
Bulgaria	420	413	417		
Malaysia	415	414	420		
Brazil	413	408	410		
Colombia	412	411	412	ml	
Brunei Darussalam	408	408	415		
Qatar	407	406	410		
Albania	405	400	402	ml	
Bosnia and Herzegovina	403	393	398		
Peru	401	406	409		
Thailand	393	395	401		
Baku (Azerbaijan)	389	380	386		
Kazakhstan	387	391	393	ml	
Georgia	380	371	373	ml	
Panama	377	370	371	ml	
Indonesia	371	373	371	ml	
Morocco	359	359	359	ml	
Kosovo	353	347	352		
Dominican Republic	342	340	344		
Philippines	340	332	341		sn

1. Relative strengths that are statistically significant are highlighted; empty cells indicate cases where the standardised subscale score is not significantly higher compared to other subscales, including cases in which it is lower. A country/economy is relatively stronger in one subscale than another if its standardised score, as determined by the mean and standard deviation of student performance in that subscale across all participating countries/economies, is significantly higher in the first subscale than in the second subscale. Text-source subscales are indicated by the following abbreviations: sn - single text; ml - multiple text.


Notes: Only countries and economies where PISA 2018 was delivered on computer are shown.

Although the OECD mean is shown in this table, the standardisation of subscale scores was performed according to the mean and standard deviation of students across all PISA-participating countries/economies.

The standardised scores that were used to determine the relative strengths of each country/economy are not shown in this table.

Countries and economies are ranked in descending order of mean reading performance.

Source: OECD, PISA 2018 Database.

StatLink  <https://doi.org/10.1787/888934028615>

What can students do in reading?

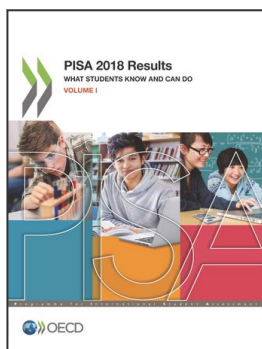
1. See Chapter 1 of this report for more details about how PISA 2018 conceptualised reading and about how reading has evolved over the past decade.
2. The cut-off scores for proficiency levels were defined in earlier PISA cycles; for further details, please see the *PISA 2018 Technical Report* (OECD, forthcoming^[11]).
3. Certain items were common to both the paper-based and computer-based assessments. These items were originally developed for the PISA 2009 reading assessment (based on the 2009 framework) and were converted into a computer-based format for PISA 2015, the first year in which PISA was primarily delivered on computer. A mode-effect study was then conducted to assure the equivalence of common items across modes; the item parameters of difficulty and discrimination were allowed to differ across modes if necessary (see Annex A5). This allowed for the comparison of countries/economies across modes of test delivery, and for the calculation of trends in performance across years as all countries, including those that delivered the test via computer in 2015 or 2018, would have delivered the test on paper in 2012 and before.
4. Based on the above description, it is possible that students who were classified at Level 1c simply responded “yes” to all reading-fluency items without making an active decision about the meaning of each sentence. An analysis of student effort in reading-fluency items (see Annex A8) shows that there were students who “straightlined” their responses over the 21 or 22 reading-fluency items (i.e. who answered “yes” to all questions or “no” to all questions), and that this proportion was larger amongst low-performing students. Indeed, between 10% and 14% of low-performing students in the Dominican Republic, Indonesia, Israel, Kazakhstan, Korea and the Philippines straightlined their responses to reading-fluency items (Table I.A8.21).

However, although most items that called for a “no” response (i.e. affirming that the sentence did not have meaning) were classified at Level 1b, two such items were classified at Level 1c. Hence, a large proportion of students at Level 1c were able to identify these sentences as being without meaning and did not simply respond “yes” to all reading-fluency items. Moreover, the presence of reading-fluency items below Level 1c indicates that students at Level 1c are able to confirm that relatively more complicated phrases have meaning, which students below Level 1c cannot do. More work is needed to fully understand what students at Level 1c can do, including further analysis of student response time and response patterns, and a description of the differences in reading-fluency items that are classified as at Level 1c and as below that level.

5. More generally, one could assume that the distribution of skills in the population not covered by the PISA sample is *stochastically dominated* by the distribution of skills in the covered population. This means that the best-performing 15-year-old who is not covered by the sample would score at the same level, at best, as the best-performing 15-year-old in the covered population, that the 90th percentile (the score above which only 10% of the population lie) of the non-covered population is, at best, equal to the 90th percentile of the covered population, and similarly for every percentile along the distribution.
6. The country-level relationship between overall mean performance in reading and differences in the single- and multiple-source subscales has a positive slope and an R^2 of 0.12.

References

- Blundell, R.** et al. (2007), “Changes in the Distribution of Male and Female Wages Accounting for Employment Composition Using Bounds”, *Econometrica*, Vol. 75/2, pp. 323-363, <http://dx.doi.org/10.1111/j.1468-0262.2006.00750.x>. [8]
- Common Core State Standards Initiative** (2010), *Common Core State Standards for English Language Arts & Literacy in History/Social Studies, Science, and Technical Subjects*, http://www.corestandards.org/wp-content/uploads/ELA_Standards1.pdf. [2]
- Elliott, S.** (2017), *Computers and the Future of Skill Demand*, Educational Research and Innovation, OECD Publishing, Paris, <http://dx.doi.org/10.1787/9789264284395-en>. [5]
- Erstad, O.** (2006), “A new direction?”, *Education and Information Technologies*, Vol. 11/3-4, pp. 415-429, <http://dx.doi.org/10.1007/s10639-006-9008-2>. [1]
- Horowitz, J.** and **C. Manski** (1995), “Identification and Robustness with Contaminated and Corrupted Data”, *Econometrica*, Vol. 63/2, pp. 282-302, <http://dx.doi.org/10.2307/2951627>. [6]
- Lee, D.** (2009), “Training, wages, and sample selection: Estimating sharp bounds on treatment effects”, *The Review of Economic Studies*, Vol. 76/3, pp. 1071-1102, <http://dx.doi.org/10.1111/j.1467-937X.2009.00536.x>. [7]
- OECD** (2018), *PISA for Development Assessment and Analytical Framework: Reading, Mathematics and Science*, PISA, OECD Publishing, Paris, <https://dx.doi.org/10.1787/9789264305274-en>. [3]
- OECD** (2010), *PISA 2009 Results: What Students Know and Can Do: Student Performance in Reading, Mathematics and Science (Volume I)*, PISA, OECD Publishing, Paris, <https://dx.doi.org/10.1787/9789264091450-en>. [4]
- OECD** (forthcoming), *PISA 2018 Technical Report*, OECD Publishing, Paris. [11]
- Spaull, N.** and **S. Taylor** (2015), “Access to What? Creating a Composite Measure of Educational Quantity and Educational Quality for 11 African Countries”, *Comparative Education Review*, Vol. 59/1, pp. 133-165, <http://dx.doi.org/10.1086/679295>. [9]
- Taylor, S.** and **N. Spaull** (2015), “Measuring access to learning over a period of increased access to schooling: The case of Southern and Eastern Africa since 2000”, *International Journal of Educational Development*, Vol. 41, pp. 47-59, <http://dx.doi.org/10.1016/j.ijedudev.2014.12.001>. [10]



From:
PISA 2018 Results (Volume I)
What Students Know and Can Do

Access the complete publication at:

<https://doi.org/10.1787/5f07c754-en>

Please cite this chapter as:

OECD (2019), "What can students do in reading?", in *PISA 2018 Results (Volume I): What Students Know and Can Do*, OECD Publishing, Paris.

DOI: <https://doi.org/10.1787/99c14352-en>

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 OECD member countries.

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>.