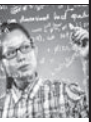
3

Main Results from the PISA 2012 Computer-Based Assessments

Computer-based tests expand the range of situations in which students' ability to apply their knowledge can be measured. Students in 32 countries and economies that participated in the PISA 2012 pencil-and-paper assessment were invited to take a test of reading and mathematics on computers. This chapter discusses the results of those computer-based assessments.



In 32 countries and economies, students who participated in the PISA 2012 pencil-and-paper assessment were also invited to take a test of reading and mathematics on computers.[1] This latter assessment included 18 reading questions originally developed for use in the 2009 assessment of digital reading, and 41 specially designed mathematics questions. This chapter reports results from the PISA 2012 computer-based assessments.

What the data tell us

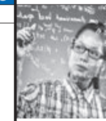
- Singapore, followed by Korea, Hong Kong-China, Japan, Canada and Shanghai-China were the top-performing countries/economies in digital reading in 2012; Singapore and Shanghai-China, followed by Korea, Hong Kong-China, Macao-China, Japan and Chinese Taipei were top performers in the 2012 computer-based mathematics assessment.
- In Korea and Singapore, students score more than 20 points higher on the digital reading scale, on average, than students in other countries with similar skills in print reading.
- Students in Australia, Austria, Canada, Japan, Slovenia and the United States, as well as students in partner countries/economies Macao-China and the United Arab Emirates, perform better on mathematics tasks that require the use of computers to solve problems compared to their success on traditional tasks. By contrast, students in Belgium, Chile, France, Ireland, Poland and Spain perform worse than expected on such tasks, given their performance on traditional mathematics tasks.

While both reading and mathematics tasks in the computer-based assessment were developed within the same framework as their corresponding paper-based tasks, the results of the former assessment are reported on separate scales. Indeed, computer-based tests expand the range of situations in which reading and mathematics are assessed in PISA. A key feature of digital reading tasks is that they use the typical text formats encountered on line; as a result, many of them require students to navigate through and across texts by using such tools as hyperlinks, browser button or scrolling, in order to access the information. The design of mathematics tasks, on the other hand, ensured that mathematical reasoning and processes take precedence over mastery of using the computer as a tool. Several tasks, however, also involve typical situations in which information and communication tools, such as using spreadsheets to collect data or create a chart, help to solve mathematics problems.

Demands for general knowledge and skills related to computers were kept to a minimum. They included using a keyboard and mouse, and knowing common conventions, such as arrows to move forward. A short introduction to the test provided all students with the opportunity to practice using the tools through which they could interact with the test items, as well as response formats.

SIMILARITIES AND DIFFERENCES BETWEEN PAPER-BASED AND COMPUTER-BASED ASSESSMENTS

This section highlights what is particular about the computer-based assessments of reading and mathematics in PISA 2012. The discussion starts by highlighting differences with paper-based



assessments in what is assessed, and ends by looking at how proficiency is assessed. More details about the framework for these assessments can be found in the framework publication (OECD, 2013); details about the test design and operational characteristics can be found in the technical report (OECD, 2014a).

Items from units *SERAING*, *SPORTS CLUB* and *LANGUAGE LEARNING* – three digital reading units used in the PISA 2012 assessment – can be seen, and tested, on the website of the Australian Council for Educational Research (http://cbasq.acer.edu.au/index.php?cmd=toEra2012). Items from three computer-based mathematics units used in the PISA 2012 main survey (*CD PRODUCTION*, *STAR POINTS* and *BODY MASS INDEX*), as well as items from four field-trial units, can also be found on the same website (http://cbasq.acer.edu.au/index.php?cmd=toMaths). All main survey items are available in 91 languages.

Differences between digital and print reading

The framework for reading treats digital and print reading as a single domain, while acknowledging the differences between reading on paper and reading on digital platforms. These differences are reflected in the assessment tasks used to assess reading in the two media.

First, in a typical Internet reading situation, the reader is generally unable to see the physical amount of text available; at the same time, he can access multiple sources more easily than in a print environment. While there are offline situations where readers need to consult several printed documents, the PISA assessment makes minimal use of such situations. All stimulus material fits onto a single page in the PISA assessment of print reading, and this limits the extent to which texts from multiple sources can be used. By contrast, because reading on the Internet usually involves referring to several pages, and often to several texts from different sources, composed by different authors and appearing in different formats, it was important that the computer-based assessment allowed for the possibility of using multiple texts simultaneously.

Another distinction between digital and print reading is the text types that are typical of each medium. Much reading in the digital medium involves personal communications and exchanges that aim to achieve a specific purpose (transactions), as in e-mails and text messages that set the date of a meeting or ask a friend for a suggestion. Narrative texts, in contrast, are more common in print reading. As a consequence, there are no assessment tasks in the digital reading assessment that are based on narrative texts, whereas transaction texts are absent from the print reading assessment in PISA.

Finally, while the major cognitive processes involved in print and digital reading are the same, performing tasks that demand these processes may pose a greater challenge in the digital medium than on paper, because navigation is required (see Chapter 4). *Access and retrieve* tasks, for instance, involve locating information: on line, readers need to search for information in a more abstract space than in printed books or documents, without seeing the full text. Search tools are also specific to each medium: search engines and menus on line, tables of contents and indices in printed documents. *Integrate and interpret* tasks require readers to contrast or compare information from different locations. In digital reading, such tasks often involve multiple texts and diverse text formats; and because the texts are usually not visible simultaneously, readers must rely on their short-term memory to perform these tasks. *Reflection and evaluation* processes tend to be required only for the most difficult tasks on paper. In contrast, when reading on line, readers must often assess the credibility of the content even when solving simple tasks, given that there are fewer filters between the author and the reader to decide what is published.

Knowledge of some techniques of navigation and some navigation tools (e.g. hyperlinks, tabs, menus, the "back" button) are part of being literate in the digital medium. Such skills and knowledge should be regarded as ICT skills that are measured, together with the mastery of reading processes, in the assessment of digital reading.

Differences between computer-based and paper-based mathematics

The computer-based assessment of mathematics recognises that mathematical competency in the 21st century includes usage of computers. Indeed, computers offer tools to describe, explain, or predict phenomena by employing mathematical concepts, facts, procedures and reasoning. Students' ability to use these tools is an aspect of mathematical literacy that could not be assessed on paper, and was only assessed in the computer-based assessment. Conversely, the mathematical competencies that are tested on paper are all represented in the computer-based assessment of mathematics (although the small number of tasks means that not all of them could be covered well).

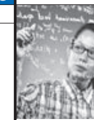
Thus, the main difference between the paper-based and computer-based mathematics assessment is that only in the latter are skills related to using ICT tools for mathematics tasks assessed. These skills include using computers to make a chart from data, produce graphs of functions, sort data sets, use on-screen calculators, use virtual instruments, or use a mouse or a dialog box to rotate, translate, or reflect a geometrical figure.

Differences in test design and operational characteristics of computer- and paper-based assessments

In addition to differences in the constructs of the reading and mathematics assessments, there are differences in how tests were administered. The obvious difference is that the paper-based assessments were completed with pen and paper as part of a two-hour test session. By contrast, computer-based assessments were completed with a keyboard and mouse, while looking at questions on a screen, and lasted only 40 minutes.

A consequence of the difference in testing time is that more items were used in the print reading and paper-based mathematics assessments than in the digital reading and computer-based mathematics assessments. For this reason, the uncertainty associated with the measurement of performance is greater in the computer-based tests, particularly at very high or very low levels of proficiency. In addition, results are only reported on a single, global scale, not on subscales.

Not all students who sat the paper-based assessments completed the computer-based assessment; nor did they necessarily encounter questions from the digital reading assessment or the computer-based mathematics assessment in their test forms. In fact, in the 32 countries that participated in the optional computer-based assessments of reading and mathematics, only about half of all students who were sampled for PISA within each participating school were also invited to take a



computer-based test. And because three domains (digital reading, computer-based mathematics and problem solving) were assessed on computers, of all students who were sampled for the computer-based test, only two out of three encountered questions from a particular domain in their forms.

STUDENT PERFORMANCE IN DIGITAL READING

PISA outcomes are reported in a variety of ways. This section gives the country results and shows how performance varies within and across countries. In addition, it shows trends in the digital reading performance of countries/economies that participated in both the PISA 2009 and PISA 2012 assessments.

When digital reading was assessed for the first time in 2009, the scale was fixed so that the average mean score and standard deviation for OECD countries would match those of the print reading scale for the same year and the same countries (OECD, 2011). In 2012, results were reported on the same scale as in 2009.

Average performance in digital reading

When comparing countries and economies on the basis of their average digital reading score, it is important to remember that not all performance differences observed between countries are statistically significant. In other words, because the PISA survey is based on a sample of students and a limited number of items, some small differences may be observed by chance, even when there are no differences in the true proficiency of students on average. When interpreting mean performance, only those differences among countries and economies that are statistically significant should be taken into account. These are differences that are large enough – so large in fact as to make it highly unlikely that the difference observed among samples of students does not reflect a true difference in the populations from which these students are drawn.

Figure 3.1 lists each participating country and economy in descending order of its mean digital-reading score (left column). The values range from a high of 567 points for partner country Singapore to a low of 396 points for partner country Colombia. Countries and economies are also divided into three broad groups: those whose mean scores are not statistically different from the mean for the 23 OECD countries participating in the assessment (highlighted in dark blue), those whose mean scores are significantly above the OECD mean (highlighted in pale blue), and those whose mean scores are significantly below the OECD mean. The best-performing OECD country is Korea, followed by Japan. Partner country Singapore performs better than all other countries and economies, including Korea, while the performance of Hong Kong-China is not statistically different from that of Korea or Japan. Canada, Shanghai-China, Estonia, Australia, Ireland, Chinese Taipei, Macao-China, the United States, France and Belgium (in decreasing order of mean performance) also perform above the OECD average, but below the four best-performing countries and economies.

Because the figures are derived from samples, it is not possible to determine a country's precise rank among the participating countries/economies. However, it is possible to determine, with confidence, a range of ranks in which the performance of the country/economy lies (Figure 3.2).

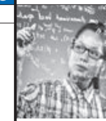
■ Figure 3.1 ■

Comparing countries' and economies' performance in digital reading

Statistically significantly **above** the OECD average
Not statistically significantly different from the OECD average
Statistically significantly **below** the OECD average

Mean score	Comparison country/economy	Countries/economies whose mean score is NOT statistically significantly different from that of the comparison country/economy
567	Singapore	
555	Korea	Hong Kong-China
550	Hong Kong-China	Korea, Japan
545	Japan	Hong Kong-China
532	Canada	Shanghai-China
531	Shanghai-China	Canada, Estonia
523	Estonia	Shanghai-China, Australia, Ireland, Chinese Taipei
521	Australia	Estonia, Ireland, Chinese Taipei, United States
520	Ireland	Estonia, Australia, Chinese Taipei, Macao-China, United States, France
519	Chinese Taipei	Estonia, Australia, Ireland, Macao-China, United States, France
515	Macao-China	Ireland, Chinese Taipei, United States, France
511	United States	Australia, Ireland, Chinese Taipei, Macao-China, France, Italy, Belgium
511	France	Ireland, Chinese Taipei, Macao-China, United States, Italy, Belgium
504	Italy	United States, France, Belgium, Norway, Sweden, Denmark
502	Belgium	United States, France, Italy, Norway, Sweden
500	Norway	Italy, Belgium, Sweden, Denmark
498	Sweden	Italy, Belgium, Norway, Denmark
495	Denmark	Italy, Norway, Sweden, Portugal
486	Portugal	Denmark, Austria, Poland
480	Austria	Portugal, Poland, Slovak Republic
477	Poland	Portugal, Austria, Slovak Republic, Slovenia, Spain, Russian Federation
474	Slovak Republic	Austria, Poland, Slovenia, Spain, Russian Federation
471	Slovenia	Poland, Slovak Republic, Spain, Russian Federation
466	Spain	Poland, Slovak Republic, Slovenia, Russian Federation, Israel
466	Russian Federation	Poland, Slovak Republic, Slovenia, Spain, Israel
461	Israel	Spain, Russian Federation, Chile, Hungary
452	Chile	Israel, Hungary
450	Hungary	Israel, Chile
436	Brazil	
407	United Arab Emirates	
396	Colombia	

Source: OECD, PISA 2012 Database.
StatLink http://dx.doi.org/10.1787/888933252891



■ Figure 3.2 ■

Where countries and economies rank in digital reading performance

	Digital reading scale					
			Range of ranks			
			OECD countries		All countries/economies	
	Mean score	S.E.	Upper rank	Lower rank	Upper rank	Lower rank
Singapore	567	(1.2)			1	1
Korea	555	(3.6)	1	1	2	3
Hong Kong-China	550	(3.6)			2	4
Japan	545	(3.3)	2	2	3	4
Canada	532	(2.3)	3	3	5	6
Shanghai-China	531	(3.7)			5	6
Estonia	523	(2.8)	4	6	7	10
Australia	521	(1.7)	4	6	7	10
Ireland	520	(3.0)	4	7	7	11
Chinese Taipei	519	(3.0)			7	11
Macao-China	515	(0.9)			10	12
United States	511	(4.5)	6	10	10	15
France	511	(3.6)	7	9	10	14
Italy	504	(4.3)	7	12	12	17
Belgium	502	(2.6)	9	12	14	17
Norway	500	(3.5)	9	13	14	18
Sweden	498	(3.4)	9	13	14	18
Denmark	495	(2.9)	11	14	16	19
Portugal	486	(4.4)	13	16	18	21
Austria	480	(3.9)	14	17	19	22
Poland	477	(4.5)	14	18	19	23
Slovak Republic	474	(3.5)	15	19	20	24
Slovenia	471	(1.3)	17	19	22	24
Spain	466	(3.9)	17	20	23	26
Russian Federation	466	(3.9)			23	26
Israel	461	(5.1)	19	22	24	28
Chile	452	(3.6)	20	22	26	28
Hungary	450	(4.4)	21	22	26	28
Brazil	436	(4.9)			29	29
United Arab Emirates	407	(3.3)			30	30
Colombia	396	(4.0)			31	31

Source: OECD, PISA 2012 Database.
StatLink http://dx.doi.org/10.1787/888933252903

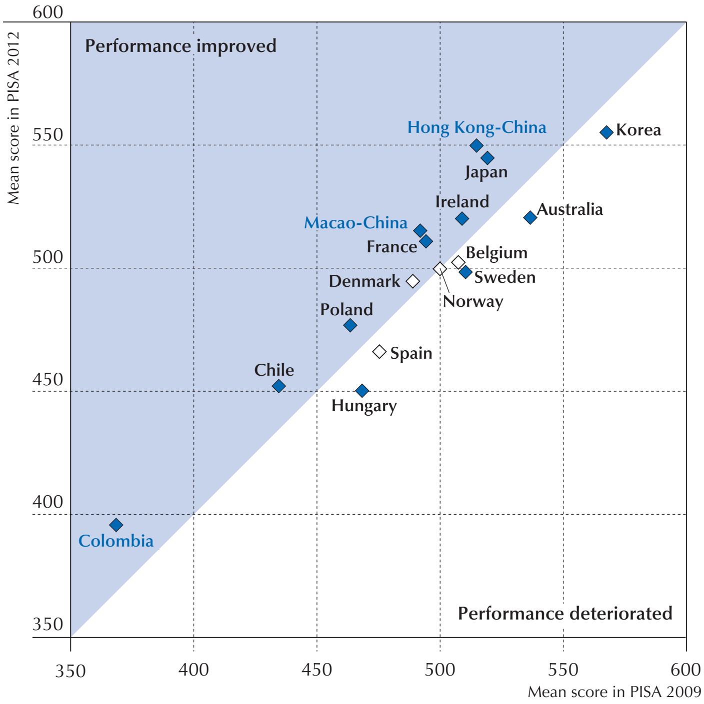
Trends in average digital reading performance

PISA 2012 marks the second time digital reading was assessed in PISA, with tasks that are built around the typical text formats encountered on line. Of the 19 countries and economies that participated in the digital reading assessment in 2009, 17 renewed their participation in 2012 (Iceland and New Zealand are the exceptions). Because the PISA 2012 assessment of digital reading uses a subset of the items developed and used in PISA 2009, results from the two assessments can be compared over time.

Among the 16 countries and economies for which results can be compared over time,[2] four show a decline in the mean performance of their students, four show no change, and eight countries and economies show a significant improvement in performance (Figure 3.3).

■ Figure 3.3 ■

Digital reading performance in 2009 and 2012



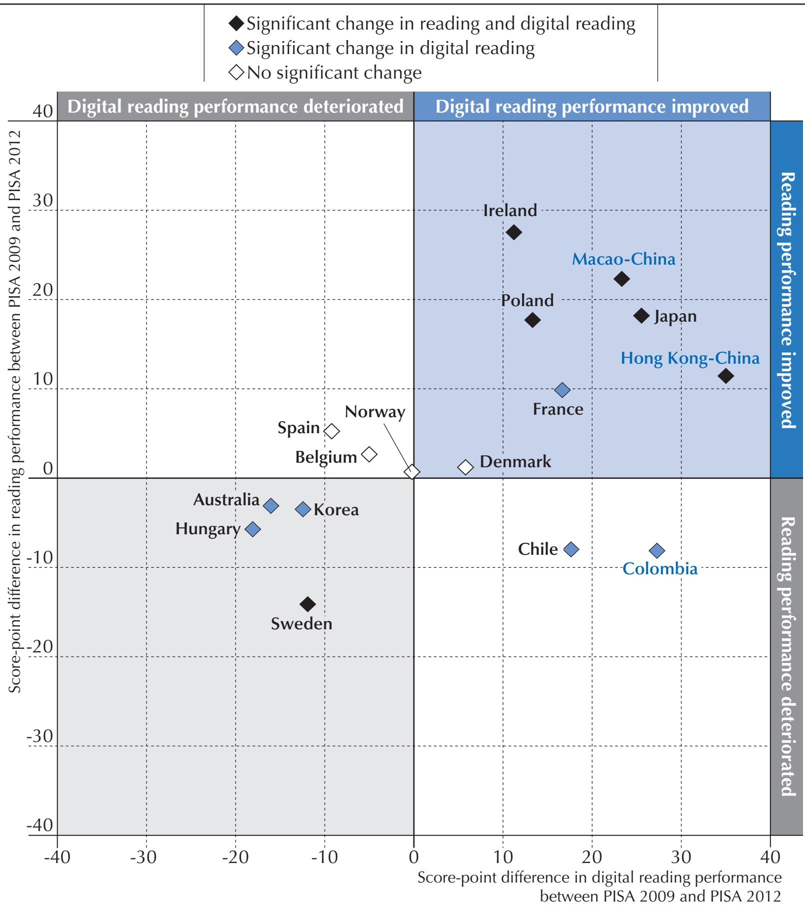
Note: Statistically significant score-point changes between PISA 2012 and PISA 2009 are marked in a darker tone.
Source: OECD, PISA 2012 Database, Table 3.2.
StatLink http://dx.doi.org/10.1787/888933252910

The largest improvement in average performance is observed in Hong Kong-China, where students scored 35 score points higher, on average, than they did in 2009. Significant improvements in average performance are also observed in Colombia, Japan, Macao-China, Chile, France, Poland and Ireland, in decreasing order of their magnitude. A stable mean performance is found in Belgium, Denmark, Norway and Spain. In Australia, Hungary, Korea and Sweden, students in 2012 performed more than ten points below the level achieved by students in 2009. Korea was the top-performing country in 2009, with a mean score of 568 points, almost 50 points above Hong Kong-China and Japan. By 2012, students in Korea performed on par with students in Hong Kong-China (Figure 3.3).

In general, trends in digital reading performance are highly correlated to trends in print reading performance. Figure 3.4 shows that most countries where digital reading performance improved also saw similar gains in their print reading performance. The most notable exceptions are Chile and Colombia, where digital reading performance improved significantly, but performance on the print reading assessment remained stable. These trends are examined further below.

■ Figure 3.4 ■

Change in digital and print reading performance between 2009 and 2012

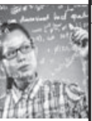


Source: OECD, PISA 2012 Database, Table 3.2.
StatLink http://dx.doi.org/10.1787/888933252926

Students at the different levels of proficiency in digital reading

This section describes performance in terms of the levels of proficiency that were constructed for reporting the PISA 2009 digital reading scale. Because the PISA digital reading assessment is a short test based on a limited number of tasks, only four proficiency levels could be described, rather than the usual six. The lowest described level of proficiency is equivalent to Level 2 on the reading scale, and corresponds to a baseline level of proficiency in digital reading. The highest described level of proficiency is equivalent to Level 5 on the reading scale.

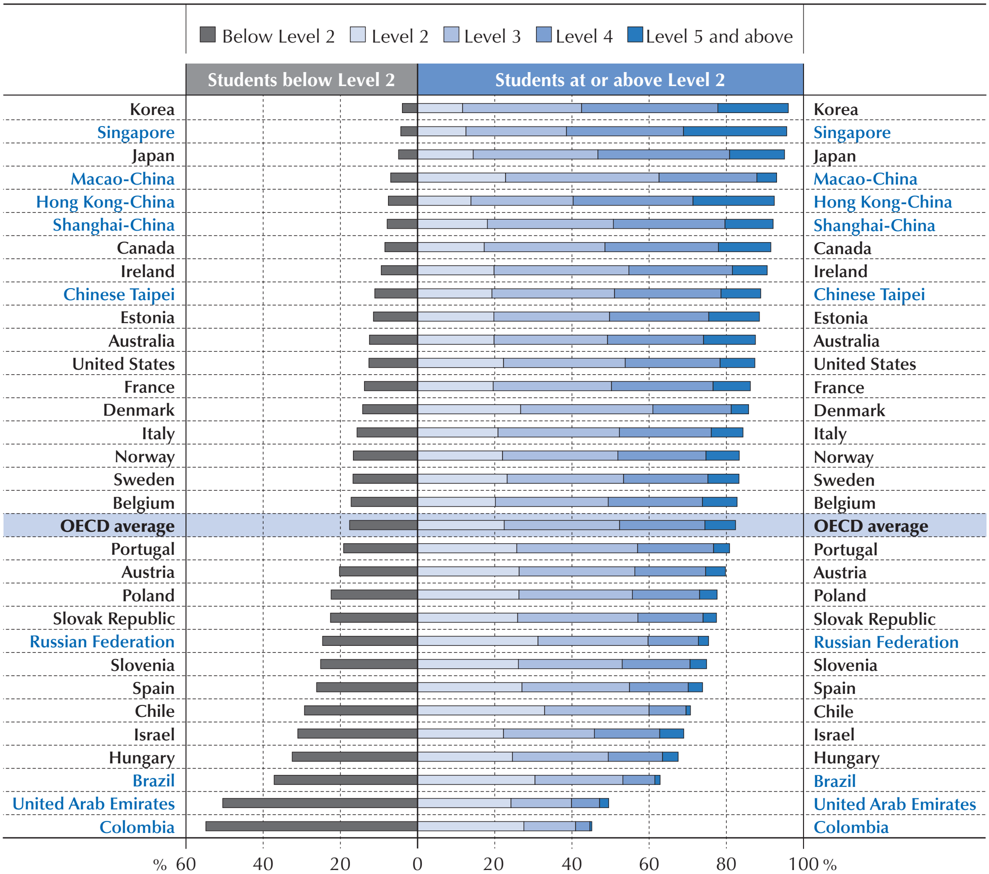
The distribution of student performance across these proficiency levels in each participating country is shown in Figure 3.5. A detailed description of proficiency levels can be found in *PISA 2009 Results: Students On Line* (OECD, 2011).



■ Figure 3.5 ■

Proficiency in digital reading

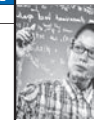
Percentage of students at the different levels of digital reading proficiency



Countries and economies are ranked in descending order of the percentage of students at or above Level 2 in digital reading.
Source: OECD, PISA 2012 Database, Table 3.3.
StatLink http://dx.doi.org/10.1787/888933252935

Top performers in digital reading

Students proficient at Level 5 or above are skilled online readers. Top performers in digital reading are able to evaluate information from several sources, assessing the credibility and utility of what they read using criteria that they have generated themselves. They are also able to solve tasks that require the reader to locate information, related to an unfamiliar context, in the presence of ambiguity and without explicit directions. In short, they are able to navigate autonomously and efficiently. Critical evaluation and expertise in locating relevant information are the key skills in online reading, given the virtually unlimited number of texts that can be accessed on line, and the variation in their credibility and trustworthiness. Students performing at Level 5 or above are able to deal with more technical material as well as with more popular and idiomatic texts.



They notice fine distinctions in the detail of the text, allowing them to draw inferences and form plausible hypotheses.

Across the 23 OECD countries that participated in the digital reading assessment in 2012, 8% of students performed at this level and can be considered top performers in digital reading. In Singapore, more than one in four students (27%) perform at Level 5 or above. So do about one in five students in Hong Kong-China (21%) and Korea (18%).

In general, a ranking of countries and economies by the proportion of top-performing students (students at Level 5 or above) matches the ranking of countries/economies by mean performance, but there are a number of exceptions. Mean performance in Israel is below the OECD average, but the share of top-performing students in Israel is similar to the share found across the OECD on average. By contrast, students in Macao-China perform above students in Belgium, Italy and Norway, but these countries all have larger proportions of top-performing students than Macao-China.

Low performers in digital reading

At the lower end of the scale, students performing below Level 2 are able to complete only the easiest digital reading tasks in the PISA 2012 assessment, if any. They have difficulties using conventional navigation tools and features, and locating links or information that are not prominently placed. Some of these students can scroll and navigate across web pages, and can locate simple pieces of information in a short text, if given explicit directions. These students are referred to as low performers in digital reading because they perform at levels that are not likely to allow them full access to education, employment and social opportunities afforded by digital devices.

Some 18% of students are considered low performers in digital reading, on average across the 23 participating OECD countries. In partner countries Colombia and the United Arab Emirates, more than half of all 15-year-old students perform at this low level. Large proportions of low-performing students are also found in Brazil (37%), Hungary (32%), Israel (31%), Chile (29%) and Spain (26%). By contrast, less than 5% of students perform below Level 2 in Japan, Korea and Singapore. These countries are close to ensuring that all students have the basic knowledge and skills required to access and use information that can be found on the Internet.

Progressions in digital reading proficiency

As students progress from the lower levels of proficiency to ever greater skill in digital reading, they become more autonomous in their navigation and better able to deal with a range of online text formats and text types, including unfamiliar ones. At Level 2 on the digital reading scale, students can successfully follow explicit instructions to locate information on line, form generalisations, such as recognising the intended audience of a website, and use typical online order forms that include drop-down menus or open text fields. At Level 3, in addition to mastering Level 2 tasks, students can cope with more complex digital reading tasks, including tasks that require integrating information from across different websites. At Level 4, students complete even more challenging tasks: they can assess the authority and relevance of sources when provided with support, and can explain the criteria on which their judgements are based.

They can also synthesise information across several sites (as is required, for instance, in the sample unit *SERAING*, Task 3: see Chapter 7) and understand texts written in technical language.

Box 3.1. The International Computer and Information Literacy Study (2013) and its relation to digital reading in PISA

In 2013, 21 education systems around the world participated in the first International Computer and Information Literacy Study (ICILS), organised by the International Association for the Evaluation of Educational Achievement (IEA). Computer and information literacy is defined as "an individual's ability to use computers to investigate, create and communicate in order to participate effectively at home, at school, in the workplace and in society". The framework highlights two strands of digital competence: "collecting and managing information", which also involves locating and evaluating information, and "producing and exchanging information", of which an understanding of online safety and security issues are part.

While some aspects highlighted by the PISA digital reading framework are covered, in particular, by the first strand of the ICILS framework, the concept of computer and information literacy is clearly distinct from digital reading.

The test was administered to eighth-grade students. Among the 12 countries that met the sampling requirements for ICILS, the Czech Republic obtained the highest mean score, followed by a group of four countries (Australia, Korea, Norway [grade 9] and Poland) with similar mean scores. While the target population differs, it is notable that the mean performance of Poland was clearly above that of countries, such as the Russian Federation, the Slovak Republic and Slovenia, whose mean scores in the PISA digital reading assessment was similar.

Source: Fraillon et al., 2014.

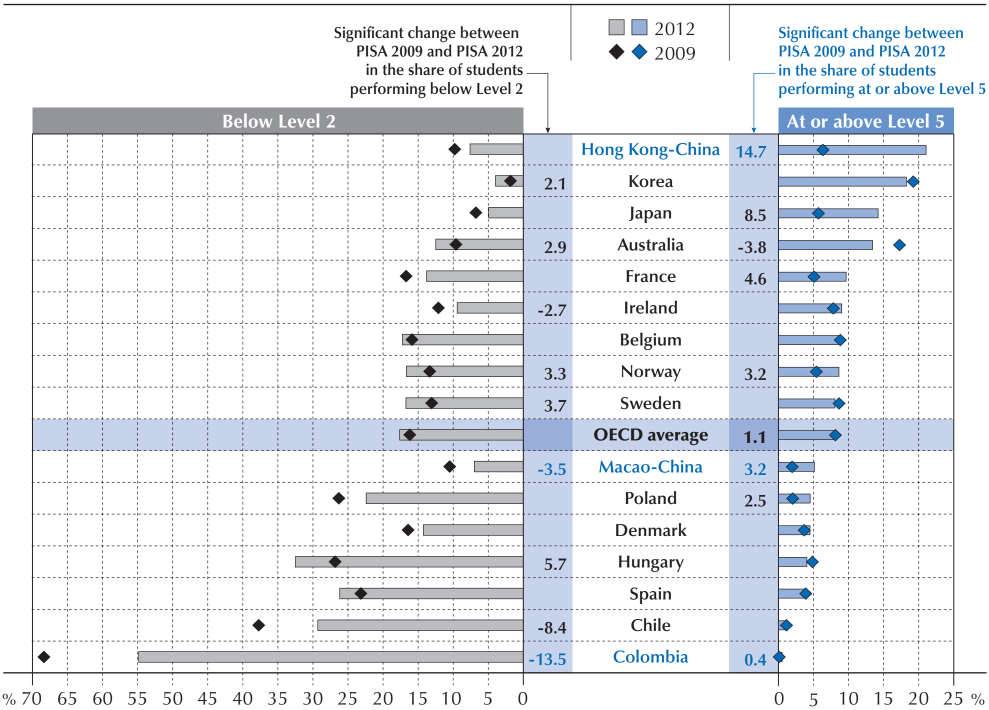
Trends at the top and bottom of the performance distribution in digital reading

Changes in a country's/economy's average performance can result from improvements or deterioration in performance at different points in the performance distribution. Trends in the proportion of low- and top-performing students indicate, in particular, what students can do better in 2012 than in 2009 (Figure 3.6).

Between 2009 and 2012, two countries, Chile and Colombia, significantly reduced the share of students performing below Level 2 in digital reading. Both countries still have large proportions of students performing at the lowest levels, but they were able to reduce underperformance significantly within only three years. The fact that no reduction was observed in these countries in the share of low achievers in print reading suggests that improvements in digital reading performance are related to improved ICT skills and better dispositions towards the use of computers among students. In the past, lack of familiarity with ICT tools and online text formats may have been a major obstacle for some students to complete even the easiest digital reading tasks.

■ Figure 3.6 ■

Percentage of low achievers and top performers in digital reading in 2009 and 2012



Notes: This figure includes only countries/economies that participated in both the PISA 2009 and PISA 2012 assessments of digital reading.
Changes that are statistically significant are reported next to the country/economy name.
For the OECD average, the diamonds denote all OECD countries that participated in the PISA 2009 assessment, the bars denote all OECD countries that participated in the PISA 2012 assessment, and the reported change applies only to OECD countries that participated in both assessments of digital reading.
Countries and economies are ranked in descending order of the percentage of students at or above Level 5 in digital reading in 2012.
Source: OECD, PISA 2012 Database, Table 3.4.
StatLink http://dx.doi.org/10.1787/888933252943

In Hong Kong-China and Japan, the share of top-performing students increased significantly between 2009 and 2012. In both, a similar, though smaller, increase in the share of top performers was observed in print reading as well (OECD, 2014b, Table I.4.1b). This may indicate that Hong Kong-China and Japan achieved gains at higher levels of proficiency by improving students' ability to master difficult reading tasks across both printed and online texts.

It is also possible to assess whether these changes in performance occurred among the countries'/economies' strongest or weakest students by looking at trends in percentiles. Eight countries/economies improved their average performance between 2009 and 2012. In Chile, improvements were largest among the lowest-performing students. By contrast, Colombia, Hong Kong-China and Japan were able to raise performance in digital reading mainly among their best-performing students. France, Ireland, Macao-China and Poland showed similar improvements among students at the top and bottom of the performance distribution (Table 3.5).

Among countries with deteriorating performance in digital reading, Hungary, Korea and Sweden show the biggest declines in performance among their weakest students. In Australia, performance declined to a similar extent across the distribution (Table 3.5).

Four countries, namely Belgium, Denmark, Norway and Spain, showed stable mean performance. However, in Norway, the lack of change in mean performance masks a significant widening of performance differences, with the lowest-achieving students performing even lower, and the highest-achieving students even higher, in 2012 compared to 2009 (Table 3.5).

DIFFERENCES IN PERFORMANCE BETWEEN PRINT AND DIGITAL READING

Overall, the correlation between the digital and print reading performance of students is 0.81, about the same correlation as observed between digital reading and (paper-based) mathematics scores (0.78) (Table 3.9).[3]

While, in general, strong readers will perform well both in print and digital reading, there is significant variation in digital reading performance at all levels of performance in print reading. The variation in digital reading performance that is not explained by differences in print reading skills is referred to as residual variation. Some of this residual variation contributes to differences in performance observed across countries/economies. It is then referred to as the relative performance of countries/economies in digital reading (Figure 3.7). This relative performance may be related to skills that are used, to a greater extent, when reading on line (see Chapter 4). It may also be related to students' dispositions towards the medium and the variation in students' familiarity with basic ICT skills, such as operating a mouse and keyboard to use hyperlinks, browser buttons, drop-down menus and text-entry fields.

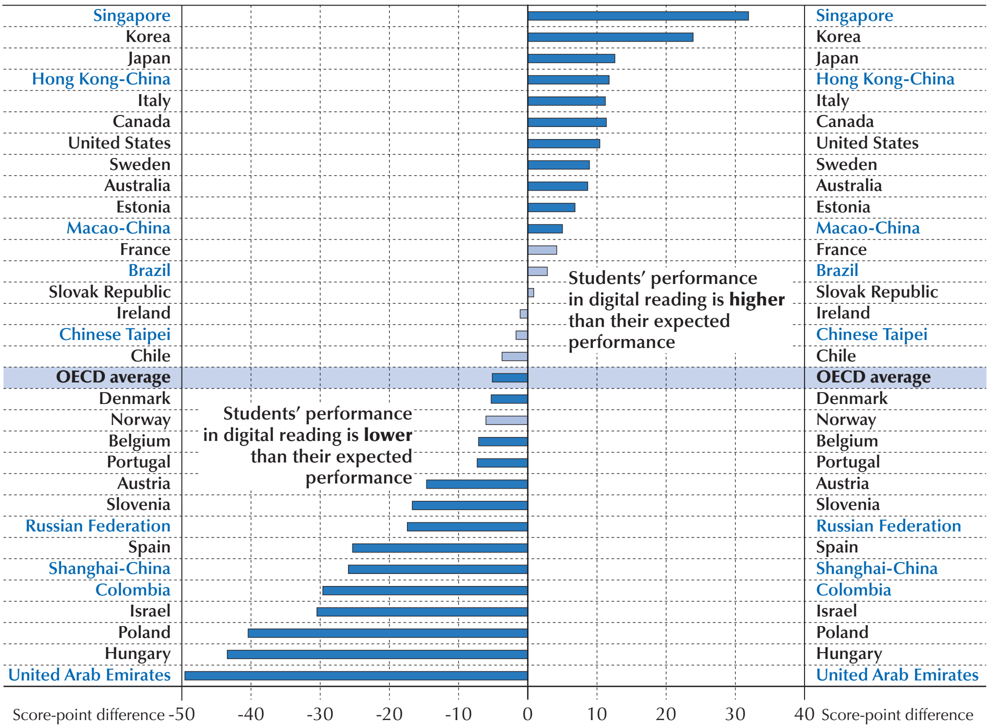
In 11 countries and economies, students perform significantly better in digital reading, on average, than students in other countries with similar skills in print reading. A large, positive difference in digital reading performance, after accounting for print reading performance, is observed in Singapore (32 score points) and Korea (24 score points). Students in Australia, Canada, Estonia, Hong Kong-China, Italy, Japan, Macao-China, Sweden and the United States also perform better than would be expected, based on their performance in print reading (Figure 3.7).

In 15 countries and economies, students perform below par in digital reading, on average, when compared to students in other participating countries and economies who display the same level of proficiency in print reading. Large gaps in relative performance in digital reading are found in the United Arab Emirates (50 score points), Hungary (43 score points), Poland (40 score points), Israel (31 score points), Colombia (30 score points), Shanghai-China (26 score points) and Spain (25 score points). Students in Austria, Belgium, Denmark, the Russian Federation, Portugal and Slovenia also perform worse in digital reading, on average, than would be expected, based on their performance in print reading (Figure 3.7).

■ Figure 3.7 ■

Relative performance in digital reading

Score-point difference between actual and expected performance

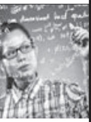


Notes: Statistically significant differences are shown in a darker tone.
Each student's expected performance is estimated, using a regression model, as the predicted performance in digital reading given his or her score in print reading.
Countries and economies are ranked in descending order of the score-point difference between actual and expected performance.
Source: OECD, PISA 2012 Database, Table 3.6.
StatLink http://dx.doi.org/10.1787/888933252959

Top performers in digital and print reading

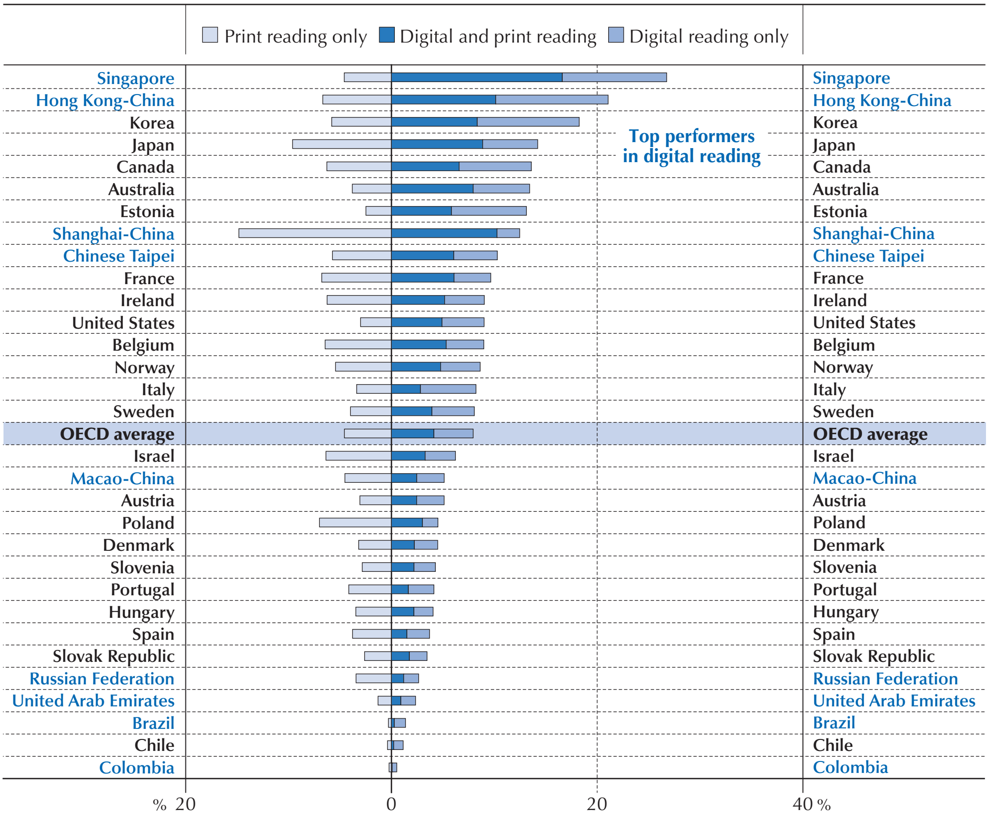
Figure 3.8 shows the proportion of top performers across countries and economies participating in the digital reading assessment, highlighting the extent to which those students who demonstrate high proficiency in print reading can perform at similar levels in digital reading as well. On average across OECD countries, 8% of students perform at Level 5 or above in digital reading. Of these, about half (4%) also perform at this level in print reading.

Conversely, in many countries and economies, about half of the top performers in print reading also perform at the top in digital reading. In Australia, Estonia and Singapore, more than two in three top performers in print reading also perform at the top in digital reading. In these countries, good readers usually are able to perform at similar levels regardless of the medium.



■ Figure 3.8 ■

Overlapping of top performers in digital and print reading



Countries and economies are ranked in descending order of the percentage of top performers in digital reading.
Source: OECD, PISA 2012 Database, Table 3.7.
StatLink http://dx.doi.org/10.1787/888933252962

In Poland, however, fewer than one in three top performers in print reading also performs at the top in digital reading. This may indicate that, more often than in other countries, in Poland, good readers of print documents lack the evaluation and navigation skills that would make them skilled online readers.

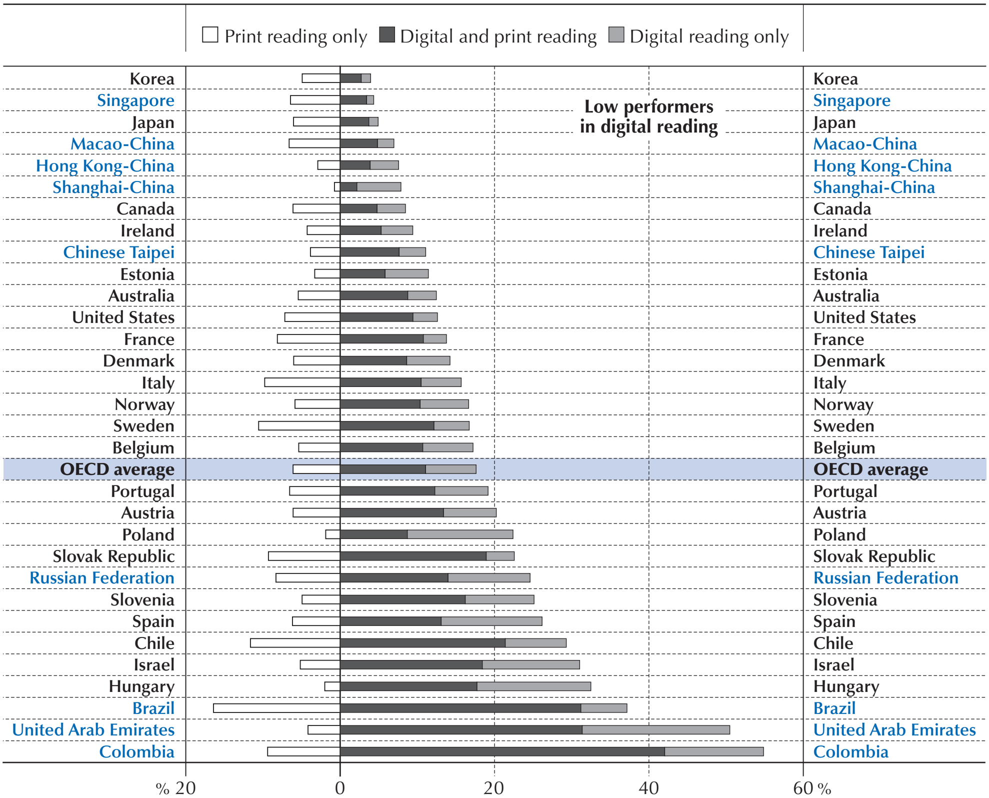
Low performers in digital and print reading

Figure 3.9 shows the proportion of low performers across OECD countries, highlighting the extent to which low-performing students in digital reading also encounter difficulties when reading print documents. In general, there is a greater overlap among low-performers than among top-performers across the two media.

At the same time, several countries and economies have significant proportions of student who, despite being able to read at Level 2 or above when assessed on paper, perform below Level 2 when assessed on computer. In Colombia, Hungary, Israel, Poland, the Russian Federation, Spain and the United Arab Emirates, more than one in ten students is a low performer in digital reading but not in print reading (Figure 3.9). In these countries, many students may have difficulties with the generic ICT skills and conventions required to interact with the test platform, and thus perform poorly in digital reading despite their relatively good reading skills.

■ Figure 3.9 ■

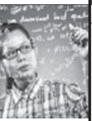
Overlapping of low performers in digital and print reading



Countries and economies are ranked in ascending order of the percentage of low performers in digital reading.
Source: OECD, PISA 2012 Database, Table 3.7.
StatLink http://dx.doi.org/10.1787/888933252976

STUDENT PERFORMANCE IN THE COMPUTER-BASED ASSESSMENT OF MATHEMATICS

Mathematics was the focus of the PISA 2012 assessment, meaning that booklets in the paper-based test contained questions measuring students' capacity to formulate, employ and interpret mathematics to a larger extent than questions for reading and science, the other domains assessed on paper. For the first time, mathematics was also assessed on computers in 2012. This section reports the results from the computer-based assessment of mathematics.



Average performance in the computer-based assessment of mathematics

The same 32 countries/economies that participated in the optional assessment of digital reading also participated in the computer-based assessment of mathematics. The scale for reporting performance on the computer-based test of mathematics was fixed so that the average mean score and standard deviation for OECD countries would match those of the paper-based mathematics scale for the same year and the same countries.

■ Figure 3.10 ■

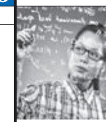
Comparing countries' and economies' performance in the computer-based assessment of mathematics

Statistically significantly **above** the OECD average
Not statistically significantly different from the OECD average
Statistically significantly **below** the OECD average

Mean score	Comparison country/economy	Countries/economies whose mean score is NOT statistically significantly different from that of the comparison country/economy
566	Singapore	Shanghai-China
562	Shanghai-China	Singapore, Korea
553	Korea	Shanghai-China, Hong Kong-China
550	Hong Kong-China	Korea, Macao-China
543	Macao-China	Hong Kong-China, Japan, Chinese Taipei
539	Japan	Macao-China, Chinese Taipei
537	Chinese Taipei	Macao-China, Japan
523	Canada	
516	Estonia	Belgium
512	Belgium	Estonia, France, Australia, Austria
508	France	Belgium, Australia, Austria, Italy, United States
508	Australia	Belgium, France, Austria
507	Austria	Belgium, France, Australia, Italy, United States
499	Italy	France, Austria, United States, Norway, Slovak Republic, Denmark, Ireland, Sweden, Russian Federation, Poland, Portugal
498	United States	France, Austria, Italy, Norway, Slovak Republic, Denmark, Ireland, Sweden, Russian Federation, Poland, Portugal
498	Norway	Italy, United States, Slovak Republic, Denmark, Ireland, Sweden, Poland
497	Slovak Republic	Italy, United States, Norway, Denmark, Ireland, Sweden, Russian Federation, Poland, Portugal
496	Denmark	Italy, United States, Norway, Slovak Republic, Ireland, Sweden, Russian Federation, Poland, Portugal
493	Ireland	Italy, United States, Norway, Slovak Republic, Denmark, Sweden, Russian Federation, Poland, Portugal
490	Sweden	Italy, United States, Norway, Slovak Republic, Denmark, Ireland, Russian Federation, Poland, Portugal, Slovenia
489	Russian Federation	Italy, United States, Slovak Republic, Denmark, Ireland, Sweden, Poland, Portugal, Slovenia
489	Poland	Italy, United States, Norway, Slovak Republic, Denmark, Ireland, Sweden, Russian Federation, Portugal, Slovenia
489	Portugal	Italy, United States, Slovak Republic, Denmark, Ireland, Sweden, Russian Federation, Poland, Slovenia
487	Slovenia	Sweden, Russian Federation, Poland, Portugal
475	Spain	Hungary
470	Hungary	Spain
447	Israel	
434	United Arab Emirates	Chile
432	Chile	United Arab Emirates
421	Brazil	
397	Colombia	

Source: OECD, PISA 2012 Database.
StatLink http://dx.doi.org/10.1787/888933252985

Figure 3.10 lists each participating country and economy in descending order of its mean score in the computer-based mathematics test (left column). The values range from a high of 566 points for partner country Singapore to a low of 397 points for partner country Colombia. Shanghai-China (562 points) performs at the same level as Singapore. Students in Korea, Hong Kong-China,



Macao-China, Japan and Chinese Taipei (in descending order of mean performance) score lower than students in Singapore, on average, but significantly higher than the mean performance of students in any other country/economy participating in the assessment.

Differences between countries' mean scores on the computer-based and paper-based mathematics assessment are smaller than those observed between the digital and print-reading assessments. Indeed, the correlation between students' results on the paper- and computer-based mathematics scale is higher (0.86) than the correlation between the digital and print-reading scores (0.81), when considering the pooled sample of students from all participating countries (Table 3.9). Table 3.10 reports differences in mean scores between the computer-based and the paper-based assessment of mathematics, by country.

Figure 3.11 shows where each country/economy ranks in its mean performance in the computer-based mathematics test. A range of ranks is presented to reflect the uncertainty associated with this estimate.

■ Figure 3.11 ■

Where countries and economies rank in computer-based mathematics performance

	Computer-based mathematics scale					
			Range of ranks			
			OECD countries		All countries/economies	
	Mean score	S.E.	Upper rank	Lower rank	Upper rank	Lower rank
Singapore	566	(1.3)			1	2
Shanghai-China	562	(3.4)			1	2
Korea	553	(4.5)	1	1	2	4
Hong Kong-China	550	(3.4)			3	4
Macao-China	543	(1.1)			5	6
Japan	539	(3.3)	2	2	5	7
Chinese Taipei	537	(2.8)			6	7
Canada	523	(2.2)	3	3	8	8
Estonia	516	(2.2)	4	5	9	10
Belgium	512	(2.5)	4	7	9	12
France	508	(3.3)	5	9	10	14
Australia	508	(1.6)	6	8	11	13
Austria	507	(3.5)	5	9	10	14
Italy	499	(4.2)	8	15	13	20
United States	498	(4.1)	8	15	13	20
Norway	498	(2.8)	9	14	14	19
Slovak Republic	497	(3.5)	9	15	13	20
Denmark	496	(2.7)	9	15	14	20
Ireland	493	(2.9)	11	17	15	22
Sweden	490	(2.9)	13	18	18	24
Russian Federation	489	(2.6)			19	24
Poland	489	(4.0)	12	18	17	24
Portugal	489	(3.1)	13	18	18	24
Slovenia	487	(1.2)	16	18	21	24
Spain	475	(3.2)	19	20	25	26
Hungary	470	(3.9)	19	20	25	26
Israel	447	(5.6)	21	21	27	27
United Arab Emirates	434	(2.2)			28	29
Chile	432	(3.3)	22	22	28	29
Brazil	421	(4.7)			30	30
Colombia	397	(3.2)			31	31

Source: OECD, PISA 2012 Database.
StatLink http://dx.doi.org/10.1787/888933252992

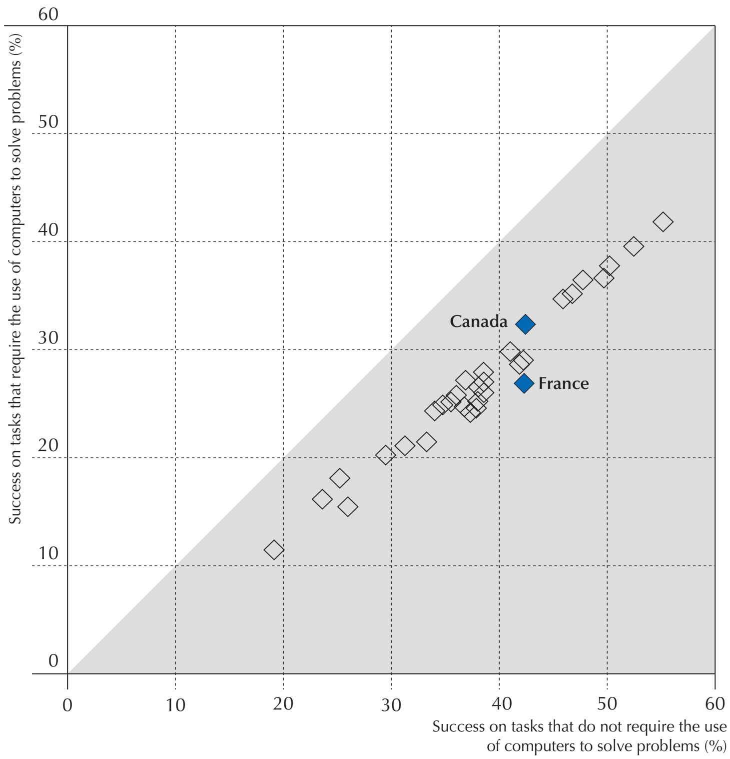
DIFFERENCES IN PERFORMANCE RELATED TO THE USE OF ICT TOOLS FOR SOLVING MATHEMATICS PROBLEMS

Computers provide a range of opportunities for developing tests that are more interactive, authentic and engaging (Stacey and Wiliam, 2012); they are also increasingly used in the workplace and in everyday life to deal with problems involving numbers, quantities, two or three-dimensional figures, and data. While the assessment framework for the PISA computer-based mathematics assessment is the same as for the paper-based test, some of the computer-based tasks could not exist in a paper test because of their response format (e.g. “drag and drop”), or because they require students to use the computer as a mathematical tool, by interacting with the stimulus to solve a mathematics problem.

■ Figure 3.12 ■

Success on mathematics tasks that require/do not require the use of computers to solve problems

Average percentage of full-credit responses across countries and economies



Notes: Each diamond represents the mean values of a country/economy.
In the computer-based assessment of mathematics, Canada and France share similar levels of performance on tasks that do not require the use of computers to solve mathematics problems, but differ in their students’ performance on tasks that do require such use; this example is discussed in the text.
Source: OECD, PISA 2012 Database, Table 3.11.
StatLink http://dx.doi.org/10.1787/888933253005

Such tasks require students to build and rotate a three-dimensional figure using a mouse, to find out how the graphical representation of a function changes depending on its parameters, to use an on-screen calculator, to sort a data set, or to produce a chart from data.

By design, not all computer-based tasks involved the use of the computer as a mathematical tool. This variation makes it possible to analyse the impact of these kinds of demands on performance. While task formats that involve the use of computers as mathematical tools may appear more engaging, not all students may react similarly to them. These types of tasks also typically require greater familiarity with computers and their application to mathematics.

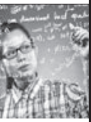
Figure 3.12 plots average success rates for tasks that require the use of computers to solve mathematics problems against average success rates for more traditional mathematics tasks.[4] While both types of tasks were presented on screen, only in the former set of tasks did the solution require the use of computer tools, or was made easier if the computer was used as a tool. Tasks in unit *CD PRODUCTION*, for instance, require students to use an on-screen calculator. Tasks in units *STAR POINTS* and *BODY MASS INDEX*, in contrast, are examples of "traditional" items. The fact that students use keyboard and mouse, instead of pens and pencils, to answer these items does not make them easier than their corresponding paper versions would be.[5]

In general, country rankings are similar across the two types of tasks. However, as Figure 3.12 shows, performance is not perfectly aligned. Countries that share similar levels of success on tasks that do not require the use of computers to solve problems do not necessarily perform similarly on tasks that require students to use mathematics-specific ICT tools in order to solve the task. Often, when considering two countries with similar performance on the first set of tasks, one country is significantly stronger than the other on the second set of tasks.

Students in Canada, for instance, have similar success rates as students in France on tasks where the use of computers as tools for solving mathematics problems is not required. In both countries, students answer around 42% of these tasks correctly. Students in Canada, however, have significantly greater success than students in France (32% vs. 27%) on tasks where the solution is only possible, or is made easier, by using computers as mathematical tools.

Figure 3.13 ranks countries and economies according to whether their students had greater success on tasks that require the use of computers to solve problems, or on the remaining tasks, relative to their overall success. This analysis accounts for differences in the difficulty of tasks across the two sets by comparing success on both types of tasks in each country/economy to the average success rate across OECD countries.

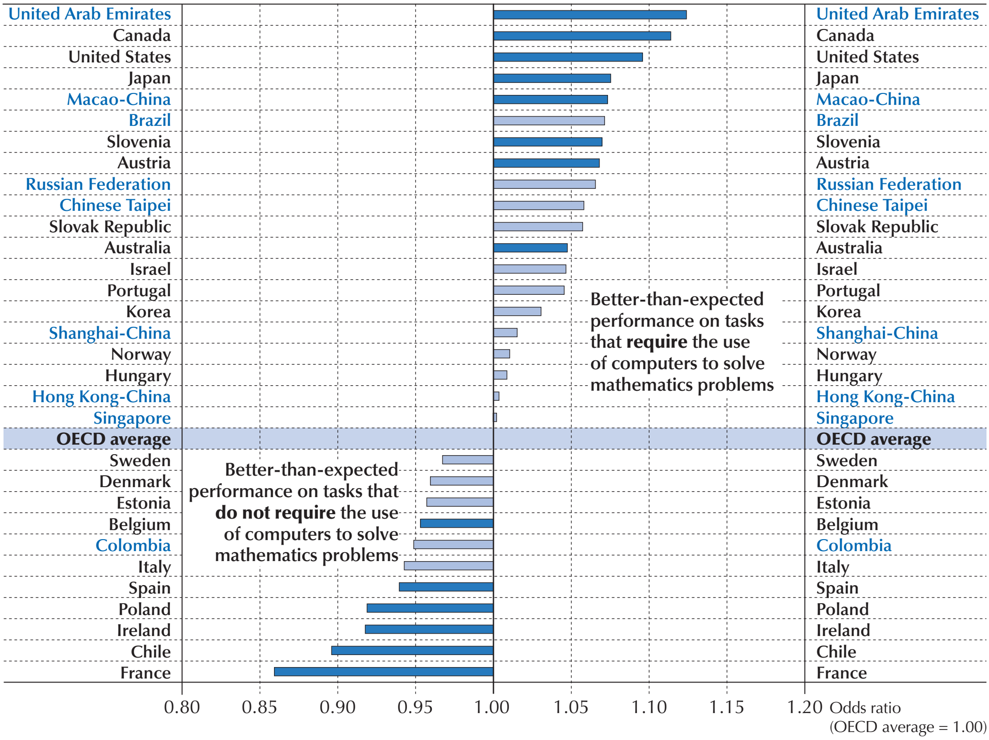
According to these adjusted figures, students in Australia, Austria, Canada, Japan, Slovenia and the United States as well as those in partner countries/economies Macao-China and the United Arab Emirates perform better on tasks that require the use of computers to solve problems, compared to their success on traditional tasks. By contrast, relative success is only 0.86 in France (significantly below par), indicating weaker-than-expected performance when students are confronted with tasks that require the use of computer-based tools to arrive at the solution. Students in Belgium, Chile, Ireland, Poland and Spain also perform worse than expected on such tasks.



■ Figure 3.13 ■

Relative success on mathematics tasks that require the use of computers to solve problems

Compared to the OECD average



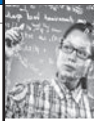
Notes: Values that are statistically significant are marked in a darker tone.

This figure shows that students in Canada are 1.11 times more likely than students across OECD countries, on average, to succeed on tasks in the computer-based mathematics assessment that require the use of computers to solve problems, given their success on other tasks in the assessment.

Countries and economies are ranked in descending order of their relative success on tasks involving the use of computers to solve problems.

Source: OECD, PISA 2012 Database, Table 3.11.

StatLink http://dx.doi.org/10.1787/888933253011



Notes

1. Germany participated in the assessments of digital reading and computer-based mathematics as a research project. Results for Germany are not reported.

2. Although Austria did participate in both assessments, the comparability of the 2009 data with data from PISA 2012 cannot be assured. A negative atmosphere surrounding educational assessment affected the conditions under which the assessment was administered in 2009, and could have adversely affected student motivation to respond to the PISA tasks.

3. Both figures refer to the latent correlation in the pooled sample of students from all countries/economies participating in computer-based assessments. Student observations are weighted with final student weights.

4. Some of the items classified as "traditional", because they do not require the use of computers to solve problems, may, however, have response formats that are only possible on screen, such as drag and drop, or may involve animated stimulus information. This classification is therefore meant to capture the difference in item demands, rather than a difference merely related to item presentation.

5. The examples refer to released computer-based mathematics items, which can be tried out on the website of the Australian Council for Educational Research (http://cbasq.acer.edu.au/index.php?cmd=toMaths).

Chapter 3 tables are available on line at http://dx.doi.org/10.1787/edu-data-en.

Note regarding Israel

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

References

Fraillon, J., J. Ainley, W. Schulz, T. Friedman and **E. Gebhardt** (2014), "Executive summary", in Fraillon, J., J. Ainley, W. Schulz, T. Friedman and E. Gebhardt (eds.), *Preparing for Life in a Digital Age*, pp. 15-25, Springer International Publishing.

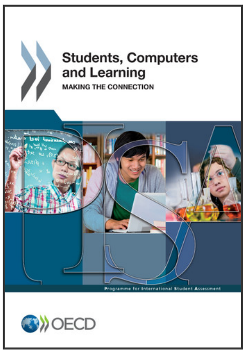
OECD (2014a), *PISA 2012 Technical Report*, PISA, OECD, Paris, www.oecd.org/pisa/pisaproducts/pisa2012technicalreport.htm.

OECD (2014b), *PISA 2012 Results: What Students Know and Can Do (Volume I, Revised edition, February 2014): Student Performance in Mathematics, Reading and Science*, PISA, OECD Publishing, Paris, http://dx.doi.org/10.1787/9789264208780-en.

OECD (2013), *PISA 2012 Assessment and Analytical Framework: Mathematics, Reading, Science, Problem Solving and Financial Literacy*, PISA, OECD Publishing, Paris, http://dx.doi.org/10.1787/9789264190511-en.

OECD (2011), *PISA 2009 Results: Students On Line (Volume VI): Digital Technologies and Performance*, PISA, OECD Publishing, Paris, http://dx.doi.org/10.1787/9789264112995-en.

Stacey, K. and **D. Wiliam** (2012), "Technology and assessment in mathematics", in Clements, M.A.K., A.J. Bishop, C. Keitel, J. Kilpatrick and F.K.S. Leung (eds.), *Third International Handbook of Mathematics Education*, Springer International Handbooks of Education, Vol. 27, Springer New York, pp. 721-751.



From:
Students, Computers and Learning
Making the Connection

Access the complete publication at:
https://doi.org/10.1787/9789264239555-en

Please cite this chapter as:

OECD (2015), “Main Results from the PISA 2012 Computer-Based Assessments”, in *Students, Computers and Learning: Making the Connection*, OECD Publishing, Paris.

DOI: https://doi.org/10.1787/9789264239555-6-en